Bruce K. Driver

# 240C Supplements

May 16, 2018 File:Supplements.tex

## Contents

1	The Polar Decomposition of Lebesgue Measure	1
2	Metric-Measure Space Regularity Results         2.1       Metric space results         2.2       Regularity Results for Borel measures on $(X, \mathcal{B}_X)$ 2.3       Dual Considerations	$5 \\ 5 \\ 6 \\ 8$
3	Fourier Series.         3.1 Dirichlet Kernel.         3.2 Fejér Kernel.         3.3 The Dirichlet Problems on D and the Poisson Kernel.         3.4 Multi-Dimensional Fourier Series.         3.5 Translation Invariant Operators.	$     \begin{array}{r}       11 \\       12 \\       14 \\       15 \\       17 \\       18 \\       \end{array} $
4	Convolution and smoothing operators         4.1       Basic Properties of Convolutions         4.2       Young's Inequalities         4.3       Convolution smoothing	19 19 22 23
5	Fourier Transform $5.1$ Motivation $5.2$ Fourier Transform formal development $5.3$ Schwartz Test Functions $5.4$ Summary of Basic Properties of $\mathcal{F}$ and $\mathcal{F}^{-1}$	27 27 28 31 36
6	Constant Coefficient partial differential equations $6.1$ Elliptic examples $6.2$ Heat Equation on $\mathbb{R}^n$ $6.3$ Poisson Semi-Group	37 37 38 42

	<u>a</u>
4	Contents

	6.4Addendum: convolutions and Fourier Transforms involving measures6.5Wave Equation on $\mathbb{R}^n$	42 43
7	Radon Measures and the Dual of $C_0(X)$	49
	7.1 The Riesz-Markov Theorem	49
	7.2 Classifying Radon Measures on $\mathbb{R}$	51
	7.3 Classifying Radon Measures on $\mathbb{R}$ using Theorem 7.11	52
	7.4 Kolmogorov's Existence of Measure on Products Spaces	53
	7.5 The dual of $C_0(X)$	54
8	Homework #6 (Spring 2018)	57
9	Spectral Theorem (Compact Operator Case)	59
	9.1 Basics of Compact Operators	59
	9.2 Compact Operators on Hilbert spaces	61
	9.3 The Spectral Theorem for Self Adjoint Compact Operators	62
	9.4 Hilbert Schmidt Operators	66

## The Polar Decomposition of Lebesgue Measure

Let

$$S^{d-1} = \{ x \in \mathbb{R}^d : |x|^2 := \sum_{i=1}^d x_i^2 = 1 \} \text{ and}$$
$$\bar{B}' := \{ x \in \mathbb{R}^d : 0 < ||x|| \le 1 \}$$

be the unit sphere and 0-deleted "closed" ball in  $\mathbb{R}^d$  and let  $\mathcal{B}_{S^{d-1}}$  and  $\mathcal{B}_{\bar{B}'}$  be the Borel  $\sigma$ -algebras on these metric spaces. We further equip  $(0,\infty) \times S^{d-1}$ with product  $\sigma$ -algebra  $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$  which is also the Borel  $\sigma$ -algebra on  $(0,\infty) \times S^{d-1}$  thought of as a product of two metric spaces. The maps  $\Phi$ :  $\mathbb{R}^d \setminus \{0\} \to (0,\infty) \times S^{d-1}$  and  $\psi: \bar{B}' \to S^{d-1}$  defined by

$$\Phi(x) := (|x|, |x|^{-1} x) \text{ for all } x \in \mathbb{R}^d \setminus \{0\} \text{ and}$$
$$\psi(x) = |x|^{-1} x \text{ for all } x \in \overline{B}',$$

are both continuous and hence measurable. Similarly the inverse map,  $\Phi^{-1}$ :  $(0,\infty) \times S^{d-1} \to \mathbb{R}^d \setminus \{0\}$ , is given by  $\Phi^{-1}(r,\omega) = r\omega$  which is continuous and therefore also measurable.

For  $E \in \mathcal{B}_{S^{d-1}}$  and a > 0, let

$$E_a := \{ r\omega : r \in (0, a] \text{ and } \omega \in E \} = \Phi^{-1}((0, a] \times E) \in \mathcal{B}_{\mathbb{R}^d}.$$

Further observe that  $E_1 = \psi^{-1}(E) \in \mathcal{B}_{\overline{B}'} \subset \mathcal{B}_{\mathbb{R}^d}$  and for  $a > 0, E_a = aE_1$ .

**Definition 1.1.** The surface measure,  $\sigma$ , on  $S^{d-1}$  is defined to be  $\sigma = d \cdot (\psi_* m)$ , *i.e.* 

$$\sigma(E) := d \cdot m(E_1) \text{ for all } E \in \mathcal{B}_{S^{d-1}}.$$

Let us now explain the intuition behind Definition 1.1. If  $E \subset S^{d-1}$  is a set and  $\varepsilon > 0$  is a small number, then the volume of

$$(1, 1 + \varepsilon] \cdot E = \{ r\omega : r \in (1, 1 + \varepsilon] \text{ and } \omega \in E \}$$

should be approximately given by  $m((1, 1 + \varepsilon] \cdot E) \cong \sigma(E)\varepsilon$ , see Figure 1.1 below. On the other hand

$$m\left((1,1+\varepsilon]E\right) = m\left(E_{1+\varepsilon}\setminus E_1\right) = \left[(1+\varepsilon)^d - 1\right]\cdot m\left(E_1\right).$$



Fig. 1.1. Motivating the definition of surface measure for a sphere.

Therefore we expect the area of E should be given by

$$\sigma(E) = \lim_{\varepsilon \downarrow 0} \frac{\left\{ (1+\varepsilon)^d - 1 \right\}}{\varepsilon} m(E_1) = d \cdot m(E_1)$$

The following theorem is an abstract version of integration in polar coordinates.

**Theorem 1.2 (Polar decomposition of** m). Let  $\rho_d$  be the measure on  $\mathcal{B}_{(0,\infty)}$  defined by  $d\rho_d(r) = r^{d-1}dr$ , *i.e.* 

$$\rho(J) = \int_{J} r^{d-1} dr \ \forall \ J \in \mathcal{B}_{(0,\infty)}, \tag{1.1}$$

Then  $\Phi_*m = \rho \otimes \sigma$  on  $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$ .

**Proof.** Let  $\mathcal{E}$  be the  $\pi$ -system in  $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$  consisting of sets of the form  $A = (a, b] \times E \in \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$  with  $0 < a < b < \infty$  and  $E \in \mathcal{B}_{S^{d-1}}$ . For such an  $A \in \mathcal{E}$  we have

$$\Phi^{-1}(A) = \{ r\omega : r \in (a, b] \text{ and } \omega \in E \} = E_b \setminus E_a = bE_1 \setminus aE_1.$$

#### 2 1 The Polar Decomposition of Lebesgue Measure

Therefore by the basic scaling properties of m and the fundamental theorem of calculus,

$$\Phi_*m)(A) = m(bE_1 \setminus aE_1) = m(bE_1) - m(aE_1) = b^d m(E_1) - a^d m(E_1) = d \cdot m(E_1) \int_a^b r^{d-1} dr \qquad (1.2) = \rho((a,b]) \sigma(E) = (\rho \otimes \sigma)(A) \qquad (1.3)$$

Since  $(\Phi_*m)(A) = (\rho \otimes \sigma)(A)$  for all  $A \in \mathcal{E}$ , we may apply the multiplicative system theorem<sup>1</sup> in the form of Proposition ?? to conclude that  $\Phi_*m = \rho \otimes \sigma$  on  $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{d-1}}$ .

**Corollary 1.3 (Polar Coordinates).** If  $f : \mathbb{R}^d \to [0,\infty]$  is a  $(\mathcal{B}_{\mathbb{R}^d}, \mathcal{B})$ -measurable function then

$$\int_{\mathbb{R}^d} f(x) dm(x) = \int_{(0,\infty) \times S^{d-1}} f(r\omega) r^{d-1} \, dr d\sigma(\omega) \,. \tag{1.4}$$

In particular if  $f : \mathbb{R}_+ \to \mathbb{R}_+$  is measurable then

$$\int_{\mathbb{R}^d} f(|x|) dx = \sigma \left( S^{d-1} \right) \int_0^\infty f(r) r^{d-1} dr = \int_0^\infty f(r) dV(r)$$
(1.5)

where

$$V(r) = m(B(0,r)) = r^d m(B(0,1)) = d^{-1} r^d \sigma(S^{d-1}).$$

[In Example ??, Exercise ??, and Proposition ?? below, we will use the general change of variables Theorem ?? to give a explicit description for the surface integrals relative to  $\sigma$ .]

**Proof.** Equation (1.4) is a direct consequence of the abstract change of variables theorem (Exercise ??), Theorem 1.2, and Tonelli's Theorem ??. Indeed we have,

$$\begin{split} \int f dm &= \int_{\mathbb{R}^d \setminus \{0\}} \left( f \circ \Phi^{-1} \right) \circ \Phi \ dm = \int_{(0,\infty) \times S^{d-1}} \left( f \circ \Phi^{-1} \right) \ d\left( \Phi_* m \right) \\ &= \int_{(0,\infty) \times S^{d-1}} \left( f \circ \Phi^{-1} \right) d\left[ \rho \otimes \sigma \right] = \int_{(0,\infty) \times S^{d-1}} f(r\omega) \rho\left( dr \right) \sigma\left( d\omega \right) \\ &= \int_{(0,\infty) \times S^{d-1}} f(r\omega) r^{d-1} \ dr d\sigma\left( \omega \right). \end{split}$$

Equation (1.5) is a special case of Eq. (1.4).

<sup>1</sup> Or you could use the  $\pi$ - $\lambda$  theorem.

Example 1.4 ( $\sigma(S^1) = 2\pi$ ). Let  $E = \{(\cos \theta, \sin \theta) \in S^1 \subset \mathbb{R}^2 : 0 \le \theta \le \pi\}$ , then  $E_1$  is the upper half of closed unit disk centered at 0 in  $\mathbb{R}^2$ . Therefore,

$$m^{2}(E_{1}) = \int 1_{E_{1}} dm = \int_{-1}^{1} dx \int_{0}^{\sqrt{1-x^{2}}} dy = \int_{-1}^{1} \sqrt{1-x^{2}} dx.$$

Letting  $x = \sin \theta$  we find,

$$\sigma(E) = 2m^2(E_1) = 2\int_{-\pi/2}^{\pi/2} \cos\theta \cdot \cos\theta \, d\theta$$
$$= 2\frac{1}{2}\int_{-\pi/2}^{\pi/2} [1 + \cos 2\theta] \, d\theta = \pi.$$

Therefore  $\sigma(S^1) = 2\sigma(E) = 2\pi$  – the circumference of a circle of radius 1 as to be expected.

**Lemma 1.5.** If a > 0 and  $d \in \mathbb{N}$ , then

$$I_{d}(a) := \int_{\mathbb{R}^{d}} e^{-a|x|^{2}} dm(x) = (\pi/a)^{d/2}.$$

**Proof.** Using Tonelli's theorem and induction,

$$I_{d}(a) = \int_{\mathbb{R}^{d-1} \times \mathbb{R}} e^{-a|y|^{2}} e^{-at^{2}} m_{d-1}(dy) dt$$
  
=  $I_{d-1}(a) I_{1}(a) = I_{1}^{d}(a).$  (1.6)

So it suffices to compute:

$$I_{2}(a) = \int_{\mathbb{R}^{2}} e^{-a|x|^{2}} dm(x) = \int_{\mathbb{R}^{2} \setminus \{0\}} e^{-a(x_{1}^{2} + x_{2}^{2})} dx_{1} dx_{2}.$$

Using polar coordinates, see Eq. (1.4), we find,

$$I_{2}(a) = \int_{(0,\infty)\times S^{1}} e^{-a|r\omega|^{2}} r \, dr d\sigma(\omega) = \sigma\left(S^{1}\right) \cdot \int_{0}^{\infty} r e^{-ar^{2}} dr$$
$$= 2\pi \lim_{M \to \infty} \int_{0}^{M} r e^{-ar^{2}} dr = 2\pi \lim_{M \to \infty} \frac{e^{-ar^{2}}}{-2a} \int_{0}^{M} = \frac{2\pi}{2a} = \pi/a.$$

This shows that  $I_2(a) = \pi/a$  and the result now follows from Eq. (1.6).

**Corollary 1.6.** The surface area  $\sigma(S^{d-1})$  of the unit sphere  $S^{d-1} \subset \mathbb{R}^d$  is

$$\sigma(S^{d-1}) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \tag{1.7}$$

where  $\Gamma$  is the gamma function is as in Example ?? and ??.

**Proof.** Using Corollary 1.3 we find

$$I_d(1) = \int_0^\infty dr \ r^{d-1} e^{-r^2} \int_{S^{d-1}} d\sigma = \sigma(S^{d-1}) \int_0^\infty r^{d-1} e^{-r^2} dr.$$

Making the making the change of variables  $u=r^2$  so that  $r=u^{1/2}$  and  $dr=\frac{1}{2}u^{-1/2}du$  we find

$$\frac{I_d(1)}{\sigma(S^{d-1})} = \int_0^\infty u^{\frac{d-1}{2}} e^{-u} \frac{1}{2} u^{-1/2} du = \frac{1}{2} \int_0^\infty u^{\frac{d}{2}-1} e^{-u} du = \frac{1}{2} \Gamma(d/2).$$
(1.8)

Solving this equation for  $\sigma(S^{d-1})$  while making use of Lemma 1.5 gives Eq. (1.7).

**Exercise 1.1 (Folland Problem 2.62 on p. 80. ).** Rotation invariance of surface measure on  $S^{n-1}$ .

**Exercise 1.2 (Folland Problem 2.64 on p. 80. ).** On the integrability of  $|x|^{a} |\log |x||^{b}$  for x near 0 and x near  $\infty$  in  $\mathbb{R}^{n}$ .

Exercise 1.3. Show, using Problem 1.1 that

$$\int_{S^{d-1}} \omega_i \omega_j d\sigma\left(\omega\right) = \frac{1}{d} \delta_{ij} \sigma\left(S^{d-1}\right).$$

**Hint:** show  $\int_{S^{d-1}} \omega_i^2 d\sigma(\omega)$  is independent of *i* and therefore

$$\int_{S^{d-1}} \omega_i^2 d\sigma\left(\omega\right) = \frac{1}{d} \sum_{j=1}^d \int_{S^{d-1}} \omega_j^2 d\sigma\left(\omega\right).$$

**Proposition 1.7.** Let  $d \in \mathbb{N}$ ,

$$\mathbb{R}^d_+ := \left\{ x \in \mathbb{R}^d : x_i \ge 0 \text{ for } 1 \le i \le d \right\}, \qquad \mathbb{Z}^d_+ = \mathbb{Z}^d \cap \mathbb{R}^d_+$$

and  $f(r) \ge 0$  is a continuous decreasing (i.e. non-increasing) function of  $r \ge 0$ . With this notation we have

$$\sum_{k \in \mathbb{Z}_+^d} f\left( \|k\| \right) < \infty \iff \int_0^\infty f\left( r \right) r^{d-1} dr < \infty.$$

**Proof.** Let us set f(r) = f(0) for  $r \leq 0$  and let  $Q = (0, 1]^d$  and for  $k \in \mathbb{Z}_+^d$ , let  $Q_k := k + Q$  be the translate of Q by k. For any  $x = k + y \in Q_k$  we have

$$||k|| \le ||x||$$
 and  $||x|| \le ||k|| + ||y|| \le ||k|| + \sqrt{d}$ ,

i.e.

$$||x|| - \sqrt{d} \le ||k|| \le ||x||$$
 for  $x \in Q_k$ .

Thus it follows that

$$f\left(\|x\| - \sqrt{d}\right) \ge f\left(\|k\|\right) \ge f\left(\|x\|\right) \text{ for } x \in Q_k.$$

Thus if let

$$F(x) := \sum_{k \in \mathbb{Z}^{d}_{+}} f(\|k\|) 1_{Q_{k}}(x)$$

we have shown

$$f\left(\|x\| - \sqrt{d}\right) \ge F\left(x\right) \ge f\left(\|x\|\right) \text{ for } x \in \mathbb{R}^{d}_{+}.$$
(1.9)

Recalling that

$$\int_{\mathbb{R}^{d}_{+}} f\left(\left\|x\right\|\right) dm\left(x\right) = c_{d} \int_{0}^{\infty} f\left(r\right) r^{d-1} dr$$

for some constant  $c_d < \infty$ , we may integrate Eq. (1.9) over  $\mathbb{R}^d_+$  to find,

$$c_d \int_0^\infty f\left(r - \sqrt{d}\right) r^{d-1} dr \ge \sum_{k \in \mathbb{Z}_+^d} f\left(\|k\|\right) \ge c_d \int_0^\infty f\left(r\right) r^{d-1} dr.$$

Since

$$\int_0^\infty f\left(r - \sqrt{d}\right) r^{d-1} dr = \int_{-\sqrt{d}}^\infty f\left(s\right) \left(s + \sqrt{d}\right)^{d-1} ds$$
$$\leq f\left(0\right) \int_{-\sqrt{d}}^{\sqrt{d}} \left(s + \sqrt{d}\right)^{d-1} ds + \int_{\sqrt{d}}^\infty f\left(s\right) \left(s + \sqrt{d}\right)^{d-1} ds$$
$$\leq f\left(0\right) \cdot C\left(d\right) + 2^{d-1} \int_{\sqrt{d}}^\infty f\left(s\right) s^{d-1} ds$$

and we have shown,

$$c_{d}f(0) \cdot C(d) + c_{d}2^{d-1} \int_{\sqrt{d}}^{\infty} f(s) s^{d-1} ds \ge \sum_{k \in \mathbb{Z}_{+}^{d}} f(\|k\|) \ge c_{d} \int_{0}^{\infty} f(r) r^{d-1} dr$$

from which the result easilyt follows.

Page: 3

**Corollary 1.8.** If  $d \in \mathbb{N}$  and  $f(r) \geq 0$  is a continuous decreasing (i.e. non-increasing) function of  $r \geq 0$ , then

$$\sum_{k\in\mathbb{Z}^{d}}f\left(\|k\|\right)<\infty\iff \int_{0}^{\infty}f\left(r\right)r^{d-1}dr<\infty.$$

**Proof.** For  $\varepsilon \in \{\pm 1\}^d$ , let  $\mathbb{Z}^d_{\varepsilon} = \{k \in \mathbb{Z}^d : \varepsilon_i k_i \ge 0 \text{ for } 1 \le i \le d\}$ . Then

$$\sum_{k \in \mathbb{Z}_{\varepsilon}^{d}} f\left(\|k\|\right) = \sum_{k \in \mathbb{Z}_{+}^{d}} f\left(\|k\|\right) \text{ for all } \varepsilon \in \{\pm 1\}^{d}$$

and since  $\mathbb{Z}^d_+\subset\mathbb{Z}^d\subset\cup_\varepsilon\mathbb{Z}^d_\varepsilon$  it follows that

$$\sum_{k \in \mathbb{Z}^d} f\left(\|k\|\right) < \infty \iff \sum_{k \in \mathbb{Z}^d_+} f\left(\|k\|\right) < \infty \iff \int_0^\infty f\left(r\right) r^{d-1} dr < \infty.$$

## Metric-Measure Space Regularity Results

This section is a self study guide to the "approximating" Borel sets in a metric space by closed and open subsets of the metric space. We will see similar results in more general topological spaces later in the book. [See Section ?? and also subsection ?? for related results.] We begin with some basic properties of metric spaces. Throughout this section we will assume that  $(X, \rho)$  is a metric space and  $\mathcal{B}_X$  denotes the Borel  $\sigma$ -algebra on X.

#### 2.1 Metric space results

**Lemma 2.1.** For any non empty subset  $A \subset X$ , let  $\rho_A(x) := \inf \{ \rho(x, a) | a \in A \}$ , then

$$\left|\rho_A\left(x\right) - \rho_A(y)\right| \le \rho\left(x, y\right) \ \forall x, y \in X \tag{2.1}$$

which shows  $\rho_A : X \to [0, \infty)$  is continuous.

**Proof.** Let  $a \in A$  and  $x, y \in X$ , then

$$\rho_A(x) \le \rho(x, a) \le \rho(x, y) + \rho(y, a).$$

Take the infimum over a in the above equation shows that

$$\rho_A(x) \le \rho(x, y) + \rho_A(y) \quad \forall x, y \in X.$$

Therefore,  $\rho_A(x) - \rho_A(y) \le \rho(x, y)$  and by interchanging x and y we also have that  $\rho_A(y) - \rho_A(x) \le \rho(x, y)$  which implies Eq. (2.1).

Corollary 2.2. The function  $\rho$  satisfies,

$$|\rho(x, y) - \rho(x', y')| \le \rho(y, y') + \rho(x, x').$$

In particular  $\rho: X \times X \to [0,\infty)$  is continuous.

**Proof.** By Lemma 2.1 for single point sets and the triangle inequality for the absolute value of real numbers,

$$\begin{aligned} |\rho(x,y) - \rho(x',y')| &\leq |\rho(x,y) - \rho(x,y')| + |\rho(x,y') - \rho(x',y')| \\ &\leq \rho(y,y') + \rho(x,x'). \end{aligned}$$

**Corollary 2.3.** Given any set  $A \subset X$  and  $\varepsilon > 0$ , then

$$A_{\varepsilon} := \{\rho_A < \varepsilon\} := \{x \in X : \rho_A(x) < \varepsilon\}$$

is an open set containing A and  $A_{\varepsilon} \downarrow \overline{A}$  as  $\varepsilon \downarrow 0$  where  $\overline{A}$  is the closure of A. Similarly,

$$F_{\varepsilon} := \{\rho_A \ge \varepsilon\} = \{x \in X : \rho_A(x) \ge \varepsilon\}$$

is a closed set and  $F_{\varepsilon} \uparrow (A^c)^o$  as  $\varepsilon \downarrow 0$  where  $(A^c)^o$  is the interior of  $A^c := X \setminus A$ .

**Proof.** Because of the continuity of  $\rho_A$  and the facts that  $(-\infty, \varepsilon)$  is open in  $\mathbb{R}$  and  $[\varepsilon, \infty)$  is closed in  $\mathbb{R}$ , it follows that  $A_{\varepsilon} = \rho_A^{-1}((-\infty, \varepsilon))$  is open and  $F_{\varepsilon} = \rho_A^{-1}([\varepsilon, \infty))$  is closed. We have  $x \in \bigcap_{\varepsilon > 0} A_{\varepsilon}$  iff  $\rho_A(x) < \varepsilon$  for all  $\varepsilon > 0$  iff  $\rho_A(x) = 0$  and hence

$$A \subset \{\rho_A = 0\} = \bigcap_{\varepsilon > 0} A_{\varepsilon}.$$

Since  $\{\rho_A = 0\}$  is closed it follows that  $\overline{A} \subset \{\rho_A = 0\}$ . Conversely if  $x \in \{\rho_A = 0\}$  then there exists  $\{x_n\} \subset A$  such that  $\lim_{n\to\infty} \rho(x, x_n) = 0$ , i.e.  $x_n \to x$  and therefore  $x \in \overline{A}$ .

To finish the proof observe that

$$\left[\bigcup_{\varepsilon>0}F_{\varepsilon}\right]^{c}=\bigcap_{\varepsilon>0}F_{\varepsilon}^{c}=\bigcap_{\varepsilon>0}\left\{\rho_{A}<\varepsilon\right\}=\bar{A}$$

and therefore

$$\cup_{\varepsilon>0} F_{\varepsilon} = \bar{A}^c = (A^c)^o$$

**Lemma 2.4 (Urysohn's Lemma for Metric Spaces).** Let (X, d) be a metric space and suppose that A and B are two disjoint closed subsets of X. Then

$$f(x) = \frac{d_B(x)}{d_A(x) + d_B(x)} \text{ for } x \in X$$

$$(2.2)$$

defines a continuous function,  $f: X \to [0,1]$ , such that f(x) = 1 for  $x \in A$  and f(x) = 0 if  $x \in B$ .

**Proof.** By Lemma 2.1,  $d_A$  and  $d_B$  are continuous functions on X. Since A and B are closed,  $d_A(x) > 0$  if  $x \notin A$  and  $d_B(x) > 0$  if  $x \notin B$ . Since

#### 6 2 Metric-Measure Space Regularity Results

 $A \cap B = \emptyset$ ,  $d_A(x) + d_B(x) > 0$  for all x and  $(d_A + d_B)^{-1}$  is continuous as well. The remaining assertions about f are all easy to verify.

Sometimes Urysohn's lemma will be use in the following form. Suppose  $F \subset V \subset X$  with F being closed and V being open, then there exists  $f \in C(X, [0, 1])$  such that f = 1 on F while f = 0 on  $V^c$ . This of course follows from Lemma 2.4 by taking A = F and  $B = V^c$ .

**Corollary 2.5.** If A and B are two disjoint closed subsets of a metric space, (X, d), then there exists disjoint open subsets U and V of X such that  $A \subset U$  and  $B \subset V$ .

**Proof.** Let f be as in Lemma 2.4 so that  $f \in C(X \to [0,1])$  such that f = 1 on A and f = 0 on B. Then set  $U = \{f > \frac{1}{2}\}$  and  $V = \{f < 1/2\}$ .

We end this subsection with the following simple variant of Proposition ??. This proposition shows how to associate a pseudo metric to any measure space.

**Proposition 2.6 (The measure pseudo metric).** Let  $(\Omega, \mathcal{B}, \mu)$  be a measure space and define

$$d_{\mu}(A,B) := \mu(A \bigtriangleup B) \in [0,\infty] \ \forall A, B \in \mathcal{B}.$$

Then  $d = d_{\mu}$  satisfies;

1. d is a pseudo metric, i.e. d(A, B) = d(B, A) and  $d(A, C) \le d(A, B) + d(B, C)$  for all  $A, B, C \in \mathcal{B}$ . 2.  $d(A^c, C^c) = d(A, C)$  for all  $A, B \in \mathcal{B}$ . 3. If  $\{A_n\}_{n=1}^{\infty}, \{B_n\}_{n=1}^{\infty} \subset \mathcal{B}$ , then

$$d\left(\cup_{n=1}^{\infty}A_n, \bigcup_{n=1}^{\infty}B_n\right) \le \sum_{n=1}^{\infty}d\left(A_n, B_n\right) \quad and \tag{2.3}$$

$$d\left(\bigcap_{n=1}^{\infty}A_n,\bigcap_{n=1}^{\infty}B_n\right) \le \sum_{n=1}^{\infty}d\left(A_n,B_n\right).$$
(2.4)

In summary,

$$\max\left\{d\left(\bigcap_{n=1}^{\infty}A_{n},\bigcap_{n=1}^{\infty}B_{n}\right),d\left(\bigcup_{n=1}^{\infty}A_{n},\bigcup_{n=1}^{\infty}B_{n}\right)\right\}\leq\sum_{n=1}^{\infty}d\left(A_{n},B_{n}\right).$$
(2.5)

**Proof.** We take each item in turn.

1. The fact that d is a pseudo metric easily follows from the fact that  $1_{A \triangle C} = |1_A - 1_C|$  and therefore,

$$d(A,C) = \|1_A - 1_C\|_1.$$

2. Item 2. follows from the fact that

$$A^{c} \bigtriangleup C^{c} = [A^{c} \cap C] \cup [C^{c} \cap A] = [C \setminus A] \cup [A \setminus C] = A \bigtriangleup C$$

which is also seen via,

$$1_{A^c \triangle C^c} = |1_{A^c} - 1_{C^c}| = |[1 - 1_A] - [1 - 1_C]| = |1_A - 1_C| = 1_{A \triangle C}.$$

3. It is a simple exercise to verify,

$$\left[\cup_{n=1}^{\infty}A_n\right] \bigtriangleup \left[\cup_{n=1}^{\infty}B_n\right] \subset \cup_{n=1}^{\infty}\left[A_n \bigtriangleup B_n\right]$$

and hence

$$d\left(\cup_{n=1}^{\infty}A_{n},\cup_{n=1}^{\infty}B_{n}\right) = \mu\left(\left[\cup_{n=1}^{\infty}A_{n}\right] \bigtriangleup \left[\cup_{n=1}^{\infty}B_{n}\right]\right) \le \mu\left(\cup_{n=1}^{\infty}\left[A_{n}\bigtriangleup B_{n}\right]\right)$$
$$\le \sum_{n=1}^{\infty}\mu\left(A_{n}\bigtriangleup B_{n}\right) = \sum_{n=1}^{\infty}d\left(A_{n},B_{n}\right),$$

which proves Eq. (2.3). Equation (2.4) may be proved similarly or by combining item 2. with Eq. (2.3) as follows;

$$d\left(\bigcap_{n=1}^{\infty} A_{n}, \bigcap_{n=1}^{\infty} B_{n}\right)$$
  
=  $d\left(\left[\bigcap_{n=1}^{\infty} A_{n}\right]^{c}, \left[\bigcap_{n=1}^{\infty} B_{n}\right]^{c}\right)$   
=  $d\left(\bigcup_{n=1}^{\infty} A_{n}^{c}, \bigcup_{n=1}^{\infty} B_{n}^{c}\right) \leq \sum_{n=1}^{\infty} d\left(A_{n}^{c}, B_{n}^{c}\right) = \sum_{n=1}^{\infty} d\left(A_{n}, B_{n}\right).$ 

2.2 Regularity Results for Borel measures on  $(X, \mathcal{B}_X)$ 

**Exercise 2.1.** If  $(X, \rho)$  is a metric space and  $\mu$  is a **finite** measure on  $(X, \mathcal{B}_X)$ , then for all  $A \in \mathcal{B}_X$  and  $\varepsilon > 0$  there exists a closed set F and open set V such that  $F \subset A \subset V$  and  $\mu(V \setminus F) = \mu(F \bigtriangleup V) < \varepsilon$ . Here are some suggestions.

- 1. Let  $\mathcal{B}_0$  denote those  $A \subset X$  such that for all  $\varepsilon > 0$  there exists a closed set F and open set V such that  $F \subset A \subset V$  and  $d_{\mu}(F, V) = \mu(V \setminus F) < \varepsilon$ .
- 2. Show  $\mathcal{B}_0$  contains all closed (or open if you like) sets using using Corollary 2.3.
- 3. Show  $\mathcal{B}_0$  is a  $\sigma$ -algebra. [You may find Proposition 2.6 to be helpful in this step.]
- 4. Explain why this proves the result.

**Exercise 2.2.** Let  $(X, \rho)$  be a metric space and  $\mu$  be a measure on  $(X, \mathcal{B}_X)$ . If there exists open sets,  $\{V_n\}_{n=1}^{\infty}$ , of X such that  $V_n \uparrow X$  and  $\mu(V_n) < \infty$  for all n, then for all  $A \in \mathcal{B}_X$  and  $\varepsilon > 0$  there exists a closed set F and open set V such that  $F \subset A \subset V$  and  $d_{\mu}(F, V) = \mu(V \setminus F) < \varepsilon$ . **Hints:** 

- 1. Show it suffices to prove; for all  $\varepsilon > 0$  and  $A \in \mathcal{B}_X$ , there exists an open set  $V \subset X$  such that  $A \subset V$  and  $\mu(V \setminus A) < \varepsilon$ .
- 2. Now you must verify the assertion above holds. For this, you may find it useful to apply Exercise 2.1 to the measures,  $\mu_n : \mathcal{B}_X \to [0, \mu(V_n)]$ , defined by  $\mu_n(A) := \mu(A \cap V_n)$  for all  $A \in \mathcal{B}_X$ . The  $\varepsilon$  in Exercise 2.1 should be replaced by judiciously chosen small quantities depending on n.

**Theorem 2.7.** Suppose that  $(X, \rho)$  is a metric space and  $\mu$  is a measure on  $(X, \mathcal{B}_X)$  such that  $\mu(K) < \infty$  whenever K is a compact subset of X. If there exists open sets,  $\{V_n\}_{n=1}^{\infty}$ , of X such that  $V_n \uparrow X$  and  $\overline{V}_n$  is compact for all  $n \in \mathbb{N}$ , then  $C_c(X, \mathbb{C})$  is dense in  $L^p(\mu)$  for all  $1 \leq p < \infty$ .

**Proof.** Suppose that  $A \in \mathcal{B}_X$  is a set such that  $\mu(A) < \infty$  and let  $\varepsilon > 0$  be given. By Exercise 2.2, there exists a closed set F and an open set V such that  $F \subset A \subset V$  and  $\mu(V \setminus F) < \varepsilon$ . [Note that  $\mu(F) \leq \mu(A) \leq \mu(V) \leq \mu(A) + \varepsilon < \infty$ .] For each  $m \in \mathbb{N}$ ,  $K_m := F \cap \overline{V}_m$  are compact subsets of F such that  $K_m \uparrow F$  as  $m \uparrow \infty$ . By DCT if follows  $\lim_{m \to \infty} \mu(V \setminus K_m) = \mu(V \setminus F) < \varepsilon$  and hence three exists a  $m \in \mathbb{N}$  so that  $\mu(V \setminus K_m) < \varepsilon$ . Thus if we let  $K := K_m$ , then K is compact,  $K \subset A \subset V$ , and  $\mu(V \setminus K) < \varepsilon$ . Moreover, since  $K \subset X = \bigcup_{n=1}^{\infty} V_n$ , there exists (by compactness) an  $n \in \mathbb{N}$  such that  $K \subset V_n \cap V$ .

We now define  $\delta := \rho \left( K_m, \left[ V_n \cap V \right]^c \right) > 0$  and then define

$$f(x) = \left[1 - \frac{2}{\delta}\rho_K(x)\right]_+$$
 for all  $x \in X$ .

Since

$$\{f > 0\} \subset \left\{1 - \frac{2}{\delta}\rho_K > 0\right\} \subset \left\{\rho_K < \frac{1}{2}\delta\right\} \subset \left\{\rho_K \le \frac{\delta}{2}\right\} \subset V_n \cap V$$

it follows that

$$\operatorname{supp}(f) = \overline{\{f > 0\}} \subset V_n \subset \overline{V}_n.$$

Thus  $f \in C_c(X, [0, 1])$ , f = 1 on K, and f = 0 on  $V^c$ , and hence

$$|f - 1_A| \le 1_{V \cap V_n \setminus K} \le 1_{V \setminus K}$$

from which it follows that

$$\|f - 1_A\|_p \le \|1_{V \setminus K}\|_p \le \varepsilon^{1/p}.$$

As  $\varepsilon > 0$  was arbitrary, we have shown  $1_A \in \overline{C_c(X, \mathbb{C})}^{L^p(\mu)}$  for all  $A \in \mathcal{B}_X$  with  $\mu(A) < \infty$ . This completes the proof since simple functions which are in  $L^p(\mu)$  are known to be dense in  $L^p(\mu)$ .

**Corollary 2.8.** If X is an open subset of  $\mathbb{R}^n$  and  $\mu$  is a measure on  $\mathcal{B}_X$  such that  $\mu(K) < \infty$  for all compact subsets,  $K \subset U$ , then  $C_c(X, \mathbb{C})$  is dense in  $L^p(\mu)$  for all  $1 \leq p < \infty$ .

**Proof.** Let  $\rho(x, y) := |y - x|$  be the usual Euclidean metric on  $X \subset \mathbb{R}^n$  and define

$$V_n := \left\{ \rho_{X^c} > \frac{1}{n} \right\} \cap B_\rho\left(0, n\right),$$

where

$$B_{\rho}(0,n) := \{x \in \mathbb{R}^n : \rho(x,0) = |x| < n\}.$$

Then  $V_n$  is an open subset of X such that

$$\bar{V}_n \subset \left\{ \rho_{X^c} \ge \frac{1}{n} \right\} \cap \overline{B_\rho\left(0,n\right)} \subset X.$$

As  $\overline{V}_n$  is closed and bounded it is compact and since  $V_n \uparrow X$  as  $n \to \infty$ , the result now follows by an application of Theorem 2.7.

**Corollary 2.9.** Suppose that  $(X, \rho)$  is a metric space with open sets,  $\{V_n\}_{n=1}^{\infty} \subset X$  such that  $V_n \uparrow X$  and  $\bar{V}_n$  is compact for all  $n \in \mathbb{N}$  and  $\nu$  is a complex measure on  $(X, \mathcal{B}_X)$ . If  $\int_X f d\nu = 0$  for all  $f \in C_c(X)$ , then  $\nu \equiv 0$ .

**Proof.** If we let  $\mu := |\nu|$ , then there is a measurable function,  $g: X \to S^1 \subset \mathbb{C}$  such that  $d\nu = gd\mu$ . Since  $C_c(X)$  is dense in  $L^1(\mu)$ , there exists  $f_n \in C_c(X)$  such that  $f_n \to \overline{g}$  in  $L^1(\mu)$  as  $n \to \infty$ . Therefore,

$$0=\int_{X}f_{n}d\nu=\int_{X}f_{n}gd\mu\rightarrow\int_{X}\bar{g}gd\mu=\int_{X}d\mu=\mu\left(X\right).$$

This shows  $\mu(X) = 0$  and hence  $\nu \equiv 0$ .

**Definition 2.10.** If  $\nu$  is a complex measure on  $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ , let  $\hat{\nu} : \mathbb{R}^n \to \mathbb{C}$  be the characteristic function of  $\nu$  defined by,

$$\hat{\nu}\left(\lambda\right):=\int_{\mathbb{R}^{n}}e^{-i\lambda\cdot x}d\nu\left(x\right) \ \forall \ \lambda\in\mathbb{R}^{n}.$$

**Corollary 2.11.** Let  $\nu$  be a complex measure on  $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ . If  $\hat{\nu} \equiv 0$ , then  $\nu = 0$ , *i.e.* the linear map,

{complex measures on  $\mathbb{R}^n$ }  $\ni \nu \to \hat{\nu} \in$  {functions on  $\mathbb{R}^n$ },

is injective.

#### 8 2 Metric-Measure Space Regularity Results

**Proof.** Suppose that  $f \in C_c(\mathbb{R}^n)$ . For N > 0 large let

$$f_N(x) := \sum_{k \in \mathbb{Z}^n} f(x + Nk).$$

Then  $f_N$  is a bounded continuous function such that  $f_N(x + Ne_j) = f_N(x)$  for all  $x \in \mathbb{R}^n$ . Given  $\varepsilon > 0$ , by the Stone–Weierstrass theorem we can then find a function  $g_{\varepsilon}(x)$  of the form

$$g_{\varepsilon}\left(x\right) = \sum_{k \in \Lambda} a_{k} e^{i\frac{1}{2\pi N}k \cdot x}$$

where  $\Lambda \subset_f \mathbb{Z}^n$  and  $a_k \in \mathbb{C}$  such that  $\max_{x \in \mathbb{R}^n} |f_N(x) - g_{\varepsilon}(x)| \leq \varepsilon$ . Under the assumption that  $\hat{\nu} \equiv 0$ , we will have

$$\int_{\mathbb{R}^{n}} g_{\varepsilon}(x) \, d\nu(x) = \sum_{k \in \Lambda} a_{k} \hat{\nu}\left(\frac{1}{2\pi N}k\right) = 0$$

and hence

$$\int_{\mathbb{R}^{n}} f_{N}(x) d\nu(x) \bigg| \leq \left| \int_{\mathbb{R}^{n}} \left[ f_{N}(x) - g_{\varepsilon}(x) \right] d\nu(x) \right| \leq \varepsilon \left| \nu \right| (\mathbb{R}^{n}).$$

As  $\varepsilon > 0$  was arbitrary, it follows that

$$\int_{\mathbb{R}^{n}} f_{N}(x) \, d\nu(x) = 0$$

and then by letting  $N \to \infty$  it follows that  $\int_{\mathbb{R}^n} f(x) d\nu(x) = 0$ . The proof is then completed by an application of Corollary 2.9.

**Definition 2.12.** If  $g \in L^1(m)$  where m is Lebesgue measure on  $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n})$ , then we define the **Fourier transform** of g by

$$\hat{g}\left(\lambda\right):=\int_{\mathbb{R}^{n}}e^{-i\lambda\cdot x}g\left(x\right)dm\left(x\right).$$

**Corollary 2.13.** If  $g \in L^1(m)$  and  $\hat{g} \equiv 0$ , then g = 0 m-a.e.

**Proof.** The measure,  $d\nu = gdm$ , is a complex measure such that  $\hat{\nu}(\lambda) = \hat{g}(\lambda)$ . So if  $\hat{g} \equiv 0$  implies  $\hat{\nu} \equiv 0$  which implies  $\nu = 0$ . Hence it follows that

$$0 = |\nu| \left( \mathbb{R}^n \right) = \int_{\mathbb{R}^n} |g| \, dm \implies g = 0 \text{ a.e}$$

#### 2.3 Dual Considerations

As in Theorem 2.7 let us suppose for simplicity that  $(X, \rho)$  is a metric space such that there exists open sets,  $\{V_n\}_{n=1}^{\infty}$ , of X such that  $V_n \uparrow X$  and  $\bar{V}_n$  is compact for all  $n \in \mathbb{N}$ . We now suppose that  $\nu$  is a complex measure on  $(X, \mathcal{B}_X)$ and for  $f \in C_0(X)$ , let

$$\nu\left(f\right):=\int_{X}fd\nu,$$

i.e. we identify  $\nu$  with an element of  $C_0(X)^*$ . Our goal is to show  $\|\nu\|_{C_0(X)^*} = |\nu|(X)$ . To do this we will use the following simple lemma.

**Lemma 2.14 (Sliding points).** Let  $\varphi : \mathbb{C} \to \overline{D} := \{z \in \mathbb{C} : |z| \le 1\}$  be defined by

$$\varphi(z) = (|z| \lor 1)^{-1} \ z = \begin{cases} z & \text{if } |z| \le 1 \\ \frac{z}{|z|} & \text{if } |z| > 1 \end{cases}$$

Then  $\varphi$  is continuous and satisfies

$$\left|\varphi\left(z\right)-w\right|\leq\left|z-w\right|\ \forall\ z\in\mathbb{C}\ and\ w\in S^{1},$$

see Figure 2.1



Fig. 2.1. Sliding points to the unit circle.

**Proof.** It is easy to verify  $\varphi$  is continuous. If  $w, z \in S^1$  then

$$\frac{d}{dt} |w - tz|^2 = \frac{d}{dt} \left[ 1 + t^2 - 2t \operatorname{Re}(\bar{w}z) \right] = 2 \left[ t - \operatorname{Re}(\bar{w}z) \right] > 0 \text{ if } t > 1.$$

This shows  $|w - \varphi(tz)| \ge |w - z|$  for all  $t \ge 1$ .

**Theorem 2.15 (Dual of**  $C_0(X)$ ). Let  $(X, \rho)$  be a metric space such that there exists open sets,  $\{V_n\}_{n=1}^{\infty}$ , of X such that  $V_n \uparrow X$  and  $\overline{V}_n$  is compact for all  $n \in \mathbb{N}$ . If  $\nu$  is a complex measure on  $(X, \mathcal{B}_X)$ , then

$$\|\nu\|_{C_0(X)^*} = |\nu|(X)$$

**Proof.** Let  $\mu = |\nu|$  and  $q: X \to S^1$  be chosen so that  $d\nu = qd\mu$ . Then for  $f \in C_0(X),$ 

$$\begin{split} |\nu\left(f\right)| &= \left|\int_{X} f d\nu\right| = \left|\int_{X} f g d\mu\right| \\ &\leq \int_{X} |f| \, d\mu \leq \|f\|_{u} \cdot \mu\left(X\right) = |\nu|\left(X\right) \|f\|_{\mu} \end{split}$$

which shows that

$$\|\nu\|_{C_0(X)^*} \le |\nu|(X).$$

To prove the reverse inequality us Corollary 2.8 to find  $f_n \in C_c(X) \subset C_0(X)$ such that  $f_n \to \overline{g}$  in  $L^1(\mu)$  as  $n \to \infty$ . Let  $g_n = \varphi(f_n)$  where  $\varphi : \mathbb{C} \to \mathbb{C}$  is the the continuous function in Lemma 2.14. Then

$$|g_n - \bar{g}| \le |\varphi(f_n) - \bar{g}| \le |f_n - \bar{g}|$$

and hence  $g_n \to \bar{g}$  in  $L^1(\mu)$  where now  $||g_n||_u \leq 1$  and hence

$$\|\nu\|_{C_0(X)^*} \ge |\nu\left(g_n\right)| = \left|\int_X g_n g d\mu\right| \to \left|\int_X \bar{g} g d\mu\right| = \mu\left(X\right) = |\nu|\left(X\right).$$

This shows that  $\|\nu\|_{C_0(X)^*} \ge |\nu|(X)$  and the proof is complete. For completeness, let me now state a form of the Riesz-Markov theorem in the context being considered here.

**Theorem 2.16 (Riesz-Markov Theorem).** Let  $(X, \rho)$  be a metric space such that there exists open sets,  $\{V_n\}_{n=1}^{\infty}$ , of X such that  $V_n \uparrow X$  and  $\overline{V}_n$  is compact for all  $n \in \mathbb{N}$ .

1. If  $\varphi$  is a positive linear functional on  $C_{c}(X)$ , then there exists a unique positive measure,  $\mu$ , on  $(X, \mathcal{B}_X)$  such that  $\mu(K) < \infty$  when K is compact and

$$\varphi \left( f \right) = \mu \left( f \right) := \int_X f d\mu$$

for all  $f \in C_c(X)$ .

2. If  $\varphi \in C_0(X)^*$ , then there exists a unique complex measure  $\nu$  on  $(X, \mathcal{B}_X)$ such that  $\varphi(f) = \nu(f)$  for all  $f \in C_0(X)$ . Moreover the map,

{complex measures on X} 
$$\ni \nu \to \left(f \to \nu(f) = \int_X f d\nu\right) \in C_0(X)^*$$

is an isometric isomorphism of Banach space where  $\|\nu\| := |\nu|(X)$  where  $\nu$ is a complex measure.

The next theorem gives an important and interesting example of using the Riesz-Markov theorem. For this theorem we will be using the following notation.

**Notation 2.17** Given a sequence of metric spaces,  $\{(X_n, \rho_n)\}_{n=1}^{\infty}$ , for each  $N \in \mathbb{N}, let X^{(N)} := X_1 \times \cdots \times X_N,$ 

$$\mathcal{B}^{(N)}=\mathcal{B}_{X_1\times\cdots\times X_N}=\mathcal{B}_{X_1}\otimes\mathcal{B}_{X_2}\otimes\cdots\otimes\mathcal{B}_{X_N},$$

 $X := X^{(\infty)} := \prod_{n=1}^{\infty} X_n, \ \mathcal{B}^{(\infty)} = \mathcal{B}_X, \ and \ \pi^{(N)} : X \to X^{(N)}$  be the projection map,

$$\pi^{(N)}(x) = (x_1, \dots, x_N) \quad \forall \ x \in X.$$

Remark 2.18. If the metrics,  $\rho_n$ , are all bounded by 1 (can do this by replacing  $\rho_n$  by  $\frac{\rho_n}{1+\rho_n}$  if necessary), then

$$\rho\left(x,y\right) := \sum_{n=1}^{\infty} \frac{1}{2^n} \rho_n\left(x_n, y_n\right)$$

defined a metric on X whose topology is consistent with the product topology on X.

Remark 2.19. Let

$$\mathbb{D} := \bigcup_{N=1}^{\infty} \left[ C\left( X^{(N)}, \mathbb{R} \right) \circ \pi^{(N)} \right] \subset C\left( X, \mathbb{R} \right).$$

An easy application of the Stone-Weirstrass Theorem shows that  $\mathbb{D}$  is a (uniformally) dense subspace of  $C(X,\mathbb{R})$ . We will use this result freely in the proof to follow.

**Theorem 2.20.** Suppose that  $\{(X_n, \rho_n)\}_{n=1}^{\infty}$  is a sequence of compact metric spaces and for each  $N \in \mathbb{N}$ ,  $\mu^{(N)}$  is a probability measure on  $(X^{(N)}, \mathcal{B}^{(N)})$ . If for each  $N \in \mathbb{N}$ ,

$$\mu^{(N+1)} \left( A \times X_{N+1} \right) = \mu^{(N)} \left( A \right) \ \forall \ A \in \mathcal{B}^{(N)}, \tag{2.6}$$

then there exists a unique probability measure,  $\mu$ , on  $(X, \mathcal{B})$  such that

$$\pi_*^{(N)}\mu = \mu^{(N)} \ \forall \ N \in \mathbb{N}.$$

**Proof. Uniqueness.** If  $\mu$  and  $\nu$  are two such measures, then for  $F \in$  $C(X^{(N)},\mathbb{R})$ , we will have

$$\int_X F \circ \pi_N d\mu = \int_{X^{(N)}} F d\mu^{(N)} = \int_X F \circ \pi_N d\nu.$$

from which it follows that

$$\int_X f d\mu = \int_X f d\nu \ \forall \ f \in \mathbb{D}.$$

By Remark 2.19 and DCT it then follows that

$$\int_{X} f d\mu = \int_{X} f d\nu \; \forall \; f \in C\left(X, \mathbb{R}\right)$$

Applying Theorem 2.15 to  $\lambda = \mu - \nu$  shows that  $\mu = \nu$ . Existence. For  $f = F \circ \pi^{(N)} \in \mathbb{D}$ , let

$$\lambda\left(f\right):=\int_{X^{(N)}}Fd\mu^{(N)}.$$

We must first show this definition is well defined. For example we could write  $f = G \circ \pi^{(N+1)}$  where  $G(x_1, \ldots, x_{N+1}) = F(x_1, \ldots, x_N)$ . In this case we have

$$\int_{X^{(N+1)}} G d\mu^{(N+1)} = \int_{X^{(N)} \times X_{N+1}} F \otimes 1_{X_{N+1}} d\mu^{(N+1)} = \int_{X^{(N)}} F d\mu^{(N)},$$

which follows by approximating F by simple functions and then making use of Eq. (2.6). It then follows by induction, if M > N and  $G(x_1, \ldots, x_M) = F(x_1, \ldots, x_N)$ , that

$$\int_{X^{(M)}} G d\mu^{(M)} = \int_{X^{(N)}} F d\mu^{(N)}$$

Thus we have shown that  $\lambda$  is well defined. It is now clearly linear and positive on  $\mathbb D$  and moreover

$$|\lambda(f)| \le \int_{X^{(N)}} |F| \, d\mu^{(N)} \le \|F\|_u = \|f\|_u \,.$$

Hence and application of the BLT theorem allows us to extend  $\lambda$  to a bounded linear functional on  $C(X, \mathbb{R})$ . If  $f \geq 0$  in  $C(X, \mathbb{R})$  and  $f_n \in \mathbb{D}$  converges to f uniformly, then max  $(f_n, 0) \in \mathbb{D}$  and

$$||f - \max(f_n, 0)||_u \le ||f - f_n||_u \to 0$$

Therefore,

$$\lambda\left(f\right) = \lim_{n \to \infty} \lambda\left(\max\left(f_n, 0\right)\right) \ge 0$$

which shows the extended  $\lambda$  is still positive. By the Riesz-Markov theorem, there exists a unique measure  $\mu$  on  $(X, \mathcal{B}_X)$  such that

$$\lambda(f) = \int_{X} f d\mu \text{ for all } f \in C(X, \mathbb{R})$$

Taking  $f = F \circ \pi^{(N)}$  with  $F \in C(X^{(N)}, \mathbb{R})$  now shows

$$\int_{X^{(N)}} F d\mu^{(N)} = \lambda(f) = \int_X F \circ \pi^{(N)} d\mu = \int_{X^{(N)}} F d\left[\pi_*^{(N)} \mu\right]$$

and hence by Theorem 2.15 again it follows that  $\mu^{(N)} = \pi^{(N)}_* \mu$ .

## **Fourier Series**

**Theorem 3.1.** Suppose that  $\lambda$  is a complex measure  $((-\pi, \pi), \mathcal{B} = \mathcal{B}_{(-\pi, \pi)})$ . If  $\int_{(-\pi, \pi)} e^{in\theta} d\nu (\theta) = 0 \text{ for all } n \in \mathbb{Z}$ 

on

then  $\nu \equiv 0$ .

**Proof.** For  $f \in C_c((-\pi,\pi),\mathbb{C})$ , let  $F(e^{i\theta}) = f(\theta)$  for  $-\pi \leq \theta \leq \pi$ . Then F is a continuous function on  $S^1$  (which is 0 in neighborhood of  $-1 \in S^1$ ) and hence by the Stone–Weierstrass theorem, given  $\varepsilon > 0$  there exists  $N < \infty$  and  $\{a_{m,n}\}_{m=0}^{N} \subset \mathbb{C}$  such that

$$\max_{z\in S^{1}}\left|F\left(z\right)-\sum_{m,n=0}^{N}a_{m,n}z^{m}\bar{z}^{n}\right|\leq\varepsilon.$$

Evaluating this expression at  $z = e^{i\theta}$  then shows

$$\left| f\left(\theta\right) - \sum_{m,n=0}^{N} a_{m,n} e^{i(m-n)\theta} \right| \le \varepsilon.$$

Therefore

$$\left| \int_{(-\pi,\pi)} f(\theta) \, d\nu(\theta) \right| = \left| \int_{(-\pi,\pi)} \left[ f(\theta) - \sum_{m,n=0}^{N} a_{m,n} e^{i(m-n)\theta} \right] d\nu(\theta) \right|$$
$$\leq \varepsilon \left| \nu \right| \left( (-\pi,\pi) \right).$$

As  $\varepsilon > 0$  was arbitrary it follows that

$$\int_{(-\pi,\pi)} f(\theta) \, d\nu(\theta) = 0 \text{ for all } f \in C_c((-\pi,\pi),\mathbb{C})$$

and we have seen this implies  $\nu \equiv 0$ .

**Corollary 3.2.** Let  $\mathbb{D} := \operatorname{span}_{\mathbb{C}} \{\theta \to e^{in\theta}\}_{n \in \mathbb{Z}}$ . If  $\mu$  is a finite positive measure on  $((-\pi, \pi), \mathcal{B} = \mathcal{B}_{(-\pi, \pi)})$ , then  $\mathbb{D}$  is dense in  $L^p(\mu)$  for all  $1 \leq p < \infty$ .

**Proof. First proof.** According to the Hahn-Banach theorem, in order to show  $\mathbb{D}$  is dense it suffices to show if  $\varphi \in L^p(\mu)^*$  satisfies  $\varphi|_{\mathbb{D}} \equiv 0$ , then  $\varphi \equiv 0$ . Since  $L^p(\mu)^* \cong L^{p^*}(\mu)$ , there exists  $g \in L^{p^*}(\mu)$  such that

$$\varphi(f) = \int_{(-\pi,\pi)} f(\theta) g(\theta) d\mu(\theta) \text{ for all } f \in L^{p}(\mu).$$

Letting  $d\nu = gd\mu$  (a complex measure) as  $g \in L^{p*}(\mu) \subset L^{1}(\mu)$ , the assumption that  $\varphi|_{\mathbb{D}} \equiv 0$  implies

$$0 = \varphi \left( \theta \to e^{in\theta} \right) = \int_{(-\pi,\pi)} e^{in\theta} dv \left( \theta \right) \text{ for all } n \in \mathbb{Z}.$$

From Theorem 3.1 this implies that  $\nu \equiv 0$  and hence  $d |\nu| = |g| d\mu$  is the zero measure and hence |g| = 0 for  $\mu$ -a.e. Thus g = 0 in  $L^{p*}(\mu)$  and so  $\varphi \equiv 0$ .

**Second proof.** In the proof of Theorem 3.1 we have shown every element,  $f \in C_c((-\pi,\pi),\mathbb{C})$  may be uniformly approximated by an element of  $\mathbb{D}$  and hence in  $L^p(\mu)$  for all  $1 \leq p < \infty$  because  $\mu$  is a finite measure. But we already know that  $C_c((-\pi,\pi),\mathbb{C})$  is dense<sup>1</sup> in  $L^p(\mu)$  and hence the proof is complete.

**Theorem 3.3.** Let m be Lebesgue measure on  $(-\pi, \pi)$  and for  $f, g \in L^2(m)$ , let

$$\langle f|g\rangle := rac{1}{2\pi} \int_{-\pi}^{\pi} f\left(\theta\right) \bar{g}\left(\theta\right) d\theta.$$

Then  $\varphi_{n}(\theta) := e^{in\theta}$  for  $n \in \mathbb{Z}$  forms an orthonormal basis for  $L^{2}(m)$ .

The above results easily generalize to the case where  $(-\pi, \pi)$  is replaced by  $(-\pi, \pi)^d$  for any  $d \in \mathbb{N}$ . We now setup some more notation.

Notation 3.4 (Periodic functions) Let  $C_{per}(\mathbb{R}^d)$  denote the  $2\pi$  – periodic functions in  $C(\mathbb{R}^d)$ , that is  $f \in C_{per}(\mathbb{R}^d)$  iff  $f \in C(\mathbb{R}^d)$  and  $f(\theta+2\pi e_i) = f(\theta)$  for all  $\theta \in \mathbb{R}^d$  and i = 1, 2, ..., d. We further let  $C_{per}^k(\mathbb{R}^d) = C_{per}(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$  for all  $k \in \mathbb{N}$ . Here  $\{e_i\}_{i=1}^d$  is the standard basis for  $\mathbb{R}^d$ .

<sup>&</sup>lt;sup>1</sup> As we have seen, the assertion that  $C_c((-\pi,\pi),\mathbb{C})$  is dense in  $L^p(\mu)$  holds even if  $\mu$  is an infinite measure which is finite on compact sets.

12 3 Fourier Series

#### Definition 3.5. Let

$$\mathbb{D} = \operatorname{span}_{\mathbb{C}} \left\{ \mathbb{R}^d \ni x \to e^{ik \cdot x} \right\}_{k \in \mathbb{Z}^d} \subset C_{per}^{\infty} \left( \mathbb{R}^d \right) := \cap_k C_{per}^k \left( \mathbb{R}^d \right).$$

In more detail,  $f \in \mathbb{D}$  iff there exists a function,  $a : \mathbb{Z}^d \to \mathbb{C}$  with finite support (i.e.  $\# \{k \in \mathbb{Z}^d : a(k) \neq 0\} < \infty$ ) such that

$$f(x) = f_a(x) = \sum_{k \in \mathbb{Z}^d} a(k) e^{ik \cdot x} \text{ for all } x \in \Omega.$$

**Theorem 3.6 (Density of Trigonometric Polynomials).** Any  $2\pi$  – periodic continuous function,  $f : \mathbb{R} \to \mathbb{C}$ , may be uniformly approximated by a trigonometric polynomial of the form

$$p\left(x\right) = \sum_{\lambda \in \Lambda} a_{\lambda} e^{i\lambda \cdot x}$$

where  $\Lambda$  is a finite subset of  $\mathbb{Z}$  and  $a_{\lambda} \in \mathbb{C}$  for all  $\lambda \in \Lambda$ .

**Proof.** For  $z \in S^1$ , define  $F(z) := f(\theta)$  where  $\theta \in \mathbb{R}$  is chosen so that  $z = e^{i\theta}$ . Since f is  $2\pi$  – periodic, F is well defined since if  $\theta$  solves  $e^{i\theta} = z$  then all other solutions are of the form  $\{\theta + 2\pi n : n \in \mathbb{Z}\}$ . Since the map  $\theta \to e^{i\theta}$  is a local homeomorphism, i.e. for any J = (a, b) with  $b - a < 2\pi$ , the map  $\theta \in J \xrightarrow{\varphi} \tilde{J} := \{e^{i\theta} : \theta \in J\} \subset S^1$  is a homeomorphism, it follows that  $F(z) = f \circ \varphi^{-1}(z)$  for  $z \in \tilde{J}$ . This shows F is continuous when restricted to  $\tilde{J}$ . Since such sets cover  $S^1$ , it follows that F is continuous.

By Example ??, the polynomials in z and  $\overline{z} = z^{-1}$  are dense in  $C(S^1)$ . Hence for any  $\varepsilon > 0$  there exists

$$p(z,\bar{z}) = \sum_{0 \le m,n \le N} a_{m,n} z^m \bar{z}^n$$

such that  $|F(z) - p(z, \bar{z})| \leq \varepsilon$  for all  $z \in S^1$ . Taking  $z = e^{i\theta}$  then implies

$$\sup_{\theta} \left| f(\theta) - p\left(e^{i\theta}, e^{-i\theta}\right) \right| \le \varepsilon$$

where

$$p\left(e^{i\theta}, e^{-i\theta}\right) = \sum_{0 \le m, n \le N} a_{m,n} e^{i(m-n)\theta}$$

is the desired trigonometry polynomial.

**Exercise 3.1.** Use Example ?? to show that any  $2\pi$  – periodic continuous function,  $f : \mathbb{R}^d \to \mathbb{C}$ , may be uniformly approximated by a trigonometric polynomial of the form

$$p\left(x\right) = \sum_{\lambda \in \Lambda} a_{\lambda} e^{i\lambda \cdot x}$$

where  $\Lambda$  is a finite subset of  $\mathbb{Z}^d$  and  $a_{\lambda} \in \mathbb{C}$  for all  $\lambda \in \Lambda$ .

**Hint:** start by showing there exists a unique continuous function,  $F : (S^1)^d \to \mathbb{C}$  such that  $F(e^{ix_1}, \ldots, e^{ix_d}) = f(x)$  for all  $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ .

**Exercise 3.2.** Let  $\Omega = (-\pi, \pi)^d$ ,  $\mathcal{B} = \mathcal{B}_\Omega$  be the Borel  $\sigma$ -algebra on  $\Omega$ ,  $\nu$  be any complex measure on  $(\Omega, \mathcal{B})$ . Show that  $\nu \equiv 0$  iff

$$\int_{\Omega} e^{ik \cdot x} d\nu \left( x \right) = 0 \text{ for all } k \in \mathbb{Z}^{d}.$$

**Hint:** each  $f \in C_c(\Omega, \mathbb{C})$  may be extended to zero on  $\mathbb{R}^d \setminus \Omega$  and in this way may be viewed as an element of  $C_c(\mathbb{R}^d, \mathbb{C})$ . Using this extended f, let  $F(\theta) := \sum_{k \in \mathbb{Z}^d} f(\theta + 2\pi k)$  so that  $F \in C_{per}(\mathbb{R}^d)$ . Given  $\varepsilon > 0$ , use the Stone–Weierstrass theorem to show there exists  $\Lambda \subset_f \mathbb{Z}^d$  and  $a : \Lambda \to \mathbb{C}$  such that

$$\sup_{x \in \mathbb{R}^d} \left| F(x) - \sum_{k \in \Lambda} a(k) e^{ik \cdot x} \right| \le \varepsilon.$$
(3.1)

**Exercise 3.3.** Let  $\Omega = (-\pi, \pi)^d$ ,  $\mathcal{B} = \mathcal{B}_\Omega$  be the Borel  $\sigma$ -algebra on  $\Omega$ , and  $\mathbb{D}$  be as in Definition ??. If  $\mu$  is a finite positive measure on  $(\Omega, \mathcal{B})$ , show  $\mathbb{D}$  is dense in  $L^p(\Omega, \mathcal{B}, \mu)$  for all  $1 \leq p < \infty$ . **Hint:** using  $L^p(\Omega, \mathcal{B}, \mu)^* \cong L^{p^*}(\Omega, \mathcal{B}, \mu)$  where  $p^* = \frac{p}{p-1}$  and a corollary of the Hahn-Banach theorem, show it suffices to show if  $q \in L^{p^*}(\Omega, \mathcal{B}, \mu)$  satisfies,

$$\int_{\Omega} e^{ik \cdot x} g(x) \, d\mu(x) = 0 \text{ for all } k \in \mathbb{Z}^d,$$
(3.2)

then g(x) = 0 for  $\mu$ -a.e. x.

#### 3.1 Dirichlet Kernel

Although the sum in Eq. (??) is guaranteed to converge relative to the Hilbertian norm on H it certainly need not converge pointwise even if  $f \in C_{per}(\mathbb{R}^d)$ as will be proved in Section ?? below. Nevertheless, if f is sufficiently regular, then the sum in Eq. (??) will converge pointwise as we will now show. In the process we will give a direct and constructive proof of the result in Exercise ??, see Theorem 3.11 below.

Let us restrict our attention to d = 1 here. Consider

$$f_{n}(\theta) = \sum_{|k| \le n} \tilde{f}(k)\varphi_{k}(\theta) = \sum_{|k| \le n} \frac{1}{2\pi} \left[ \int_{[-\pi,\pi]} f(x) e^{-ik \cdot x} dx \right] \varphi_{k}(\theta)$$
$$= \frac{1}{2\pi} \int_{[-\pi,\pi]} f(x) \sum_{|k| \le n} e^{ik \cdot (\theta - x)} dx$$
$$= \frac{1}{2\pi} \int_{[-\pi,\pi]} f(x) D_{n}(\theta - x) dx$$
(3.3)

where

$$D_n\left(\theta\right) := \sum_{k=-n}^n e^{ik\theta}$$

is called the **Dirichlet kernel**. Letting  $\alpha = e^{i\theta/2}$ , we have

$$D_n(\theta) = \sum_{k=-n}^n \alpha^{2k} = \frac{\alpha^{2(n+1)} - \alpha^{-2n}}{\alpha^2 - 1} = \frac{\alpha^{2n+1} - \alpha^{-(2n+1)}}{\alpha - \alpha^{-1}}$$
$$= \frac{2i\sin(n+\frac{1}{2})\theta}{2i\sin\frac{1}{2}\theta} = \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta}.$$

and therefore

$$D_n(\theta) := \sum_{k=-n}^n e^{ik\theta} = \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta},$$
(3.4)

see Figure 3.1. with the understanding that the right side of this equation is



**Fig. 3.1.** This is a plot  $D_1$  and  $D_{10}$ .

2n+1 whenever  $\theta \in 2\pi\mathbb{Z}$ .

**Theorem 3.7.** Suppose  $f \in L^1([-\pi,\pi], dm)$  and f is differentiable at some  $\theta \in [-\pi,\pi]$ , then  $\lim_{n\to\infty} f_n(\theta) = f(\theta)$  where  $f_n$  is as in Eq. (3.3).

**Proof.** Observe that

$$\frac{1}{2\pi}\int_{[-\pi,\pi]}D_n(\theta-x)dx = \frac{1}{2\pi}\int_{[-\pi,\pi]}\sum_{|k|\leq n}e^{ik\cdot(\theta-x)}dx = 1$$

and therefore,

$$f_{n}(\theta) - f(\theta) = \frac{1}{2\pi} \int_{[-\pi,\pi]} [f(x) - f(\theta)] D_{n}(\theta - x) dx$$
  
$$= \frac{1}{2\pi} \int_{[-\pi,\pi]} [f(x) - f(\theta - x)] D_{n}(x) dx$$
  
$$= \frac{1}{2\pi} \int_{[-\pi,\pi]} \left[ \frac{f(\theta - x) - f(\theta)}{\sin\frac{1}{2}x} \right] \sin(n + \frac{1}{2}) x dx.$$
(3.5)

If f is differentiable at  $\theta$ , then

$$\lim_{x \to 0} \frac{f(\theta - x) - f(\theta)}{\sin \frac{1}{2}x} = -2f'(x)$$

and hence there exists  $\varepsilon > 0$  such that

$$M_{\varepsilon} := \sup_{|x| \le \varepsilon} \left| \frac{f(\theta - x) - f(\theta)}{\sin \frac{1}{2}x} \right| < \infty.$$

Using this remark it is now easily seen that

$$1_{[-\pi,\pi]}(x) \, \frac{f(\theta - x) - f(\theta)}{\sin \frac{1}{2}x} \in L^1\left([-\pi,\pi], dm\right)$$

and hence the last expression in Eq. (3.5) tends to 0 as  $n \to \infty$  by the Riemann Lebesgue Lemma, see Corollary ?? or Lemma 4.20).

**Proposition 3.8 (Lack of pointwise convergence).** For each  $\alpha \in [-\pi, \pi] / \sim$ , there exists a residual (non-meager set) set  $R_{\alpha} \subset C_{per}(\mathbb{R})$  such that  $\sup_{n} |f_{n}(\alpha)| = \infty$  for all  $f \in R_{\alpha}$ .<sup>2</sup> Recall that  $C_{per}(\mathbb{R})$  is a complete metric space, hence  $R_{\alpha}$  is a dense subset of  $C_{per}(\mathbb{R})$ .

**Proof.** By symmetry considerations, it suffices to assume  $\alpha = 0 \in [-\pi, \pi]$ . Let  $d\nu_n(\theta) := \frac{1}{2\pi} D_n(\theta) d\theta$  which is a complex measure on  $(-\pi, \pi)$  which identify with an element of  $C_{per}(\mathbb{R})^*$  by

<sup>&</sup>lt;sup>2</sup> Recall this means that  $R_{\alpha}$  contains a countable union of dense open subsets of  $C_{per}(\mathbb{R})$  and such a set is dense in a complete metric space!

$$\nu_{n}(f) := f_{n}(0) = \frac{1}{2\pi} \int_{(-\pi,\pi)} f(\theta) D_{n}(\theta) d\theta.$$

Recall that

$$\|\nu_n\|_{\rm op} = \frac{1}{2\pi} \|D_n\|_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(e^{-i\theta})| \, d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\sin(n+\frac{1}{2})\theta}{\sin\frac{1}{2}\theta} \right| \, d\theta.$$

Using

$$\left|\sin x\right| = \left|\int_0^x \cos y \, dy\right| \le \left|\int_0^x \left|\cos y\right| \, dy\right| \le |x|$$

$$\begin{aligned} \left|\nu_{n}\right|_{\text{op}} &\geq \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|\frac{\sin(n+\frac{1}{2})\theta}{\frac{1}{2}\theta}\right| d\theta = \frac{2}{\pi} \int_{0}^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta} \\ &= \frac{2}{\pi} \int_{0}^{\pi} \left|\sin(n+\frac{1}{2})\theta\right| \frac{d\theta}{\theta} = \int_{0}^{(n+\frac{1}{2})\pi} \left|\sin y\right| \frac{dy}{y} \to \infty \text{ as } n \to \infty \end{aligned} (3.6)$$

and hence  $\sup_n \left\| \nu_n \right\|_{\mathrm{op}} = \infty.$  So by uniform boundedness principal it follows that

$$R_0 = \{ f \in C_{per} \left( \mathbb{R} \right) : \sup_{n \in \mathbb{R}} |\nu_n f| = \infty \}$$

is a residual set. [See Rudin [?, Chapter 5] for more details.]

Lemma 3.9 (Fourier Series on  $L^1$ ). For  $f \in L^1((-\pi, \pi))$ , let

$$\tilde{f}\left(n\right) := \left\langle f | \varphi_n \right\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(\theta\right) e^{-in\theta} d\theta$$

Then  $\tilde{f} \in c_0 := C_0(\mathbb{Z})$  (i.e.  $\lim_{n \to \infty} \tilde{f}(n) = 0$ ) and the map  $f \in L^1(T) \xrightarrow{A} \tilde{f} \in c_0$ is a one to one bounded linear transformation into but **not** onto  $c_0$ .

**Proof.** By the Riemann Lebesgue Lemma we know that  $\lim_{|n|\to\infty} \tilde{f}(n) = 0$ so that  $\tilde{f} \in c_0$  as claimed. Moreover if  $\tilde{f} \equiv 0$ , then by Theorem 3.1 we know that  $d\nu(\theta) := \frac{1}{2\pi}f(\theta) d\theta$  is the zero measure and hence  $f(\theta) = 0$  for a.e.  $\theta$ . This shows that  $\Lambda$  is injective. If  $\Lambda$  were surjective, the open mapping theorem would imply that  $\Lambda^{-1}: c_0 \to L^1(T)$  is bounded. In particular this implies there exists  $C < \infty$  such that

$$\|f\|_{L^1} \le C \left\|\tilde{f}\right\|_{c_0} \text{ for all } f \in L^1(T).$$

$$(3.7)$$

Taking  $f = D_n$ , we find (because  $\tilde{D}_n(k) = 1_{|k| \le n}$ ) that  $\left\| \tilde{D}_n \right\|_{c_0} = 1$  while by Eq. (3.6)  $\lim_{n \to \infty} \left\| \tilde{D}_n \right\|_{L^1} = \infty$  contradicting Eq. (3.7). Therefore  $\operatorname{Ran}(\Lambda) \neq c_0$ .

## 3.2 Fejér Kernel

Despite the Dirichlet kernel not being positive, it still satisfies the approximate  $\delta$  – sequence property,  $\frac{1}{2\pi}D_n \to \delta_0$  as  $n \to \infty$ , when acting on  $C^1$  – periodic functions in  $\theta$ . In order to improve the convergence properties it is reasonable to try to replace  $\{f_n : n \in \mathbb{N}_0\}$  by the sequence of averages (see Exercise ??),

$$F_{N}(\theta) = \frac{1}{N+1} \sum_{n=0}^{N} f_{n}(\theta) = \frac{1}{N+1} \sum_{n=0}^{N} \frac{1}{2\pi} \int_{[-\pi,\pi]} f(x) \sum_{|k| \le n} e^{ik \cdot (\theta-x)} dx$$
$$= \frac{1}{2\pi} \int_{[-\pi,\pi]} K_{N}(\theta-x) f(x) dx$$

where

$$K_{N}(\theta) := \frac{1}{N+1} \sum_{n=0}^{N} \sum_{|k| \le n} e^{ik \cdot \theta} = \frac{1}{N+1} \sum_{n=0}^{N} D_{n}(\theta)$$
(3.8)

is the  ${\bf Fejér}\ {\bf kernel}$  which we now compute more explicitly.

**Lemma 3.10.** The Fejér kernel  $K_N$  in Eq. (3.8) is given by

$$K_N(\theta) = \frac{1}{N+1} \frac{\sin^2\left(\frac{N+1}{2}\theta\right)}{\sin^2\left(\frac{\theta}{2}\right)}$$
(3.9)

and we also have the identity,

$$K_N(\theta) = \sum_{k=-N}^{N} \left[ 1 - \frac{|k|}{N+1} \right] e^{ik\theta}$$
(3.10)

**Proof.** Recall the trigonometric identities,

$$\cos\left(A\pm B\right) = \cos A \cdot \cos B \mp \sin A \cdot \sin B$$

which subtracted gives

$$\cos\left(A+B\right) - \cos\left(A-B\right) = -2\sin A \cdot \sin B.$$

Further taking A = B in the last equality also gives,

$$\cos\left(2A\right) - 1 = -2\sin^2 A.$$

Using these identities and a telescoping sum argument proves Eq.  $\left( 3.9\right)$  as follows,

Page: 14

$$(N+1) K_N(\theta) := \sum_{n=0}^{N} D_n(\theta)$$
  
=  $\frac{1}{2 \sin^2 \frac{1}{2}\theta} \sum_{n=0}^{N} 2 \sin \frac{1}{2}\theta \cdot \sin(n + \frac{1}{2})\theta$   
=  $-\frac{1}{2 \sin^2 \frac{1}{2}\theta} \sum_{n=0}^{N} [\cos((n+1)\theta) - \cos(n\theta)]$   
=  $\frac{1}{2 \sin^2 \frac{1}{2}\theta} [1 - \cos(N+1)\theta]$   
=  $\frac{1}{2 \sin^2 \frac{1}{2}\theta} 2 \sin^2 \left(\frac{1}{2} (N+1)\theta\right).$ 

Equation (3.10) is a consequence of the identity,

$$(N+1) K_N(\theta) = \sum_{n=0}^{N} \sum_{|k| \le n} e^{ik \cdot \theta} = \sum_{|k| \le n \le N} e^{ik \cdot \theta} = \sum_{|k| \le N} (N+1-|k|) e^{ik \cdot \theta}.$$

**Theorem 3.11.** The Fejér kernel  $K_N$  in Eq. (3.8) satisfies:

- 1.  $K_N(\theta) \ge 0$ . 2.  $\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(\theta) d\theta = 1$
- 3.  $\sup_{\varepsilon \le |\theta| \le \pi} K_N(\theta) \to 0 \text{ as } N \to \infty \text{ for all } \varepsilon > 0, \text{ see Figure 3.2.}$ 4. For any continuous  $2\pi$  periodic function f on  $\mathbb{R}$ ,  $K_N * f(\theta) \to f(\theta)$ uniformly in  $\theta$  as  $N \to \infty$ , where

$$K_N * f(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(\theta - \alpha) f(\alpha) \, d\alpha$$
$$= \sum_{n=-N}^{N} \left[ 1 - \frac{|n|}{N+1} \right] \tilde{f}(n) \, e^{in\theta}. \tag{3.11}$$

**Proof.** Items 1. is obvious form Eq. (3.9) and item 2. follows from the fact that  $K_N$  is an average of Dirichlet kernels which all integrate to 1, i.e.

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(\theta) \, d\theta = \frac{1}{N+1} \sum_{n=0}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} D_n(\theta) \, d\theta$$
$$= \frac{1}{N+1} \sum_{n=0}^{N} 1 = 1.$$



Fig. 3.2. Plots of  $K_N(\theta)$  for N = 2, 7 and 13.

We can also prove item 2. by integrating Eq. (3.10). Item 3. is a consequence of the elementary estimate;

$$\sup_{\varepsilon \le |\theta| \le \pi} K_N(\theta) \le \frac{1}{N+1} \frac{1}{\sin^2\left(\frac{\varepsilon}{2}\right)}$$

and is clearly indicated in Figure 3.2.

Finally, item 4. now follows by the standard approximate  $\delta$  – function arguments, namely,

$$|K_N * f(\theta) - f(\theta)| = \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} K_N(\theta - \alpha) \left[ f(\alpha) - f(\theta) \right] d\alpha \right|$$
  
$$\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} K_N(\alpha) \left| f(\theta - \alpha) - f(\theta) \right| d\alpha$$
  
$$\leq \frac{1}{\pi} \frac{1}{N+1} \frac{1}{\sin^2\left(\frac{\varepsilon}{2}\right)} \left\| f \right\|_{\infty} + \frac{1}{2\pi} \int_{|\alpha| \le \varepsilon} K_N(\alpha) \left| f(\theta - \alpha) - f(\theta) \right| d\alpha$$
  
$$\leq \frac{1}{\pi} \frac{1}{N+1} \frac{1}{\sin^2\left(\frac{\varepsilon}{2}\right)} \left\| f \right\|_{\infty} + \sup_{|\alpha| \le \varepsilon} \left| f(\theta - \alpha) - f(\theta) \right|.$$

Therefore,

$$\lim_{N \to \infty} \sup_{0 \to \infty} \|K_N * f - f\|_{\infty} \leq \sup_{\theta} \sup_{|\alpha| \leq \varepsilon} |f(\theta - \alpha) - f(\theta)| \to 0 \text{ as } \varepsilon \downarrow 0.$$

### 3.3 The Dirichlet Problems on D and the Poisson Kernel

Let  $D := \{z \in \mathbb{C} : |z| < 1\}$  be the open unit disk in  $\mathbb{C} \cong \mathbb{R}^2$ , write  $z \in \mathbb{C}$  as z = x + iy or  $z = re^{i\theta}$ , and let  $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$  be the **Laplacian** acting on  $C^{2}(D)$ .

**Theorem 3.12 (Dirichlet problem for** D). To every continuous function  $g \in C(\mathrm{bd}(D))$  there exists a unique function  $u \in C(\bar{D}) \cap C^2(D)$  solving

$$\Delta u(z) = 0 \text{ for } z \in D \text{ and } u|_{\partial D} = g. \tag{3.12}$$

Moreover for r < 1, u is given by,

$$u(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r(\theta - \alpha) u(e^{i\alpha}) d\alpha =: P_r * u(e^{i\theta})$$
(3.13)

$$= \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} \frac{1 + re^{i(\theta - \alpha)}}{1 - re^{i(\theta - \alpha)}} u(e^{i\alpha}) d\alpha$$
(3.14)

where  $P_r$  is the **Poisson kernel** defined by

$$P_r(\delta) := \frac{1 - r^2}{1 - 2r\cos\delta + r^2}.$$

(The problem posed in Eq. (3.12) is called the **Dirichlet problem for** D.)

**Proof.** In this proof, we are going to be identifying  $S^1 = \operatorname{bd}(D) := \{z \in \overline{D} : |z| = 1\}$  with  $[-\pi, \pi] / (\pi \sim -\pi)$  by the map  $\theta \in [-\pi, \pi] \to e^{i\theta} \in S^1$ . Also recall that the Laplacian  $\Delta$  may be expressed in polar coordinates as,

$$\Delta u = r^{-1}\partial_r \left(r^{-1}\partial_r u\right) + \frac{1}{r^2}\partial_\theta^2 u,$$

where

$$(\partial_r u) (re^{i\theta}) = \frac{\partial}{\partial r} u (re^{i\theta}) \text{ and } (\partial_\theta u) (re^{i\theta}) = \frac{\partial}{\partial \theta} u (re^{i\theta})$$

**Uniqueness.** Suppose u is a solution to Eq. (3.12) and let

$$\tilde{g}(k) := \frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{ik\theta}) e^{-ik\theta} d\theta$$

and

$$\tilde{u}(r,k) := \frac{1}{2\pi} \int_{-\pi}^{\pi} u(re^{i\theta}) e^{-ik\theta} d\theta$$
(3.15)

be the Fourier coefficients of  $g(\theta)$  and  $\theta \to u(re^{i\theta})$  respectively. Then for  $r \in (0,1)$ ,

$$\begin{aligned} r^{-1}\partial_r \left(r\partial_r \tilde{u}(r,k)\right) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} r^{-1}\partial_r \left(r^{-1}\partial_r u\right) \left(re^{i\theta}\right) e^{-ik\theta} d\theta \\ &= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{r^2} \partial_{\theta}^2 u(re^{i\theta}) e^{-ik\theta} d\theta \\ &= -\frac{1}{r^2} \frac{1}{2\pi} \int_{-\pi}^{\pi} u(re^{i\theta}) \partial_{\theta}^2 e^{-ik\theta} d\theta \\ &= \frac{1}{r^2} k^2 \tilde{u}(r,k) \end{aligned}$$

or equivalently

$$r\partial_r \left( r\partial_r \tilde{u}(r,k) \right) = k^2 \tilde{u}(r,k). \tag{3.16}$$

Recall the general solution to

$$r\partial_r \left( r\partial_r y(r) \right) = k^2 y(r) \tag{3.17}$$

may be found by trying solutions of the form  $y(r) = r^{\alpha}$  which then implies  $\alpha^2 = k^2$  or  $\alpha = \pm k$ . From this one sees that  $\tilde{u}(r,k)$  solving Eq. (3.16) may be written as  $\tilde{u}(r,k) = A_k r^{|k|} + B_k r^{-|k|}$  for some constants  $A_k$  and  $B_k$  when  $k \neq 0$ . If k = 0, the solution to Eq. (3.17) is gotten by simple integration and the result is  $\tilde{u}(r,0) = A_0 + B_0 \ln r$ . Since  $\tilde{u}(r,k)$  is bounded near the origin for each k it must be that  $B_k = 0$  for all  $k \in \mathbb{Z}$ . Hence we have shown there exists  $A_k \in \mathbb{C}$  such that, for all  $r \in (0, 1)$ ,

$$A_k r^{|k|} = \tilde{u}(r,k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(re^{i\theta}) e^{-ik\theta} d\theta.$$
 (3.18)

Since all terms of this equation are continuous for  $r \in [0, 1]$ , Eq. (3.18) remains valid for all  $r \in [0, 1]$  and in particular we have, at r = 1, that

$$A_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(e^{i\theta}) e^{-ik\theta} d\theta = \tilde{g}(k)$$

Hence if u is a solution to Eq. (3.12) then u must be given by

$$u(re^{i\theta}) = \sum_{k \in \mathbb{Z}} \tilde{g}(k) r^{|k|} e^{ik\theta} \text{ for } r < 1.$$
(3.19)

or equivalently,

$$u(z) = \sum_{k \in \mathbb{N}_0} \tilde{g}(k) z^k + \sum_{k \in \mathbb{N}} \tilde{g}(-k) \bar{z}^k.$$

Notice that the theory of the Fourier series implies Eq. (3.19) is valid in the  $L^2(d\theta)$  - sense. However more is true, since for r < 1, the series in Eq. (3.19) is absolutely convergent and in fact defines a  $C^{\infty}$  – function (see Exercise ?? or Corollary ??) which must agree with the continuous function,  $\theta \to u(re^{i\theta})$ , for almost every  $\theta$  and hence for all  $\theta$ . This completes the proof of uniqueness.

**Existence.** Given  $g \in C(\mathrm{bd}(D))$ , let u be defined as in Eq. (3.19). Then, again by Exercise ?? or Corollary ??,  $u \in C^{\infty}(D)$ . So to finish the proof it suffices to show  $\lim_{x\to y} u(x) = g(y)$  for all  $y \in \mathrm{bd}(D)$ . Inserting the formula for  $\tilde{g}(k)$  into Eq. (3.19) gives

$$u(re^{i\theta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r \left(\theta - \alpha\right) u(e^{i\alpha}) d\alpha \text{ for all } r < 1$$

where

$$P_r(\delta) = \sum_{k \in \mathbb{Z}} r^{|k|} e^{ik\delta} = \sum_{k=0}^{\infty} r^k e^{ik\delta} + \sum_{k=0}^{\infty} r^k e^{-ik\delta} - 1 =$$

$$= \operatorname{Re}\left[2\frac{1}{1 - re^{i\delta}} - 1\right] = \operatorname{Re}\left[\frac{1 + re^{i\delta}}{1 - re^{i\delta}}\right]$$

$$= \operatorname{Re}\left[\frac{\left(1 + re^{i\delta}\right)\left(1 - re^{-i\delta}\right)}{|1 - re^{i\delta}|^2}\right] = \operatorname{Re}\left[\frac{1 - r^2 + 2ir\sin\delta}{1 - 2r\cos\delta + r^2}\right] \quad (3.20)$$

$$= \frac{1 - r^2}{1 - 2r\cos\delta + r^2}.$$

The Poisson kernel again solves the usual approximate  $\delta$  – function properties (see Figure 2), namely:

1.  $P_r(\delta) > 0$  and

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} P_r \left(\theta - \alpha\right) d\alpha &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{k \in \mathbb{Z}} r^{|k|} e^{ik(\theta - \alpha)} d\alpha \\ &= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} r^{|k|} \int_{-\pi}^{\pi} e^{ik(\theta - \alpha)} d\alpha = 1 \end{aligned}$$

and 2.

$$\sup_{\varepsilon < |\theta| < \pi} P_r\left(\theta\right) \leq \frac{1 - r^2}{1 - 2r\cos\varepsilon + r^2} \to 0 \text{ as } r \uparrow 1.$$



Therefore by the same argument used in the proof of Theorem 3.11,

$$\lim_{r\uparrow 1} \sup_{\theta} \left| u\left(re^{i\theta}\right) - g\left(e^{i\theta}\right) \right| = \lim_{r\uparrow 1} \sup_{\theta} \left| \left(P_r * g\right)\left(e^{i\theta}\right) - g\left(e^{i\theta}\right) \right| = 0$$

which certainly implies  $\lim_{x \to y} u(x) = g(y)$  for all  $y \in bd(D)$ .

Remark 3.13 (Harmonic Conjugate). Writing  $z = re^{i\theta}$ , Eq. (3.14) may be rewritten as

$$u(z) = \frac{1}{2\pi} \operatorname{Re} \int_{-\pi}^{\pi} \frac{1 + z e^{-i\alpha}}{1 - z e^{-i\alpha}} u(e^{i\alpha}) d\alpha$$

which shows  $u = \operatorname{Re} F$  where

$$F\left(z\right):=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1+ze^{-i\alpha}}{1-ze^{-i\alpha}}u(e^{i\alpha})d\alpha.$$

Moreover it follows from Eq. (3.20) that

$$\operatorname{Im} F(re^{i\theta}) = \frac{1}{\pi} \operatorname{Im} \int_{-\pi}^{\pi} \frac{r \sin(\theta - \alpha)}{1 - 2r \cos(\theta - \alpha) + r^2} g(e^{i\alpha}) d\alpha$$
$$=: (Q_r * u) (e^{i\theta})$$

where

$$Q_r(\delta) := \frac{r\sin(\delta)}{1 - 2r\cos(\delta) + r^2}.$$

From these remarks it follows that  $v =: (Q_r * g) (e^{i\theta})$  is the harmonic conjugate of u and  $\tilde{P}_r = Q_r$ . For more on this point see Section ?? below.

#### 3.4 Multi-Dimensional Fourier Series

In this subsection we will let  $d\theta$ , dx,  $d\alpha$ , etc. denote standard Lebesgue measure (m) on  $\mathbb{R}^d$ , and  $Q := (-\pi, \pi]^d$ , and  $H := L^2([-\pi, \pi]^d)$ , with inner product given by

$$\langle f|g\rangle := \left(\frac{1}{2\pi}\right)^d \int_Q f(\theta)\bar{g}(\theta)d\theta = \left(\frac{1}{2\pi}\right)^d \int_Q f(\theta)\bar{g}(\theta)dm\left(\theta\right).$$

We also let  $\varphi_k(\theta) := e^{ik \cdot \theta}$  for all  $k \in \mathbb{Z}^d$  so that  $\{\varphi_k\}_{k \in \mathbb{Z}^d}$  is an orthonormal basis for H. For  $f \in L^1(Q)$ , we will write  $\tilde{f}(k)$  for the **Fourier coefficient**,

$$\tilde{f}(k) := \langle f | \varphi_k \rangle = \left(\frac{1}{2\pi}\right)^d \int_Q f(\theta) e^{-ik \cdot \theta} d\theta.$$
(3.21)

Since any  $2\pi$  – periodic functions on  $\mathbb{R}^d$  may be identified with function on the d - dimensional torus,  $\mathbb{T}^d \cong \mathbb{R}^d / (2\pi\mathbb{Z})^d \cong (S^1)^d$ . I may also write  $C(\mathbb{T}^d)$ for  $C_{per}(\mathbb{R}^d)$  and  $L^p(\mathbb{T}^d)$  for  $L^p(Q)$  where elements in  $f \in L^p(Q)$  are to be thought of as there extensions to  $2\pi$  – periodic functions on  $\mathbb{R}^d$ . **Theorem 3.14 (Fourier Series).** The functions  $\beta := \{\varphi_k : k \in \mathbb{Z}^d\}$  form an orthonormal basis for H, *i.e.* if  $f \in H$  then

$$f = \sum_{k \in \mathbb{Z}^d} \langle f | \varphi_k \rangle \varphi_k = \sum_{k \in \mathbb{Z}^d} \tilde{f}(k) \varphi_k \tag{3.22}$$

where the convergence takes place in  $L^2([-\pi,\pi]^d)$ .

**Proof.** Simple computations show  $\beta := \{\varphi_k : k \in \mathbb{Z}^d\}$  is an orthonormal set. This fact coupled with Exercise ?? which states span  $\beta$  is dense in  $L^2\left([-\pi,\pi]^d\right)^3$  completes the proof.

#### 3.5 Translation Invariant Operators

**Proposition 3.15.** Consider. For  $f \in L^2([-\pi,\pi])$  which we identify with  $2\pi$ periodic functions. Let  $f_{\alpha}(\theta) := f(\theta - \alpha) = U_{\alpha}f$  for  $\alpha \in \mathbb{R}$  which is now unitary operator on  $L^2$ . Suppose that  $T \in B(L^2([-\pi,\pi]))$  and  $TU_{\alpha} = U_{\alpha}T$  for all  $\alpha \in \mathbb{R}$ , then  $T\varphi_n = \lambda_n\varphi_n$  for some  $\lambda_n \in \mathbb{C}$  for all  $n \in \mathbb{Z}$ . If we further assume that  $\sum_n |\lambda_n| < \infty$ , then

$$(Tf)(\theta) = \int_{-\pi}^{\pi} k(\theta - \alpha) f(\alpha) \, d\alpha \text{ for a.e. } \theta, \qquad (3.23)$$

where

$$k\left(\theta\right) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \lambda_n e^{in\theta}$$

**Proof.** Since  $U_{\alpha}\varphi_n = e^{-in\alpha}\varphi_n$  we have

$$e^{-in\alpha}T\varphi_n = TU_\alpha\varphi_n = U_\alpha T\varphi_n$$

and then taking the inner product of this equation with  $\varphi_m$  shows

$$\begin{split} e^{-in\alpha} \left\langle T\varphi_n | \varphi_m \right\rangle &= \left\langle U_\alpha T\varphi_n | \varphi_m \right\rangle = \left\langle T\varphi_n | U_{-\alpha}\varphi_m \right\rangle \\ &= \left\langle T\varphi_n | e^{im\alpha}\varphi_m \right\rangle = e^{-im\alpha} \left\langle T\varphi_n | \varphi_m \right\rangle \text{ for all } \alpha \in \mathbb{R}. \end{split}$$

From this it follows that

$$\langle T\varphi_n | \varphi_m \rangle = 0 \text{ if } n \neq m$$
<sup>3</sup> Note that  $m\left( [-\pi, \pi]^d \setminus (-\pi, \pi)^d \right) = 0$  so that may identify  $L^p\left( [-\pi, \pi]^d \right)$  with  $L^p\left( (-\pi, \pi)^d \right).$ 

and hence

$$T\varphi_n = \sum_{m \in \mathbb{Z}} \left\langle T\varphi_n | \varphi_m \right\rangle \varphi_m = \left\langle T\varphi_n | \varphi_n \right\rangle \varphi_n = \lambda_n \varphi_n$$

where  $\lambda_n := \langle T\varphi_n | \varphi_n \rangle$ .

Now let us further suppose that  $\sum_{n \in \mathbb{Z}} |\lambda_n| < \infty$ . Then for  $f \in L^2([-\pi, \pi])$  we have

$$Tf = T\sum_{n \in \mathbb{Z}} \langle f | \varphi_n \rangle \varphi_n = \sum_{n \in \mathbb{Z}} \langle f | \varphi_n \rangle T\varphi_n = \sum_{n \in \mathbb{Z}} \lambda_n \langle f | \varphi_n \rangle \varphi_n.$$

Now (leaving the details to the reader) we have

$$\sum_{n \in \mathbb{Z}} \lambda_n \langle f | \varphi_n \rangle \varphi_n(\theta) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \lambda_n \int_{-\pi}^{\pi} d\alpha f(\alpha) e^{-in\alpha} e^{in\theta}$$
$$= \int_{-\pi}^{\pi} d\alpha f(\alpha) \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \lambda_n e^{-in\alpha} e^{in\theta}$$
$$= \int_{-\pi}^{\pi} d\alpha f(\alpha) k(\theta - \alpha)$$

where the sums are pointwise convergent uniformly in  $\theta.$  The usual arguments now shows that

$$(Tf)(\theta) = \int_{-\pi}^{\pi} k(\theta - \alpha) f(\alpha) d\alpha$$
 for a.e.  $\theta$ .

**Exercise 3.4.** Suppose that  $(X, \mathcal{B}, \mu)$  is a  $\sigma$ -finite measure space and B:  $L^2(\mu) \to L^2(\mu)$  is a bounded linear operator such that  $[B, M_{\varphi}] = 0$  for all  $\varphi: X \to \mathbb{C}$  which are bounded and measurable. Show there exists a bounded measurable function,  $\psi: X \to \mathbb{C}$  such that  $B = M_{\psi}$ .

**Exercise 3.5.** Let  $\mu$  be a  $\sigma$ -finite measure on  $(\mathbb{R}^d, \mathcal{B})$  and suppose that  $B : L^2(\mu) \to L^2(\mu)$  is a bounded linear operator such that  $[B, M_{\varphi_\lambda}] = 0$  for all  $\lambda \in \mathbb{R}^d$  where  $\varphi_\lambda(x) = e^{i\lambda \cdot x}$  for all  $\lambda \in \mathbb{R}^d$ . Show  $B = M_\psi$  for some  $\psi \in L^\infty(\mu)$ .

**Corollary 3.16.** If  $B : L^2(m) \to L^2(m)$  is an operator such that  $BT_a = T_a B$ for all  $a \in \mathbb{R}^d$  where  $(T_a f)(\cdot) = f(\cdot - a)$  for each  $a \in \mathbb{R}^d$ , then  ${}^{-1}B\mathcal{F} = M_{\psi}$ for some  $\psi \in L^{\infty}(m)$ . The converse holds as well.

**Proof.** This is just a matter of noting that the given assumptions on B holds iff  $\left[\mathcal{F}^{-1}B\mathcal{F}, M_{\varphi_a}\right] = 0$  for all  $a \in \mathbb{R}^d$ . We then apply the previous results.

Page: 18

## Convolution and smoothing operators

Throughout this chapter we will be solely concerned with d – dimensional Lebesgue measure, m, and we will simply write  $L^p$  for  $L^p(\mathbb{R}^d, m)$ . The main object of study here is the convolution of two functions.

**Definition 4.1 (Convolution).** Let  $f, g : \mathbb{R}^d \to \mathbb{C}$  be measurable functions. We define

$$f * g(x) = \int_{\mathbb{R}^d} f(x - y)g(y)dy$$
(4.1)

whenever the integral is defined, i.e. either  $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^d,m)$  or  $f(x-\cdot)g(\cdot) \geq 0$ . Notice that the condition that  $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^d,m)$  is equivalent to writing  $|f| * |g|(x) < \infty$ . By convention, if the integral in Eq. (4.1) is not defined, let f \* g(x) := 0.

**Notation 4.2** Given a multi-index  $\alpha \in \mathbb{Z}_+^d$ , let  $|\alpha| = \alpha_1 + \cdots + \alpha_d$ ,

$$x^{\alpha} := \prod_{j=1}^{d} x_{j}^{\alpha_{j}}, \text{ and } \partial_{x}^{\alpha} = \left(\frac{\partial}{\partial x}\right)^{\alpha} := \prod_{j=1}^{d} \left(\frac{\partial}{\partial x_{j}}\right)^{\alpha_{j}}.$$

For  $z \in \mathbb{R}^d$  and  $f : \mathbb{R}^d \to \mathbb{C}$ , let  $\tau_z f : \mathbb{R}^d \to \mathbb{C}$  be defined by  $\tau_z f(x) = f(x - z)$ .

Remark 4.3 (The Significance of Convolution).

1. Suppose that  $f, g \in L^1(m)$  are positive functions and let  $\mu$  be the measure on  $(\mathbb{R}^d)^2$  defined by

$$d\mu(x, y) := f(x) g(y) dm(x) dm(y)$$

Then if  $h : \mathbb{R} \to [0, \infty]$  is a measurable function we have

$$\begin{split} \int_{(\mathbb{R}^d)^2} h\left(x+y\right) d\mu\left(x,y\right) &= \int_{(\mathbb{R}^d)^2} h\left(x+y\right) f\left(x\right) g\left(y\right) dm\left(x\right) dm\left(y\right) \\ &= \int_{(\mathbb{R}^d)^2} h\left(x\right) f\left(x-y\right) g\left(y\right) dm\left(x\right) dm\left(y\right) \\ &= \int_{\mathbb{R}^d} h\left(x\right) f * g\left(x\right) dm\left(x\right). \end{split}$$

In other words, this shows the measure (f \* g)m is the same as  $S_*\mu$  where S(x, y) := x + y. In probability lingo, the distribution of a sum of two "independent" (i.e. product measure) measurable functions is the the convolution of the individual distributions.

2. Suppose that  $L = \sum_{|\alpha| \le k} a_{\alpha} \partial^{\alpha}$  is a constant coefficient differential operator and suppose that we can solve (uniquely) the equation Lu = g in the form

$$u(x) = Kg(x) := \int_{\mathbb{R}^d} k(x, y)g(y)dy$$

where k(x, y) is an "integral kernel." (This is a natural sort of assumption since, in view of the fundamental theorem of calculus, integration is the inverse operation to differentiation.) Since  $\tau_z L = L\tau_z$  for all  $z \in \mathbb{R}^d$ , (this is another way to characterize constant coefficient differential operators) and  $L^{-1} = K$  we should have  $\tau_z K = K\tau_z$ . Writing out this equation then says

$$\begin{split} \int_{\mathbb{R}^d} k(x-z,y)g(y)dy &= (Kg)\left(x-z\right) = \tau_z Kg\left(x\right) = \left(K\tau_z g\right)(x) \\ &= \int_{\mathbb{R}^d} k(x,y)g(y-z)dy = \int_{\mathbb{R}^d} k(x,y+z)g(y)dy. \end{split}$$

Since g is arbitrary we conclude that k(x-z, y) = k(x, y+z). Taking y = 0 then gives

$$k(x, z) = k(x - z, 0) =: \rho(x - z)$$

We thus find that  $Kg = \rho * g$ . Hence we expect the convolution operation to appear naturally when solving constant coefficient partial differential equations. More about this point later.

#### 4.1 Basic Properties of Convolutions

**Proposition 4.4.** Suppose that  $p \in [1, \infty)$ , then  $\tau_z : L^p \to L^p$  is an isometric isomorphism and for  $f \in L^p$ ,  $z \in \mathbb{R}^d \to \tau_z f \in L^p$  is uniformly continuous.

**Proof.** The assertion that  $\tau_z : L^p \to L^p$  is an isometric isomorphism follows from translation invariance of Lebesgue measure and the fact that  $\tau_{-z} \circ \tau_z = id$ . Since

#### 20 4 Convolution and smoothing operators

$$\|\tau_{z+h}f - \tau_z f\|_p = \|\tau_h f - f\|_p,$$

to see that  $z \to \tau_z f$  is uniformly continuous it suffices to show it is continuous at 0. When  $g \in C_c(\mathbb{R}^d)$  a relatively simple use of the dominated convergence theorem<sup>1</sup> shows  $\lim_{h\to 0} \|\tau_h g - g\|_p = 0$  and hence  $z \to \tau_z g$  is continuous in this case. As  $C_c(\mathbb{R}^d)$  is dense in  $L^p(\mathbb{R}^d)$ , for any  $f \in L^p$  there exists  $f_n \in C_c(\mathbb{R}^d)$ such that  $\lim_{n\to\infty} \|f - f_n\|_p = 0$ . It then follows that

$$\sup_{z \in \mathbb{R}^d} \left\| \tau_z f - \tau_z f_n \right\|_p = \left\| f - f_n \right\|_p \to 0 \text{ as } n \to \infty$$

and hence  $\mathbb{R}^{d} \ni z \to \tau_{z} f \in L^{p}(m)$  is the uniform limit of continuous functions,  $z \to \tau_{z} f_{n}$ , and therefore is itself continuous.

**Proposition 4.5.** Suppose that  $p, q \in [1, \infty]$  and p and  $q = p^* = \frac{p}{p-1}$  are conjugate exponents,  $f \in L^p$  and  $g \in L^q$ , then  $f * g \in BC(\mathbb{R}^d)$  with f \* g being uniformly continuous and satisfying,  $\|f * g\|_{\infty} \leq \|f\|_p \|g\|_q$ . If we further assume that  $p, q \in (1, \infty)$  then  $f * g \in C_0(\mathbb{R}^d)$ .

**Proof.** The existence of f \* g(x) and the estimate  $|f * g|(x) \le ||f||_p ||g||_q$  for all  $x \in \mathbb{R}^d$  is a simple consequence of Holders inequality and the translation invariance of Lebesgue measure. In particular this shows  $||f * g||_{\infty} \le ||f||_p ||g||_q$ . By relabeling p and q if necessary we may assume that  $p \in [1, \infty)$ . Since

$$|f * g(x + h) - f * g(x)| = \left| \int_{\mathbb{R}^d} [f(x + h - y) - f(x - y)] g(y) dy \right|$$
  
$$\leq ||\tau_{-h}f - f||_p ||g||_p$$

it follows that

$$\sup_{x \in \mathbb{R}^{d}} |f * g(x+h) - f * g(x)| \le ||\tau_{-h}f - f||_{p} ||g||_{p} \to 0 \text{ as } h \to 0$$

proving the uniform continuity.

If  $1 , we let <math>f_n(x) = f(x) \mathbf{1}_{|x| \le n}$  and  $g_n(x) = g(x) \mathbf{1}_{|x| \le n}$  so that  $f_n \to f$  in  $L^p(m)$  and  $g_n \to g$  in  $L^q(m)$  as  $n \to \infty$ . By what we just proved  $f_n * g_n$  is continuous and it is easily verified that  $f_n * g_n$  is supported in  $\overline{B_{2n}(0)}$ , i.e.  $f_n * g_n \in C_c(\mathbb{R}^d)$ . The proof will be completed by showing  $f_n * g_n(x) \to f * g(x)$  uniformally in  $x \in \mathbb{R}^d$  and hence  $f * g \in \overline{C_c(\mathbb{R}^d)} = C_0(\mathbb{R}^d)$ . The uniform convergence is a consequence of the simple estimates,

$$\begin{split} \|f * g - f_n * g_n\|_{\infty} &\leq \|f * g - f_n * g\|_{\infty} + \|f_n * g - f_n * g_n\|_{\infty} \\ &\leq \|f - f_n\|_p \|g\|_q + \|f_n\|_p \|g - g_n\|_q \\ &\leq \|f - f_n\|_p \|g\|_q + \|f\|_p \|g - g_n\|_q \to 0 \text{ as } n \to \infty. \end{split}$$

Alternative proof. First suppose that g is compactly supported and let  $f_n(x) = f(x) \mathbf{1}_{|x| \le n}$ . Then  $f_n * g \in C_c(\mathbb{R}^d)$  and

$$\left\|f_n * g \to f * g\right\|_{\infty} \le \left\|f - f_n\right\|_p \left\|g\right\|_q \to 0 \text{ as } n \to \infty$$

and hence  $f * g \in C_0(\mathbb{R}^d)$ . Now for general g, let  $g_n(x) = g(x) \mathbf{1}_{|x| \le n}$  so that  $f * g_n \in C_0(\mathbb{R}^d)$  and

$$\left\|f\ast g_{n}\rightarrow f\ast g\right\|_{\infty}\leq\left\|f\right\|_{p}\left\|g-g_{n}\right\|_{q}\rightarrow0$$
 as  $n\rightarrow\infty$ 

and so  $f * g \in C_0(\mathbb{R}^d)$ .

**Theorem 4.6 (Approximate**  $\delta$  – functions). Let  $p \in [1, \infty]$ ,  $\varphi \in L^1(\mathbb{R}^d)$ ,  $a := \int_{\mathbb{R}^d} \varphi(x) dx$ , and for t > 0 let  $\varphi_t(x) = t^{-d} \varphi(x/t)$ . Then

- 1. If  $f \in L^p$  with  $p < \infty$  then  $\varphi_t * f \to af$  in  $L^p$  as  $t \downarrow 0$ .
- 2. If  $f \in BC(\mathbb{R}^d)$  and f is uniformly continuous then  $\|\varphi_t * f af\|_{\infty} \to 0$  as  $t \downarrow 0$ .
- 3. If  $f \in L^{\infty}$  and f is continuous on  $U \subset_o \mathbb{R}^d$  then  $\varphi_t * f \to af$  uniformly on compact subsets of U as  $t \downarrow 0$ .

(See Proposition 4.21 below and for a statement about almost everywhere convergence.)  $% \left( \mathcal{S}_{a}^{A}\right) =\left( \mathcal{S}_{a}^{A}\right) \left( \mathcal{S}_{a}$ 

**Proof.** Making the change of variables y = tz implies

$$\varphi_t * f(x) = \int_{\mathbb{R}^d} f(x - y)\varphi_t(y)dy = \int_{\mathbb{R}^d} f(x - tz)\varphi(z) dz$$

so that

$$\varphi_t * f(x) - af(x) = \int_{\mathbb{R}^d} \left[ f(x - tz) - f(x) \right] \varphi(z) dz$$
$$= \int_{\mathbb{R}^d} \left[ \tau_{tz} f(x) - f(x) \right] \varphi(z) dz.$$
(4.2)

Hence by Minkowski's inequality for integrals (Theorem ??), Proposition 4.4 and the dominated convergence theorem,

$$\left\|\varphi_{t}*f-af\right\|_{p} \leq \int_{\mathbb{R}^{d}} \left\|\tau_{tz}f-f\right\|_{p} \left|\varphi\left(z\right)\right| dz \to 0 \text{ as } t \downarrow 0.$$

Item 2. is proved similarly. Indeed, form Eq. (4.2)

$$\left\|\varphi_{t}*f-af\right\|_{\infty} \leq \int_{\mathbb{R}^{d}}\left\|\tau_{tz}f-f\right\|_{\infty}\left|\varphi\left(z\right)\right|dz$$

<sup>&</sup>lt;sup>1</sup> The reader should construct the appropriate dominating function.

which again tends to zero by the dominated convergence theorem because  $\lim_{t\downarrow 0} \|\tau_{tz}f - f\|_{\infty} = 0$  uniformly in z by the uniform continuity of f.

Item 3. Let  $B_R = B(0, R)$  be a large ball in  $\mathbb{R}^d$  and  $K \sqsubset U$ , then

$$\begin{split} \sup_{x \in K} & \left| \varphi_t * f\left( x \right) - af\left( x \right) \right| \\ & \leq \left| \int_{B_R} \left[ f(x - tz) - f\left( x \right) \right] \varphi\left( z \right) dz \right| + \left| \int_{B_R^c} \left[ f(x - tz) - f\left( x \right) \right] \varphi\left( z \right) dz \right| \\ & \leq \int_{B_R} \left| \varphi\left( z \right) \right| dz \cdot \sup_{x \in K, z \in B_R} \left| f(x - tz) - f\left( x \right) \right| + 2 \left\| f \right\|_{\infty} \int_{B_R^c} \left| \varphi\left( z \right) \right| dz \\ & \leq \left\| \varphi \right\|_1 \cdot \sup_{x \in K, z \in B_R} \left| f(x - tz) - f\left( x \right) \right| + 2 \left\| f \right\|_{\infty} \int_{|z| > R} \left| \varphi\left( z \right) \right| dz \end{split}$$

so that using the uniform continuity of f on compact subsets of U,

$$\limsup_{t \downarrow 0} \sup_{x \in K} \left| \varphi_t * f(x) - af(x) \right| \le 2 \left\| f \right\|_{\infty} \int_{|z| > R} \left| \varphi(z) \right| dz \to 0 \text{ as } R \to \infty.$$

**Exercise 4.1 (Similar to Exercise ??.).** Let  $p \in [1, \infty]$  and  $\|\tau_z - I\|_{L(L^p(m))}$  be the operator norm  $\tau_z - I$ . Show  $\|\tau_z - I\|_{L(L^p(m))} = 2$  for all  $z \in \mathbb{R}^d \setminus \{0\}$  and conclude from this that  $z \in \mathbb{R}^d \to \tau_z \in L(L^p(m))$  is **not** continuous.

**Hints:** 1) Show  $\|\tau_z - I\|_{L(L^p(m))} = \|\tau_{|z|e_1} - I\|_{L(L^p(m))}$ . 2) Let  $z = te_1$  with t > 0 and look for  $f \in L^p(m)$  such that  $\tau_z f$  is approximately equal to -f. (In fact, if  $p = \infty$ , you can find  $f \in L^{\infty}(m)$  such that  $\tau_z f = -f$ .) (BRUCE: add on a problem somewhere showing that spec  $(\tau_z) = S^1 \subset \mathbb{C}$ . This is very simple to prove if p = 2 by using the Fourier transform.)

**Proposition 4.7.** Suppose  $p \in [1, \infty]$ ,  $f \in L^1$  and  $g \in L^p$ , then f \* g(x) exists for almost every x,  $f * g \in L^p$  and

$$\|f * g\|_p \le \|f\|_1 \|g\|_p.$$

**Proof.** This follows directly from Minkowski's inequality for integrals, Theorem ??, and was explained in Example ??.

**Definition 4.8.** Suppose that  $(X, \tau)$  is a topological space and  $\mu$  is a measure on  $\mathcal{B}_X = \sigma(\tau)$ . For a measurable function  $f : X \to \mathbb{C}$  we define the essential support of f by

$$\operatorname{supp}_{\mu}(f) = \{ x \in X : \mu(\{ y \in V : f(y) \neq 0 \} \}) > 0 \ \forall \ neighborhoods \ V \ of \ x \}.$$

$$(4.3)$$

Equivalently,  $x \notin \operatorname{supp}_{\mu}(f)$  iff there exists an open neighborhood V of x such that  $1_V f = 0$  a.e.

It is not hard to show that  $\underline{\text{if supp}}(\mu) = X$  (see Definition ??) and  $f \in C(X)$  then  $\underline{\text{supp}}_{\mu}(f) = \underline{\text{supp}}(f) := \overline{\{f \neq 0\}}$ , see Exercise ??.

**Lemma 4.9.** Suppose  $(X, \tau)$  is second countable and  $f : X \to \mathbb{C}$  is a measurable function and  $\mu$  is a measure on  $\mathcal{B}_X$ . Then  $X := U \setminus \operatorname{supp}_{\mu}(f)$  may be described as the largest open set W such that  $f1_W(x) = 0$  for  $\mu$ -a.e. x. Equivalently put,  $C := \operatorname{supp}_{\mu}(f)$  is the smallest closed subset of X such that  $f = f1_C$  a.e.

**Proof.** To verify that the two descriptions of  $\operatorname{supp}_{\mu}(f)$  are equivalent, suppose  $\operatorname{supp}_{\mu}(f)$  is defined as in Eq. (4.3) and  $W := X \setminus \operatorname{supp}_{\mu}(f)$ . Then

$$\begin{split} W &= \{ x \in X : \exists \ \tau \ni V \ni x \text{ such that } \mu(\{ y \in V : f(y) \neq 0 \} \}) = 0 \} \\ &= \cup \{ V \subset_o X : \mu \left( f 1_V \neq 0 \right) = 0 \} \\ &= \cup \{ V \subset_o X : f 1_V = 0 \text{ for } \mu\text{-a.e.} \} \,. \end{split}$$

So to finish the argument it suffices to show  $\mu(f1_W \neq 0) = 0$ . To to this let  $\mathcal{U}$  be a countable base for  $\tau$  and set

$$\mathcal{U}_f := \{ V \in \mathcal{U} : f \mathbb{1}_V = 0 \text{ a.e.} \}.$$

Then it is easily seen that  $W = \bigcup \mathcal{U}_f$  and since  $\mathcal{U}_f$  is countable

$$\mu\left(f1_W\neq 0\right)\leq \sum_{V\in\mathcal{U}_f}\mu\left(f1_V\neq 0\right)=0.$$

**Lemma 4.10.** Suppose  $f, g, h : \mathbb{R}^d \to \mathbb{C}$  are measurable functions and assume that x is a point in  $\mathbb{R}^d$  such that  $|f| * |g|(x) < \infty$  and  $|f| * (|g| * |h|)(x) < \infty$ , then

1. 
$$f * g(x) = g * f(x)$$
  
2.  $f * (g * h)(x) = (f * g) * h(x)$   
3. If  $z \in \mathbb{R}^d$  and  $\tau_z(|f| * |g|)(x) = |f| * |g|(x - z) < \infty$ , then

$$\tau_z(f*g)(x) = \tau_z f * g(x) = f * \tau_z g(x)$$

4. If  $x \notin \operatorname{supp}_{m}(f) + \operatorname{supp}_{m}(g)$  then f \* g(x) = 0 and in particular,

$$\operatorname{supp}_{m}(f * g) \subset \overline{\operatorname{supp}_{m}(f) + \operatorname{supp}_{m}(g)}$$

where in defining  $\operatorname{supp}_{m}(f*g)$  we will use the convention that " $f*g(x) \neq 0$ " when  $|f|*|g|(x) = \infty$ .

#### 22 4 Convolution and smoothing operators

**Proof.** For item 1.,

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

where in the second equality we made use of the fact that Lebesgue measure invariant under the transformation  $y \to x - y$ . Similar computations prove all of the remaining assertions of the first three items of the lemma. Item 4. Since  $f * g(x) = \tilde{f} * \tilde{g}(x)$  if  $f = \tilde{f}$  and  $g = \tilde{g}$  a.e. we may, by replacing f by  $f1_{\operatorname{supp}_m(f)}$  and g by  $g1_{\operatorname{supp}_m(g)}$  if necessary, assume that  $\{f \neq 0\} \subset \operatorname{supp}_m(f)$  and  $\{g \neq 0\} \subset \operatorname{supp}_m(g)$ . So if  $x \notin (\operatorname{supp}_m(f) + \operatorname{supp}_m(g))$  then  $x \notin (\{f \neq 0\} + \{g \neq 0\})$  and for all  $y \in \mathbb{R}^d$ , either  $x - y \notin \{f \neq 0\}$  or  $y \notin \{g \neq 0\}$ . That is to say either  $x - y \in \{f = 0\}$  or  $y \in \{g = 0\}$  and hence f(x - y)g(y) = 0 for all y and therefore f \* g(x) = 0. This shows that f \* g = 0 on  $\mathbb{R}^d \setminus (\overline{\operatorname{supp}_m(f) + \operatorname{supp}_m(g)})$  and therefore

$$\mathbb{R}^{d} \setminus \left( \overline{\operatorname{supp}_{m}(f) + \operatorname{supp}_{m}(g)} \right) \subset \mathbb{R}^{d} \setminus \operatorname{supp}_{m}(f * g),$$

i.e.  $\operatorname{supp}_{m}(f\ast g)\subset\operatorname{supp}_{m}\left(f\right)+\operatorname{supp}_{m}\left(g\right).$ 

Remark 4.11. Let A, B be closed sets of  $\mathbb{R}^d$ , it is not necessarily true that A+B is still closed. For example, take

$$A = \{(x,y) : x > 0 \text{ and } y \ge 1/x\} \text{ and } B = \{(x,y) : x < 0 \text{ and } y \ge 1/|x|\},\$$

then every point of A + B has a positive y - component and hence is not zero. On the other hand, for x > 0 we have  $(x, 1/x) + (-x, 1/x) = (0, 2/x) \in A + B$  for all x and hence  $0 \in \overline{A + B}$  showing A + B is not closed. Nevertheless if one of the sets A or B is compact, then A + B is closed again. Indeed, if A is compact and  $x_n = a_n + b_n \in A + B$  and  $x_n \to x \in \mathbb{R}^d$ , then by passing to a subsequence if necessary we may assume  $\lim_{n\to\infty} a_n = a \in A$  exists. In this case

$$\lim_{n \to \infty} b_n = \lim_{n \to \infty} (x_n - a_n) = x - a \in B$$

exists as well, showing  $x = a + b \in A + B$ .

#### 4.2 Young's Inequalities

**Theorem 4.12 (Young's Inequality).** Let  $p, q, r \in [1, \infty]$  satisfy

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}.$$
(4.4)

If  $f \in L^p$  and  $g \in L^q$  then  $|f| * |g|(x) < \infty$  for m-a.e. x and

$$\|f * g\|_{r} \le \|f\|_{p} \|g\|_{q}.$$
(4.5)

In particular  $L^1$  is closed under convolution. (The space  $(L^1, *)$  is an example of a "Banach algebra" without unit.) [See Section ?? for an interpolation proof of this theorem.]

**Proof.** By the usual sorts of arguments, we may assume f and g are positive functions. Let  $\alpha, \beta \in [0, 1]$  and  $p_1, p_2 \in (0, \infty]$  satisfy  $p_1^{-1} + p_2^{-1} + r^{-1} = 1$ . Then by Hölder's inequality, Corollary ??,

$$\begin{split} f * g (x) &= \int_{\mathbb{R}^d} \left[ f(x-y)^{(1-\alpha)} g(y)^{(1-\beta)} \right] f(x-y)^{\alpha} g(y)^{\beta} dy \\ &\leq \left( \int_{\mathbb{R}^d} f(x-y)^{(1-\alpha)r} g(y)^{(1-\beta)r} dy \right)^{1/r} \left( \int_{\mathbb{R}^d} f(x-y)^{\alpha p_1} dy \right)^{1/p_1} \times \\ &\qquad \times \left( \int_{\mathbb{R}^d} g(y)^{\beta p_2} dy \right)^{1/p_2} \\ &= \left( \int_{\mathbb{R}^d} f(x-y)^{(1-\alpha)r} g(y)^{(1-\beta)r} dy \right)^{1/r} \|f\|_{\alpha p_1}^{\alpha} \|g\|_{\beta p_2}^{\beta}. \end{split}$$

Taking the  $r^{\text{th}}$  power of this equation and integrating on x gives

$$\|f * g\|_{r}^{r} \leq \int_{\mathbb{R}^{d}} \left( \int_{\mathbb{R}^{d}} f(x - y)^{(1 - \alpha)r} g(y)^{(1 - \beta)r} dy \right) dx \cdot \|f\|_{\alpha p_{1}}^{\alpha} \|g\|_{\beta p_{2}}^{\beta}$$
$$= \|f\|_{(1 - \alpha)r}^{(1 - \alpha)r} \|g\|_{(1 - \beta)r}^{(1 - \beta)r} \|f\|_{\alpha p_{1}}^{\alpha r} \|g\|_{\beta p_{2}}^{\beta r}.$$
(4.6)

Let us now suppose,  $(1 - \alpha)r = \alpha p_1$  and  $(1 - \beta)r = \beta p_2$ , in which case Eq. (4.6) becomes,

$$\|f * g\|_{r}^{r} \le \|f\|_{\alpha p_{1}}^{r} \|g\|_{\beta p_{2}}^{r}$$

which is Eq. (4.5) with

$$p := (1 - \alpha)r = \alpha p_1 \text{ and } q := (1 - \beta)r = \beta p_2.$$
 (4.7)

So to finish the proof, it suffices to show p and q are arbitrary indices in  $[1,\infty]$  satisfying  $p^{-1} + q^{-1} = 1 + r^{-1}$ . If  $\alpha, \beta, p_1, p_2$  satisfy the relations above, then

$$\alpha = \frac{r}{r+p_1}$$
 and  $\beta = \frac{r}{r+p_2}$ 

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{\alpha p_1} + \frac{1}{\alpha p_2} = \frac{1}{p_1} \frac{r+p_1}{r} + \frac{1}{p_2} \frac{r+p_2}{r}$$
$$= \frac{1}{p_1} + \frac{1}{p_2} + \frac{2}{r} = 1 + \frac{1}{r}.$$

and

Conversely, if p, q, r satisfy Eq. (4.4), then let  $\alpha$  and  $\beta$  satisfy  $p = (1 - \alpha)r$  and  $q = (1 - \beta)r$ , i.e.

$$\alpha:=\frac{r-p}{r}=1-\frac{p}{r}\leq 1 \text{ and } \beta=\frac{r-q}{r}=1-\frac{q}{r}\leq 1.$$

Using Eq. (4.4) we may also express  $\alpha$  and  $\beta$  as

$$\alpha = p(1 - \frac{1}{q}) \ge 0$$
 and  $\beta = q(1 - \frac{1}{p}) \ge 0$ 

and in particular we have shown  $\alpha, \beta \in [0, 1]$ . If we now define  $p_1 := p/\alpha \in (0, \infty]$  and  $p_2 := q/\beta \in (0, \infty]$ , then

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{r} = \beta \frac{1}{q} + \alpha \frac{1}{p} + \frac{1}{r}$$
$$= (1 - \frac{1}{q}) + (1 - \frac{1}{p}) + \frac{1}{r}$$
$$= 2 - \left(1 + \frac{1}{r}\right) + \frac{1}{r} = 1$$

as desired.

Remark 4.13. Here is a scaling argument that explains why Eq. (4.4) is the only possible relationship for which Eq. (4.5) can hold. For  $\lambda > 0$ , let  $f_{\lambda}(x) := f(\lambda x)$ , then after a few simple change of variables we find

$$\|f_{\lambda}\|_{p} = \lambda^{-d/p} \|f\|$$
 and  $(f * g)_{\lambda} = \lambda^{d} f_{\lambda} * g_{\lambda}$ .

Therefore if Eq. (4.5) holds for some  $p, q, r \in [1, \infty]$ , we would also have

$$\|f\ast g\|_r = \lambda^{d/r} \left\| (f\ast g)_\lambda \right\|_r \le \lambda^{d/r} \lambda^d \left\| f_\lambda \right\|_p \left\| g_\lambda \right\|_q = \lambda^{(d+d/r-d/p-d/q)} \left\| f \right\|_p \left\| g \right\|_q$$

for all  $\lambda > 0$ . This is only possible if Eq. (4.4) holds.

#### 4.3 Convolution smoothing

We will often wish to take  $\varphi$  in Theorem 4.6 to be a smooth function with compact support. The existence of such functions is a simple consequence of the result of the next exercise, see Lemma 4.14.

Exercise 4.2. Let

$$f(t) = \begin{cases} e^{-1/t} \text{ if } t > 0\\ 0 \text{ if } t \le 0. \end{cases}$$

Show  $f \in C^{\infty}(\mathbb{R}, [0, 1])$ . **Hints:** you might start by first showing  $\lim_{t\downarrow 0} f^{(n)}(t) = 0$  for all  $n \in \mathbb{N}_0$ .

**Lemma 4.14 (Smooth bump functions).** There exists  $\varphi \in C_c^{\infty}(\mathbb{R}^d, [0, \infty))$ such that  $\varphi(0) > 0$ ,  $\operatorname{supp}(\varphi) \subset \overline{B}(0, 1)$  and  $\int_{\mathbb{R}^d} \varphi(x) \, dx = 1$ .

**Proof.** Define h(t) = f(1-t)f(t+1) where f is as in Exercise 4.2. Then  $h \in C_c^{\infty}(\mathbb{R}, [0, 1])$ ,  $\operatorname{supp}(h) \subset [-1, 1]$  and  $h(0) = e^{-2} > 0$ . Define  $c = \int_{\mathbb{R}^d} h(|x|^2) dx$ . Then  $\varphi(x) = c^{-1}h(|x|^2)$  is the desired function.

The reader asked to prove the following proposition in Exercise ?? below.

**Proposition 4.15.** Suppose that  $f \in L^1_{loc}(\mathbb{R}^d, m)$  and  $\varphi \in C^1_c(\mathbb{R}^d)$ , then  $f * \varphi \in C^1(\mathbb{R}^d)$  and  $\partial_i(f * \varphi) = f * \partial_i \varphi$ . Moreover if  $\varphi \in C^\infty_c(\mathbb{R}^d)$  then  $f * \varphi \in C^\infty(\mathbb{R}^d)$ .

The existence of smooth bump functions along with Proposition 4.15 allows us to construct smooth functions approximating most any function we like. Here are some useful results along this vein.

**Corollary 4.16.** Let  $X \subset \mathbb{R}^d$  be an open set and  $\mu$  be a K-finite measure on  $\mathcal{B}_X$ .

1. Then  $C_c^{\infty}(X)$  is dense in  $L^p(\mu)$  for all  $1 \leq p < \infty$ . 2. If  $h \in L^1_{loc}(\mu)$  satisfies

$$\int_X fhd\mu = 0 \text{ for all } f \in C_c^\infty(X)$$
(4.8)

then h(x) = 0 for  $\mu$ -a.e. x.

**Proof.** Let  $f \in C_c(X)$ ,  $\varphi$  be as in Lemma 4.14,  $\varphi_t$  be as in Theorem 4.6 and set  $\psi_t := \varphi_t * (f1_X)$ . Then by Proposition 4.15  $\psi_t \in C^{\infty}(X)$  and by Lemma 4.10 there exists a compact set  $K \subset X$  such that  $\operatorname{supp}(\psi_t) \subset K$  for all t sufficiently small. By Theorem 4.6,  $\psi_t \to f$  uniformly on X as  $t \downarrow 0$ 

- 1. The dominated convergence theorem (with dominating function being  $||f||_{\infty} 1_K$ ), shows  $\psi_t \to f$  in  $L^p(\mu)$  as  $t \downarrow 0$ . This proves Item 1., since Theorem ?? guarantees that  $C_c(X)$  is dense in  $L^p(\mu)$ .
- 2. Keeping the same notation as above, the dominated convergence theorem (with dominating function being  $\|f\|_{\infty} |h| \mathbf{1}_K$ ) implies

$$0 = \lim_{t\downarrow 0} \int_X \psi_t h d\mu = \int_X \lim_{t\downarrow 0} \psi_t h d\mu = \int_X f h d\mu.$$

The proof is now finished by an application of Lemma ??. Alternatively: Let  $\{\varphi_t\}_{t>0}$  be an approximate  $\delta$ -sequence as above and

for  $f \in C_c(X)$  and K = supp(f) we will have  $\varphi_t * f \in C_c^{\infty}(\mathbb{R}^d)$  with

$$\operatorname{supp}\left(\varphi_{t} \ast f\right) \subset K_{t} = \left\{x \in \mathbb{R}^{d} : d_{K}\left(x\right) \leq t\right\}$$

#### 24 4 Convolution and smoothing operators

where  $K_t \subset X$  for t small enough. By assumption and the dominated convergence theorem we find,

$$\int_{X} fhd\mu = \lim_{t \downarrow 0} \int_{X} \varphi_{t} * f \cdot hd\mu = 0 \text{ for all } f \in C_{c}\left(X\right).$$

Now choose open sets,  $V_n \subset X$ , such that  $\overline{V}_n$  is a compact subset of X and  $V_n \uparrow X$  as  $n \to \infty$  and define  $d\nu_n = 1_{V_n} h d\mu$  – a complex measure. Then for any  $f \in C_c(V_n)$  we have

$$\int_X f d\nu_n = \int_X f \mathbf{1}_{V_n} h d\mu = \int_X f h d\mu = 0$$

and therefore  $\nu_n \equiv 0$  from which it follows that

$$0 = |\nu_n| (X) = \int_X 1_{V_n} |h| \, d\mu.$$

Letting  $n \to \infty$  then shows  $0 = \int_X |h| d\mu$ , i.e.  $h = 0 \mu$  a.e.

**Exercise 4.3.** Show  $C_c^{\infty}(\mathbb{R}^d)$  is dense in  $L^p(\mathbb{R}^d, m)$  for any  $1 \leq p < \infty$ .

**Lemma 4.17.** Given a rectangle R in  $\mathbb{R}^d$ , say  $R = [a_1, b_1) \times \cdots \times [a_n, b_n)$ , then there exists  $f_k \in C_c^{\infty}(\mathbb{R}^d)$  such that  $f_k \to 1_R$  boundedly.

**Proof.** It suffices to consider the one dimensional case. Let  $\varphi \in C_c^{\infty}(\mathbb{R})$  such that  $\varphi \geq 0$ ,  $\varphi$  is supported in (-1, 0) and  $\int_{\mathbb{R}} \varphi(x) dx = 1$ . Set  $\varphi_{\varepsilon}(x) = \frac{1}{\varepsilon} \varphi(\frac{x}{\varepsilon})$ . Then

$$\begin{split} \varphi_{\varepsilon} * \mathbf{1}_{[a,b)} \left( x \right) &= \int_{\mathbb{R}} \varphi_{\varepsilon} \left( y \right) \mathbf{1}_{[a,b)} \left( x - y \right) dy = \int_{\mathbb{R}} \varphi \left( y \right) \mathbf{1}_{[a,b)} \left( x - \varepsilon y \right) dy \\ &= \int_{-1}^{0} \varphi \left( y \right) \mathbf{1}_{[a,b)} \left( x - \varepsilon y \right) dy \to \mathbf{1}_{[a,b)} \left( x \right) \text{ as } \varepsilon \downarrow 0 \end{split}$$

for all  $x \in \mathbb{R}$ .

4

**Corollary 4.18** ( $C^{\infty}$  – Uryshon's Lemma). Given  $K \sqsubset U \subset_o \mathbb{R}^d$ , there exists  $f \in C_c^{\infty}(\mathbb{R}^d, [0, 1])$  such that supp  $(f) \subset U$  and f = 1 on K.

**Proof.** Let d be the standard metric on  $\mathbb{R}^d$  and  $\varepsilon := d(K, U^c)$  which is positive since K is compact and  $d(x, U^c) > 0$  for all  $x \in K$ . Further let  $V := \{x \in \mathbb{R}^d : d(x, K) < \varepsilon/3\}$  and then take  $f = \varphi_{\varepsilon/3} * 1_V$  where  $\varphi_t(x) = t^{-d}\varphi(x/t)$  as in Theorem 4.6 and  $\varphi$  is as in Lemma 4.14. It then follows that

$$\operatorname{supp}(f) \subset \overline{\operatorname{supp}(\varphi_{\varepsilon/3}) + V_{\varepsilon/3}} \subset \overline{V}_{2\varepsilon/3} \subset U.$$

Since  $\overline{V}_{2\varepsilon/3}$  is closed and bounded,  $f \in C_c^{\infty}(U)$  and for  $x \in K$ ,

$$f(x) = \int_{\mathbb{R}^d} \mathbf{1}_{d(y,K) < \varepsilon/3} \cdot \varphi_{\varepsilon/3}(x-y) dy = \int_{\mathbb{R}^d} \varphi_{\varepsilon/3}(x-y) dy = 1.$$

The proof will be finished after the reader (easily) verifies  $0 \le f \le 1$ .

Here is an application of this Corollary 4.18 whose proof is left to the reader, Exercise **??**.

**Lemma 4.19 (Integration by Parts).** Suppose f and g are measurable functions on  $\mathbb{R}^d$  such that  $t \to f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_d)$  and  $t \to g(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_d)$  are continuously differentiable functions on  $\mathbb{R}$  for each fixed  $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ . Moreover assume  $f \cdot g$ ,  $\frac{\partial f}{\partial x_i} \cdot g$  and  $f \cdot \frac{\partial g}{\partial x_i}$  are in  $L^1(\mathbb{R}^d, m)$ . Then

$$\int_{\mathbb{R}^d} \frac{\partial f}{\partial x_i} \cdot g dm = - \int_{\mathbb{R}^d} f \cdot \frac{\partial g}{\partial x_i} dm$$

With this result we may give another proof of the Riemann Lebesgue Lemma.

Lemma 4.20 (Riemann Lebesgue Lemma). For  $f \in L^1(\mathbb{R}^d, m)$  let

$$\hat{f}(\xi) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-i\xi \cdot x} dm(x)$$

be the Fourier transform of f. Then  $\hat{f} \in C_0(\mathbb{R}^d)$  and  $\left\|\hat{f}\right\|_{\infty} \leq (2\pi)^{-d/2} \|f\|_1$ . (The choice of the normalization factor,  $(2\pi)^{-d/2}$ , in  $\hat{f}$  is for later convenience.)

**Proof.** The fact that  $\hat{f}$  is continuous is a simple application of the dominated convergence theorem. Moreover,

$$\left| \hat{f}(\xi) \right| \le \int_{\mathbb{R}^d} \left| f(x) \right| dm(x) \le (2\pi)^{-d/2} \left\| f \right\|_{2}$$

so it only remains to see that  $\hat{f}(\xi) \to 0$  as  $|\xi| \to \infty$ . First suppose that  $f \in C_c^{\infty}(\mathbb{R}^d)$  and let  $\Delta = \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2}$  be the Laplacian on  $\mathbb{R}^d$ . Notice that  $\frac{\partial}{\partial x_j}e^{-i\xi \cdot x} = -i\xi_j e^{-i\xi \cdot x}$  and  $\Delta e^{-i\xi \cdot x} = -|\xi|^2 e^{-i\xi \cdot x}$ . Using Lemma 4.19 repeatedly,

$$\begin{split} \int_{\mathbb{R}^d} \Delta^k f\left(x\right) e^{-i\xi \cdot x} dm\left(x\right) &= \int_{\mathbb{R}^d} f\left(x\right) \Delta^k_x e^{-i\xi \cdot x} dm\left(x\right) = -\left|\xi\right|^{2k} \int_{\mathbb{R}^d} f\left(x\right) e^{-i\xi \cdot x} dm\left(x\right) \\ &= -(2\pi)^{d/2} \left|\xi\right|^{2k} \hat{f}(\xi) \end{split}$$

job: Supplements

for any  $k \in \mathbb{N}$ . Hence

$$(2\pi)^{d/2} \left| \widehat{f}(\xi) \right| \le \left| \xi \right|^{-2k} \left\| \varDelta^k f \right\|_1 \to 0$$

as  $|\xi| \to \infty$  and  $\hat{f} \in C_0(\mathbb{R}^d)$ . Suppose that  $f \in L^1(m)$  and  $f_k \in C_c^{\infty}(\mathbb{R}^d)$  is a sequence such that  $\lim_{k\to\infty} \|f - f_k\|_1 = 0$ , then  $\lim_{k\to\infty} \|\hat{f} - \hat{f}_k\|_{\infty} = 0$ . Hence  $\hat{f} \in C_0(\mathbb{R}^d)$  by an application of Proposition ??.

The next two results give a version of Theorem 4.6 where the convergence holds almost everywhere by making use of the Lebesgue differentiation Theorem ??. Recall for  $f \in L^1_{loc}(\mathbb{R}^d)$  that, by Theorem ??, the **Lebesgue set** of f,

$$\mathcal{L}(f) := \left\{ x \in \mathbb{R}^d : \lim_{r \downarrow 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)| \, dy = 0 \right\},$$

is a set of full Lebesgue measure, i.e.  $m\left(\mathbb{R}^{d} \setminus \mathcal{L}(f)\right) = 0.$ 

**Proposition 4.21 (Theorem 4.6 continued).** Let  $p \in [1, \infty)$ ,  $\rho > 0$  and  $\varphi \in L^{\infty}(\mathbb{R}^d)$  such that  $0 \leq \varphi \leq C1_{B(0,\rho)}$  for some  $C < \infty$  and  $\int_{\mathbb{R}^d} \varphi(x) dx = 1$ . If  $f \in L^1_{loc}(m)$ , and  $x \in \mathcal{L}(f)$ , then

$$\lim_{t\downarrow 0} \left(\varphi_t * f\right)(x) = f(x),$$

where  $\varphi_t(x) := t^{-d}\varphi(x/t)$ . In particular,  $\varphi_t * f \to f$  a.e. as  $t \downarrow 0$ .

**Proof.** Notice that  $0 \leq \varphi_t \leq Ct^{-d} \mathbf{1}_{B(0,\rho t)}$  and therefore for  $x \in \mathcal{L}(f)$  we have, using Theorem ??, that

$$\begin{aligned} |\varphi_t * f(x) - f(x)| &= \left| \int_{\mathbb{R}^d} \left[ f(x - y) - f(x) \right] \varphi_t(y) \, dy \right| \\ &\leq \int_{\mathbb{R}^d} |f(x - y) - f(x)| \, \varphi_t(y) \, dy \\ &\leq Ct^{-d} \int_{B(0,\rho t)} |f(x - y) - f(x)| \, dy \\ &= C\left(\rho, d\right) \frac{1}{|B\left(0, \rho t\right)|} \int_{B(0,\rho t)} |f(x - y) - f(x)| \, dy \to 0 \text{ as } t \downarrow 0. \end{aligned}$$

The following theorem is an extension of Proposition 4.21.

**Theorem 4.22 (\* Theorem 8.15 of Folland).** More general version, assume that  $|\varphi(x)| \leq C (1 + |x|)^{-(d+\varepsilon)}$  and  $\int_{\mathbb{R}^d} \varphi(x) dx = a$ . Then for all  $x \in \mathcal{L}(f)$ ,

and in fact,

$$L(x) := \limsup_{t \downarrow 0} \int |f(x-y) - f(x)| |\varphi_t(y)| dy = 0.$$

 $\lim_{t \to 0} \left(\varphi_t * f\right)(x) = af(x)$ 

**Proof.** Throughout this proof  $f \in L^1(\mathbb{R}^d)$  and  $x \in \mathcal{L}(f)$  be fixed and for b > 0 let

$$\delta(b) := \frac{1}{b^d} \int_{|y| \le b} |f(x-y) - f(x)| \, dy$$

From the definition if  $\mathcal{L}(f)$  we know that  $\lim_{b\downarrow 0} \delta(b) = 0$ . The remainder of the proof will be broken into a number of steps.

**1.** For any  $\eta > 0$ ,

$$L(x) = \limsup_{t \downarrow 0} \int_{|y| \le \eta} |f(x-y) - f(x)| |\varphi_t(y)| dy$$

which is seen as follows;

$$\begin{split} \int_{|y|>\eta} \left| f\left(x-y\right) - f\left(x\right) \right| \left| \varphi_t\left(y\right) \right| dy \\ &\leq \int_{|y|>\eta} \left| f\left(x-y\right) \right| \left| \varphi_t\left(y\right) \right| dy + \left| f\left(x\right) \right| \int_{|y|>\eta} \left| \varphi_t\left(y\right) \right| dy \\ &\leq Ct^{-n} \int_{|y|>\eta} \left| f\left(x-y\right) \right| \left(\frac{1}{1+|y|/t}\right)^{n+\varepsilon} dy + \left| f\left(x\right) \right| \int_{|z|>\eta/t} \left| \varphi\left(z\right) \right| dy \\ &\leq \frac{Ct^{\varepsilon}}{\left(t+\eta\right)^{n+\varepsilon}} \left\| f \right\|_1 + \left| f\left(x\right) \right| \int_{|z|>\eta/t} \left| \varphi\left(z\right) \right| dy \to 0 \text{ as } t \downarrow 0. \end{split}$$

**2.** For any  $\rho > 0$ ,

$$\begin{split} \int_{|y| \le \rho} \left| f\left(x - y\right) - f\left(x\right) \right| \left| \varphi_t\left(y\right) \right| dy &= t^{-d} \int_{|y| \le \rho} \left| f\left(x - y\right) - f\left(x\right) \right| \left| \varphi\left(y/t\right) \right| dy \\ &\le C t^{-d} \delta\left(\rho\right) \cdot \rho^d = C \delta\left(\rho\right) \cdot \left(\frac{\rho}{t}\right)^d. \end{split}$$

In particular  $\rho \leq kt$  for some k, then

$$\int_{|y| \le \rho} |f(x-y) - f(x)| |\varphi_t(y)| \, dy \le Ck^d \delta(kt) \to 0 \text{ as } t \downarrow 0.$$

**3.** Given items 1. and 2., in order to finish the proof we must estimate the integral over the annular region  $\{y \in \mathbb{R}^d : kt \le |y| \le \eta\}$ . In order to control this

Page: 25

#### 26 4 Convolution and smoothing operators

integral we are going to have to divide this annular region up into a number of concentric annular regions which we will do shortly. For the moment, let  $0 < a < b < \infty$  be given, then

$$\begin{split} \int_{a<|y|\leq b} |f\left(x-y\right)-f\left(x\right)| \left|\varphi_{t}\left(y\right)\right| dy \\ &\leq Ct^{-d} \int_{a<|y|\leq b} |f\left(x-y\right)-f\left(x\right)| \left(1+\left|\frac{y}{t}\right|\right)^{-(d+\varepsilon)} dy \\ &\leq Ct^{-d} \int_{a<|y|\leq b} |f\left(x-y\right)-f\left(x\right)| \left(1+\frac{a}{t}\right)^{-(d+\varepsilon)} dy \\ &\leq Ct^{-d} \delta\left(b\right) b^{d} \left(1+\frac{a}{t}\right)^{-(d+\varepsilon)} \\ &= Ct^{-(d+\varepsilon)} \delta\left(b\right) b^{(d+\varepsilon)} \left(1+\frac{a}{t}\right)^{-(d+\varepsilon)} t^{\varepsilon} b^{-\varepsilon} \\ &= C\delta\left(b\right) \left(\frac{t}{b}\right)^{\varepsilon} \frac{1}{\left(t+\frac{a}{b}\right)^{d+\varepsilon}}. \end{split}$$

Taking a = b/2 in this expression shows,

$$\int_{\frac{b}{2} < |y| \le b} |f(x-y) - f(x)| \cdot |\varphi_t(y)| \, dy$$
$$\le C\delta(b) \left(\frac{t}{b}\right)^{\varepsilon} \frac{1}{\left(t + \frac{1}{2}\right)^{d+\varepsilon}}$$
$$= C\delta(b) \left(\frac{t}{b}\right)^{\varepsilon} \frac{1}{\left(2t + 1\right)^{d+\varepsilon}}.$$

Taking  $b = 2^{-k}\eta$  and summing the result on  $0 \le k \le K - 1$  shows

$$\begin{split} \sum_{k=0}^{K-1} \int_{2^{-(k+1)}\eta < |y| \le 2^{-k}\eta} \left| f\left(x-y\right) - f\left(x\right) \right| \left| \varphi_t\left(y\right) \right| dy \\ & \le C \sum_{k=0}^{K-1} \delta\left(2^{-k}\eta\right) \left(\frac{t}{2^{-k}\eta}\right)^{\varepsilon} \frac{1}{(2t+1)^{d+\varepsilon}} \\ & = \frac{C\delta\left(\eta\right)}{(2t+1)^{d+\varepsilon}} \left(\frac{t}{\eta}\right)^{\varepsilon} \sum_{k=0}^{K-1} 2^{\varepsilon k} \\ & = \frac{C\delta\left(\eta\right)}{(2t+1)^{d+\varepsilon}} \left(\frac{t}{\eta}\right)^{\varepsilon} \frac{2^{\varepsilon K} - 1}{2^{\varepsilon} - 1}. \end{split}$$

We now choose K so that  $2^K \frac{t}{\eta} \sim 1$  (i.e.  $2^{-K} \eta \sim t)$  and we have shown,

$$\int_{2^{-K}\eta < |y| \le \eta} |f(x-y) - f(x)| |\varphi_t(y)| dy$$
  
=  $\sum_{k=0}^{K-1} \int_{2^{-(k+1)}\eta < |y| \le 2^{-k}\eta} |f(x-y) - f(x)| |\varphi_t(y)| dy \le C\delta(\eta).$ 

4. Combining item 2. with  $\rho = 2^{-K} \eta \sim t$  with item 3. shows

$$\int_{|y| \le \eta} |f(x-y) - f(x)| |\varphi_t(y)| \, dy \le C\delta(\eta) \, .$$

Combining this result with item 1. implies,

$$\begin{split} L\left(x\right) &= \limsup_{t \downarrow 0} \int_{|y| \le \eta} \left| f\left(x - y\right) - f\left(x\right) \right| \left| \varphi_t\left(y\right) \right| dy \\ &\le C\delta\left(\eta\right) \to 0 \text{ as } \eta \downarrow 0. \end{split}$$

## Fourier Transform

#### 5.1 Motivation

Our first goal is to motivate the Fourier inversion formula from the inversion formula for Fourier series (see Exercise ?? below for more details). To do this, for L > 0, let  $H_L := L^2([-\pi L, \pi L])$  be the  $L^2$ -Hilbert space equipped with the inner product

$$\langle f|g\rangle_L := \frac{1}{2\pi L} \int_{[-\pi L,\pi L]} f(x) \,\bar{g}(x) \,dx.$$

The linear map,  $U_L: H_1 \to H_L$  defined by

$$(U_L f)(x) := f(L^{-1}x)$$
 for  $f \in H_1$ 

is unitary since

$$|U_L f||_L^2 = \frac{1}{2\pi L} \int_{[-\pi L, \pi L]} |f(L^{-1}x)|^2 dx = \frac{1}{2\pi} \int_{[-\pi, \pi]} |f(\theta)|^2 d\theta = ||f||_1^2$$

Letting  $\varphi_{\lambda}(x) = e^{i\lambda \cdot x}$ , we know that  $\{\varphi_k\}_{k=-\infty}^{\infty}$  is an orthonormal basis for  $H_1$ and therefore  $\{\varphi_k^L := U_L \varphi_k = \varphi_{L^{-1}k}\}_{k=-\infty}^{\infty}$  is an orthonormal basis for  $H_L$ .

Suppose, for simplicity, that  $f \in C_c^1(\mathbb{R})$ . For sufficiently large L we will have for  $|x| \leq \pi L$  that

$$\begin{split} f\left(x\right) &= \sum_{k \in \mathbb{Z}} \left\langle f | \varphi_{k}^{L} \right\rangle_{L} \varphi_{k}^{L}\left(x\right) \\ &= \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \frac{1}{L} \left( \frac{1}{\sqrt{2\pi}} \int_{\left[-\pi L, \pi L\right]} f\left(y\right) e^{-iky/L} dy \right) \varphi_{k}^{L}\left(x\right) \\ &= \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \frac{1}{L} \hat{f}\left(\frac{k}{L}\right) e^{ikx/L} \end{split}$$
(5.1)

where  $\hat{f}$  is the **Fourier transform** of f defined by,

$$\hat{f}\left(\xi\right) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f\left(y\right) e^{-i\xi y} dy.$$

Moreover,

$$\begin{split} \left|f\right|_{L^{2}(m)}^{2} &= 2\pi L \left\langle f|f\right\rangle_{L} = 2\pi L \sum_{k\in\mathbb{Z}} \left|\left\langle f|\varphi_{k}^{L}\right\rangle_{L}\right|^{2} \\ &= \frac{1}{2\pi L} \sum_{k\in\mathbb{Z}} \left|\int_{[-\pi L,\pi L]} f\left(y\right) e^{-iky/L} dy\right|^{2} \\ &= \sum_{k\in\mathbb{Z}} \left|\hat{f}\left(\frac{k}{L}\right)\right|^{2} \frac{1}{L}. \end{split}$$
(5.2)

Formally passing to the limit in Eqs. (5.1) and (5.2) suggests that

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\xi) e^{i\xi x} d\xi \text{ and } \|f\|_{L^{2}(m)}^{2} = \left\|\hat{f}\right\|_{L^{2}(m)}^{2}$$

which leads one to suspect that the Fourier transform,  $f \to \hat{f}$ , is a unitary operator on  $L^2(\mathbb{R})$ . We will eventually show this is the case after first showing how to interpret  $\hat{f}$  for  $f \in L^2(\mathbb{R})$ .

**Exercise 5.1 (Wirtinger's inequality, Folland 8.18).** Given a > 0 and  $f \in C^1([0, a], \mathbb{C})$  such that f(0) = f(a) = 0, show<sup>1</sup>

$$\int_{0}^{a} |f(x)|^{2} dx \leq \left(\frac{a}{\pi}\right)^{2} \int_{0}^{a} |f'(x)|^{2} dx$$

**Hint:** to use the notation above, let  $\pi L = a$  and extend f to [-a, 0] by setting f(-x) = -f(x) for  $0 \le x \le a$ . Now compute  $\int_0^a |f(x)|^2 dx$  and  $\int_0^a |f'(x)|^2 dx$  in terms of their Fourier coefficients,  $\langle f | \varphi_k^L \rangle_L$  and  $\langle f' | \varphi_k^L \rangle_L$  respectively.

We now generalize to the d – dimensionsal case. The underlying space in this section is  $\mathbb{R}^d$  with Lebesgue measure. As suggested above, the Fourier inversion formula is going to state that

$$f(x) = \left(\frac{1}{2\pi}\right)^d \int_{\mathbb{R}^d} d\xi e^{i\xi \cdot x} \left[ \int_{\mathbb{R}^d} dy f(y) e^{-iy \cdot \xi} \right].$$
 (5.3)

If we let  $\xi = 2\pi\eta$ , this may be written as

<sup>&</sup>lt;sup>1</sup> This inequality is sharp as is seen by taking  $f(x) = \sin(\pi x/a)$ .

#### 285 Fourier Transform

$$f(x) = \int_{\mathbb{R}^d} d\eta e^{i2\pi\eta \cdot x} \int_{\mathbb{R}^d} dy f(y) e^{-i2\pi y \cdot \eta}$$

and we have removed the multiplicative factor of  $\left(\frac{1}{2\pi}\right)^d$  in Eq. (5.3) at the expense of placing factors of  $2\pi$  in the arguments of the exponentials. [This is what Folland does.] Another way to avoid writing the  $2\pi$ 's altogether is to redefine dx and  $d\xi$  and this is what we will do here.

Notation 5.1 Let m be Lebesgue measure on  $\mathbb{R}^d$ ,  $c = c_d = (2\pi)^{-d/2}$ , and define:  $d\lambda$ 

$$\lambda(x) := \mathbf{d}x := c_d dm(x) \text{ and } \mathbf{d}\xi := c_d dm(\xi).$$

To be consistent with this new normalization of Lebesgue measure we will redefine  $||f||_p$ , and  $\langle f, g \rangle$ , as

$$\|f\|_{p} = \left(\int_{\mathbb{R}^{d}} |f(x)|^{p} \, \mathrm{d}x\right)^{1/p} = \left(\left(\frac{1}{2\pi}\right)^{d/2} \int_{\mathbb{R}^{d}} |f(x)|^{p} \, dm(x)\right)^{1/p},$$

and

$$\left\langle f,g
ight
angle :=\int_{\mathbb{R}^{d}}f\left(x
ight)g\left(x
ight)\mathbf{d}x\ when\ fg\in L^{1}.$$

We also define

$$\langle f|g \rangle = \langle f, \bar{g} \rangle = \int_{\mathbb{R}^d} f(x) g(x) \, \mathbf{d}x \text{ when } fg \in L^1$$

and a renormalized convolution by  $f \star \mathbf{g} := c_d \cdot f * g$ , *i.e.* 

$$f \bigstar g(x) = \int_{\mathbb{R}^d} f(x-y)g(y) \, \mathrm{d}y = \int_{\mathbb{R}^d} f(x-y)g(y) \, c_d dm(y)$$

The following notation will also be convenient; given a multi-index  $\alpha \in \mathbb{Z}_+^d$ , let  $|\alpha| = \alpha_1 + \cdots + \alpha_d$ ,

$$\begin{aligned} x^{\alpha} &:= \prod_{j=1}^{d} x_{j}^{\alpha_{j}}, \ \partial_{x}^{\alpha} = \left(\frac{\partial}{\partial x}\right)^{\alpha} := \prod_{j=1}^{d} \left(\frac{\partial}{\partial x_{j}}\right)^{\alpha_{j}} \text{ and} \\ D_{x}^{\alpha} &= \left(\frac{1}{i}\right)^{|\alpha|} \left(\frac{\partial}{\partial x}\right)^{\alpha} = \left(\frac{1}{i}\frac{\partial}{\partial x}\right)^{\alpha}. \end{aligned}$$

When  $x \in \mathbb{R}^d$  we let  $|x| = \sqrt{\sum_{j=1}^d x_j^2}$  (which is inconsistent with  $|\alpha|$  for  $\alpha \in \mathbb{Z}_+^d$ ) and further let

$$\langle x \rangle := (1 + |x|^2)^{1/2}$$
 and  $\nu_s (x) = (1 + |x|)^s$  for  $s \in \mathbb{R}$ .

#### 5.2 Fourier Transform formal development

**Definition 5.2 (Fourier Transform).** For  $f \in L^1(\mathbb{R}^d, m)$ , let

$$\hat{f}(\xi) = (\mathcal{F}f)(\xi) := \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x) \, \mathrm{d}x = c_d \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x) \, dm(x) \qquad (5.4)$$

and let  $f^{\vee}(\xi) = (\mathcal{F}^{-1}f)(\xi) := \hat{f}(-\xi)$ , that is

$$f^{\vee}(\xi) = \left(\mathcal{F}^{-1}f\right)(\xi) = \int_{\mathbb{R}^d} e^{ix\cdot\xi} f(\xi) \,\mathrm{d}\xi = c_d \int_{\mathbb{R}^d} e^{ix\cdot\xi} f(\xi) \,dm\left(\xi\right), \qquad (5.5)$$

where as above.

$$c_d = \left(\frac{1}{\sqrt{2\pi}}\right)^d$$

Remark 5.3. For  $f \in L^1(m)$ , then both  $\hat{f}$  and  $f^{\vee}$  are in  $C_0(\mathbb{R}^d)$ . The vanishing at infinity is exactly the Riemann Lebesgue lemma and the continuity easily follows from the dominated convergence theorem.

Example 5.4. If  $-\infty < a < b < \infty$ , then

$$\hat{1}_{[a,b]}(k) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} \frac{e^{-ikx}}{-ik} \Big|_{x=a}^{b} = \frac{1}{\sqrt{2\pi}} \frac{e^{-ika} - e^{-ikb}}{ik}.$$

If b > 0 and a = -b, this reduces to

$$\hat{1}_{[-b,b]}(k) = \frac{2}{\sqrt{2\pi}} \frac{\sin(kb)}{k}.$$

Notice in this case that  $\hat{1}_{[a,b]} \in L^2(m) \setminus L^1(m)$ .

Lemma 5.5 (Transforms of Gaussians). For a > 0 and  $\lambda \in \mathbb{R}$ ,

$$\int_{\mathbb{R}} e^{-\frac{1}{2}ax^2} e^{i\lambda x} dx = \sqrt{\frac{2\pi}{a}} e^{-\frac{1}{2a}\lambda^2}$$

**Proof.** Letting

$$f(\lambda) := \int_{\mathbb{R}} e^{-\frac{1}{2}ax^2} e^{i\lambda x} dx,$$

we have

$$f'(\lambda) = i \int_{\mathbb{R}} x e^{-\frac{1}{2}ax^2} e^{i\lambda x} dx = -\frac{i}{a} \int_{\mathbb{R}} \frac{d}{dx} e^{-\frac{1}{2}ax^2} e^{i\lambda x} dx$$
$$= \frac{i}{a} \int_{\mathbb{R}} e^{-\frac{1}{2}ax^2} \frac{d}{dx} e^{i\lambda x} dx = -\frac{1}{a} \lambda f(\lambda) \,.$$

Page: 28

Solving this equation for  $f(\lambda)$  shows,

$$f\left(\lambda\right) = e^{-\frac{\lambda^2}{a}}f\left(0\right)$$

where (upon letting  $y = \sqrt{ax}$ )

$$f(0) = \int_{\mathbb{R}} e^{-\frac{1}{2}ax^2} dx = \int_{\mathbb{R}} e^{-\frac{1}{2}y^2} \frac{dy}{\sqrt{a}} = \sqrt{\frac{2\pi}{a}}$$

and the proof is complete.

**Corollary 5.6.** If a > 0 and  $\lambda \in \mathbb{R}^d$ , then

$$\int_{\mathbb{R}} e^{-\frac{1}{2}a|x|^2} e^{i\lambda \cdot x} dx = \left(\frac{2\pi}{a}\right)^{d/2} e^{-\frac{1}{2a}|\lambda|^2}.$$

**Corollary 5.7.** For t > 0 and  $x \in \mathbb{R}^d$  we let

$$p_t(x) := t^{-d/2} e^{-\frac{1}{2t}|x|^2}, \tag{5.6}$$

then

$$\widehat{p}_t(\xi) = e^{-\frac{t}{2}|\xi|^2} \text{ and } (\widehat{p}_t)^{\vee}(x) = p_t(x).$$
 (5.7)

**Theorem 5.8 (Fourier Inversion Theorem 1).** Suppose that  $f \in L^1$  and  $\hat{f} \in L^{1,2}$  then

1. there exists 
$$f_0 \in C_0(\mathbb{R}^d)$$
 such that  $f = f_0$  a.e.,  
2.  $f_0 = \mathcal{F}^{-1}\mathcal{F}f$  and  $f_0 = \mathcal{F}\mathcal{F}^{-1}f$ ,  
3.  $f$  and  $\hat{f}$  are in  $L^1 \cap L^\infty$  and  
4.  $\|f\|_2 = \|\hat{f}\|_2$ .

In particular,  $\mathcal{F} : \mathcal{S} \to \mathcal{S}$  is a linear isomorphism of vector spaces. [This comment is now out of place.]

**Proof.** First notice that  $\hat{f} \in C_0(\mathbb{R}^d) \subset L^\infty$  and  $\hat{f} \in L^1$  by assumption, so that  $\hat{f} \in L^1 \cap L^\infty$ . Define  $f_0 := \hat{f}^{\vee} \in C_0(\mathbb{R}^d)$ , then

$$f_0(x) = (\hat{f})^{\vee}(x) = \left(\frac{1}{2\pi}\right)^{d/2} \int_{\mathbb{R}^d} dk \hat{f}(k) e^{ik \cdot x}$$
$$= \lim_{a \downarrow 0} \left(\frac{1}{2\pi}\right)^{d/2} \int_{\mathbb{R}^d} dk \hat{f}(k) e^{ik \cdot x} e^{-\frac{a}{2}|k|^2} \text{ by DCT.}$$

<sup>2</sup> We will see shortly that  $\hat{f}$  will be in  $L^{1}(m)$  provided f has sufficiently many derivatives on  $L^{1}(m)$ .

For fixed a > 0 we have

$$\begin{split} \int_{\mathbb{R}^d} dk \hat{f}\left(k\right) e^{ik \cdot x} e^{-\frac{a}{2}|k|^2} &= \left(\frac{1}{2\pi}\right)^{d/2} \int_{\mathbb{R}^d} dk \int_{\mathbb{R}^d} dy f\left(y\right) e^{-ik \cdot y} e^{ik \cdot x} e^{-\frac{a}{2}|k|^2} \\ &= \left(\frac{1}{2\pi}\right)^{d/2} \int_{\mathbb{R}^d} dy f\left(y\right) \int_{\mathbb{R}^d} dk e^{ik \cdot (x-y)} e^{-\frac{a}{2}|k|^2} \\ &= \left(\frac{1}{2\pi}\right)^{d/2} \left(\frac{2\pi}{a}\right)^{d/2} \int_{\mathbb{R}^d} e^{-\frac{1}{2a}|x-y|^2} f\left(y\right) dy. \end{split}$$

Thus we have shown,  $f_0(x) = \lim_{a \downarrow 0} (\delta_a * f)(x)$  where

$$\delta_a\left(x\right) = \left(\frac{2\pi}{a}\right)^{d/2} e^{-\frac{1}{2a}|x-y|^2} = \left(\frac{1}{\sqrt{a}}\right)^d \delta_1\left(\frac{x}{\sqrt{a}}\right).$$

By Theorem 4.6,  $\delta_a * f \to f$  in  $L^1(m)$  as  $a \downarrow 0$  and hence we conclude that  $f_0(x) = f(x)$  for a.e. x.Along the way we have shown  $\mathcal{F}^{-1}\mathcal{F}f = f_0 = f$  a.e.. A similar computation shows  $f_1 := \mathcal{F}\mathcal{F}^{-1}f = f$  a.e. and as both  $f_1$  and  $f_0$  are continuous it follows that  $f_1 = f_0$ .

For the last item we note,

$$\begin{split} \left\| \hat{f} \right\|_{2}^{2} &= \int_{\mathbb{R}^{d}} \hat{f}\left(\xi\right) \overline{\hat{f}\left(\xi\right)} \mathbf{d}\xi = \int_{\mathbb{R}^{d}} \mathbf{d}\xi \ \hat{f}\left(\xi\right) \int_{\mathbb{R}^{d}} \mathbf{d}x \overline{f\left(x\right)} e^{ix \cdot \xi} \\ &= \int_{\mathbb{R}^{d}} \mathbf{d}x \ \overline{f\left(x\right)} \int_{\mathbb{R}^{d}} \mathbf{d}\xi \ \hat{f}\left(\xi\right) e^{ix \cdot \xi} \text{ (by Fubini)} \\ &= \int_{\mathbb{R}^{d}} \mathbf{d}x \ \overline{f\left(x\right)} f\left(x\right) = \|f\|_{2}^{2} \end{split}$$

because

$$\int_{\mathbb{R}^d} \mathbf{d}\xi \hat{f}\left(\xi\right) e^{ix\cdot\xi} = \mathcal{F}^{-1}\hat{f}\left(x\right) = f\left(x\right) \text{ for a.e. } x.$$

The next theorem summarizes some more basic properties of the Fourier transform.

**Theorem 5.9.** Suppose that  $f, g \in L^1$ . Then

1. 
$$\hat{f} \in C_0(\mathbb{R}^d)$$
 and  $\left\|\hat{f}\right\|_{\infty} \leq \|f\|_{L^1(\lambda)}$  or equivalently,  
 $\left\|\hat{f}\right\|_{\infty} \leq c_d \|f\|_{L^1(m)}$ .

2. For 
$$y \in \mathbb{R}^d$$
,  $(\tau_y f) \hat{}(\xi) = e^{-iy \cdot \xi} \hat{f}(\xi)$  where, as usual,  $\tau_y f(x) := f(x - y)$ .

#### 30 5 Fourier Transform

3. The Fourier transform takes convolution to products, i.e.  $(f \bigstar g)^{\hat{}} = \hat{f}\hat{g}$ , i.e.

$$\left(\frac{1}{\sqrt{2\pi}}\right)^{d}\left(f*g\right)^{\hat{}} = \hat{f}\left(\xi\right)\hat{g}\left(\xi\right)$$

4. The operations of "<sup>^</sup>" and "<sup>∨</sup>" are interchanged under complex conjugation. For example, (f) = (f<sup>∨</sup>).
5. For f, g ∈ L<sup>1</sup>,

$$\langle \hat{f},g 
angle = \langle f,\hat{g} 
angle \, \, and \, \left\langle \hat{f}|g 
ight
angle = \langle f|g^{ee} 
angle \, .$$

6. If  $T : \mathbb{R}^d \to \mathbb{R}^d$  is an invertible linear transformation, then

$$(f \circ T)^{\wedge}(\xi) = |\det T|^{-1} \hat{f}((T^{-1})^* \xi)$$
 and  
 $(f \circ T)^{\vee}(\xi) = |\det T|^{-1} f^{\vee}((T^{-1})^* \xi)$ 

7. If  $(1+|x|)^k f(x) \in L^1$ , then  $\hat{f} \in C^k$  and  $\partial^{\alpha} \hat{f} \in C_0$  for all  $|\alpha| \leq k$ . Moreover,

$$\partial_{\xi}^{\alpha} \hat{f}(\xi) = \mathcal{F}\left[\left(-ix\right)^{\alpha} f(x)\right](\xi)$$
(5.8)

for all  $|\alpha| \leq k$ .

8. If  $f \in C^k$  and  $\partial^{\alpha} f \in L^1$  for all  $|\alpha| \leq k$ , then  $(1 + |\xi|)^k \hat{f}(\xi) \in C_0$  and

$$\left(\partial^{\alpha}f\right)^{\hat{}}(\xi) = (i\xi)^{\alpha}\hat{f}(\xi) \tag{5.9}$$

for all  $|\alpha| \leq k$ .

9. Suppose  $g \in L^1(\mathbb{R}^k)$  and  $h \in L^1(\mathbb{R}^{d-k})$  and  $f = g \otimes h$ , i.e.

$$f(x) = g(x_1, \ldots, x_k)h(x_{k+1}, \ldots, x_d)$$

then  $\hat{f} = \hat{g} \otimes \hat{h}$ .

**Proof.** Item 1. is the Riemann Lebesgue Lemma 4.20. Items 2. - 6. are proved by the following straight forward computations:

$$\begin{aligned} (\tau_y f) \,\widehat{}\,(\xi) &= \int_{\mathbb{R}^d} e^{-ix \cdot \xi} f(x-y) \mathrm{d}x = \int_{\mathbb{R}^d} e^{-i(x+y) \cdot \xi} f(x) \,\mathrm{d}x = e^{-iy \cdot \xi} \widehat{f}(\xi) \,, \\ (\widehat{f}) \,\widehat{}\,(\xi) &= \int_{\mathbb{R}^d} e^{-ix \cdot \xi} \overline{f(x)} \mathrm{d}x = \overline{\int_{\mathbb{R}^d} e^{ix \cdot \xi} f(x) \,\mathrm{d}x} = \overline{(f^{\vee}(\xi))}, \\ \langle \widehat{f}, g \rangle &= \int_{\mathbb{R}^d} \widehat{f}(\xi) \, g(\xi) \,\mathrm{d}\xi = \int_{\mathbb{R}^d} \mathrm{d}\xi g(\xi) \int_{\mathbb{R}^d} \mathrm{d}x e^{-ix \cdot \xi} f(x) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathrm{d}x \mathrm{d}\xi e^{-ix \cdot \xi} g(\xi) \, f(x) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathrm{d}x \widehat{g}(x) \, f(x) = \langle f, \widehat{g} \rangle, \\ \langle \widehat{f} | g \rangle &= \langle \widehat{f}, \overline{g} \rangle = \langle f, (\overline{g}) \,\widehat{} \rangle = \langle f, \overline{g^{\vee}} \rangle = \langle f | g \rangle \,, \\ (f \bigstar g) \,\widehat{}\,(\xi) &= \int_{\mathbb{R}^d} e^{-ix \cdot \xi} f \bigstar g(x) \,\mathrm{d}x = \int_{\mathbb{R}^d} e^{-ix \cdot \xi} \left( \int_{\mathbb{R}^d} f(x-y) g(y) \,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_{\mathbb{R}^d} \mathrm{d}y \int_{\mathbb{R}^d} \mathrm{d}x e^{-ix \cdot \xi} f(x-y) g(y) \\ &= \int_{\mathbb{R}^d} \mathrm{d}y \int_{\mathbb{R}^d} \mathrm{d}x e^{-i(x+y) \cdot \xi} f(x) \, g(y) \\ &= \int_{\mathbb{R}^d} \mathrm{d}y e^{-iy \cdot \xi} g(y) \int_{\mathbb{R}^d} \mathrm{d}x e^{-ix \cdot \xi} f(x) = \widehat{f}(\xi) \, \widehat{g}(\xi) \end{aligned}$$

and letting y = Tx so that  $\mathbf{d}x = |\det T|^{-1} \mathbf{d}y$ 

$$(f \circ T)^{\hat{}}(\xi) = \int_{\mathbb{R}^d} e^{-ix \cdot \xi} f(Tx) dx = \int_{\mathbb{R}^d} e^{-iT^{-1}y \cdot \xi} f(y) |\det T|^{-1} dy$$
$$= |\det T|^{-1} \hat{f}((T^{-1})^* \xi).$$

Item 7. is simply a matter of differentiating under the integral sign which is easily justified because  $(1 + |x|)^k f(x) \in L^1$ . Item 8. follows by using Lemma 4.19 repeatedly (i.e. integration by parts) to find

$$\begin{split} (\partial^{\alpha} f)^{\hat{}}(\xi) &= \int_{\mathbb{R}^d} \partial^{\alpha}_x f(x) \, e^{-ix \cdot \xi} \mathbf{d}x = (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(x) \, \partial^{\alpha}_x e^{-ix \cdot \xi} \mathbf{d}x \\ &= (-1)^{|\alpha|} \int_{\mathbb{R}^d} f(x) \, (-i\xi)^{\alpha} e^{-ix \cdot \xi} \mathbf{d}x = (i\xi)^{\alpha} \hat{f}(\xi) \,. \end{split}$$

Since  $\partial^{\alpha} f \in L^1$  for all  $|\alpha| \leq k$ , it follows that  $(i\xi)^{\alpha} \hat{f}(\xi) = (\partial^{\alpha} f)^{\hat{}}(\xi) \in C_0$  for all  $|\alpha| \leq k$ . Since

$$(1+|\xi|)^k \le \left(1+\sum_{i=1}^d |\xi_i|\right)^k = \sum_{|\alpha|\le k} c_{\alpha} |\xi^{\alpha}|$$

where  $0 < c_{\alpha} < \infty$ ,

$$\left|\left(1+\left|\xi\right|\right)^k\hat{f}\left(\xi\right)\right|\leq\sum_{\left|\alpha\right|\leq k}c_\alpha\left|\xi^\alpha\hat{f}\left(\xi\right)\right|\rightarrow 0 \text{ as } \xi\rightarrow\infty.$$

Item 9. is a simple application of the Tonelli/Fubini theorems. **Note:** Let  $c := (2\pi)^{-d/2}$ , then

$$\begin{split} (f*g)^{\widehat{}}(\xi) &= c \int_{\mathbb{R}^d} e^{-ix \cdot \xi} f*g\left(x\right) dx = c \int_{\mathbb{R}^d} e^{-ix \cdot \xi} \left( \int_{\mathbb{R}^d} f(x-y)g\left(y\right) dy \right) dx \\ &= c \int_{\mathbb{R}^d} dy \int_{\mathbb{R}^d} dx e^{-ix \cdot \xi} f(x-y)g\left(y\right) \\ &= c \int_{\mathbb{R}^d} dy \int_{\mathbb{R}^d} dx e^{-i(x+y) \cdot \xi} f\left(x\right) g\left(y\right) \\ &= c^{-1} \cdot c \int_{\mathbb{R}^d} dy e^{-iy \cdot \xi} g\left(y\right) c \int_{\mathbb{R}^d} dx e^{-ix \cdot \xi} f\left(x\right) = c^{-1} \hat{f}\left(\xi\right) \hat{g}\left(\xi\right) \end{split}$$

*Remark 5.10.* The key point of items 7. and 8. of Theorem 5.9 above is that the Fourier transform interchanges multiplication with differentiation. The fundamental results are;

1. If  $f \in L^1(\mathbb{R}^d)$  is such that  $x \to x_j f(x)$  is also integrable, then (by Corollary ??)

$$(x \to x_j f(x))^{\hat{}}(\xi) = i \frac{\partial}{\partial \xi_j} \hat{f}(\xi)$$

2. If  $f \in L^1(\mathbb{R}^d)$  is such that  $(\partial_j f)(x)$  exists, is continuous in x (this may be weakened),  $\partial_j f \in L^1(\mathbb{R}^d)$ , then (by Exercise ??)

$$\widehat{\partial_{j}f}\left(\xi\right)=i\xi_{j}\widehat{f}\left(\xi\right).$$

## 5.3 Schwartz Test Functions

**Definition 5.11.** A function  $f \in C(\mathbb{R}^d, \mathbb{C})$  is said to have rapid decay or rapid decrease if

$$\sup_{x \in \mathbb{R}^d} (1 + |x|)^N |f(x)| < \infty \text{ for } N = 1, 2, \dots$$

Equivalently, for each  $N \in \mathbb{N}$  there exists constants  $C_N < \infty$  such that  $|f(x)| \leq C_N(1+|x|)^{-N}$  for all  $x \in \mathbb{R}^d$ . A function  $f \in C(\mathbb{R}^d, \mathbb{C})$  is said to have (at most) **polynomial growth** if there exists  $N < \infty$  such

$$\sup (1 + |x|)^{-N} |f(x)| < \infty,$$

*i.e.* there exists  $N \in \mathbb{N}$  and  $C < \infty$  such that  $|f(x)| \leq C(1+|x|)^N$  for all  $x \in \mathbb{R}^d$ .

**Definition 5.12 (Schwartz Test Functions).** Let S denote the space of functions  $f \in C^{\infty}(\mathbb{R}^d)$  such that f and all of its partial derivatives have rapid decay and let

$$\|f\|_{N,\alpha} = \sup_{x \in \mathbb{R}^d} \left| (1+|x|)^N \partial^{\alpha} f(x) \right|$$

 $so\ that$ 

$$\mathcal{S} = \left\{ f \in C^{\infty}\left(\mathbb{R}^{d}\right) : \left\|f\right\|_{N,\alpha} < \infty \text{ for all } N \text{ and } \alpha \right\}.$$

Also let  $\mathcal{P}$  denote those functions  $g \in C^{\infty}(\mathbb{R}^d)$  such that g and all of its derivatives have at most polynomial growth, i.e.  $g \in C^{\infty}(\mathbb{R}^d)$  is in  $\mathcal{P}$  iff for all multiindices  $\alpha$ , there exists  $N_{\alpha} < \infty$  such

$$\sup (1+|x|)^{-N_{\alpha}} |\partial^{\alpha} g(x)| < \infty.$$

(Notice that any polynomial function on  $\mathbb{R}^d$  is in  $\mathcal{P}$ .)

*Remark 5.13.* Since  $C_c^{\infty}(\mathbb{R}^d) \subset \mathcal{S} \subset L^2(\mathbb{R}^d)$ , it follows that  $\mathcal{S}$  is dense in  $L^2(\mathbb{R}^d)$ .

Exercise 5.2. Let

$$L = \sum_{|\alpha| \le k} a_{\alpha}(x) \,\partial^{\alpha} \tag{5.10}$$

with  $a_{\alpha} \in \mathcal{P}$ . Show  $L(\mathcal{S}) \subset \mathcal{S}$  and in particular  $\partial^{\alpha} f$  and  $x^{\alpha} f$  are back in  $\mathcal{S}$  for all multi-indices  $\alpha$ .

**Notation 5.14** Suppose that  $p(x,\xi) = \Sigma_{|\alpha| \leq N} a_{\alpha}(x) \xi^{\alpha}$  where each function  $a_{\alpha}(x)$  is a smooth function. We then set

$$p(x, D_x) := \Sigma_{|\alpha| \le N} a_\alpha(x) D_x^\alpha$$

and if each  $a_{\alpha}(x)$  is also a polynomial in x we will let

$$p(-D_{\xi},\xi) := \Sigma_{|\alpha| \le N} a_{\alpha}(-D_{\xi}) M_{\xi^{\alpha}}$$

where  $M_{\xi^{\alpha}}$  is the operation of multiplication by  $\xi^{\alpha}$ .

**Proposition 5.15.** Let  $p(x,\xi)$  be as above and assume each  $a_{\alpha}(x)$  is a polynomial in x. Then for  $f \in S$ ,

$$(p(x, D_x)f)^{\wedge}(\xi) = p(-D_{\xi}, \xi)\hat{f}(\xi)$$
(5.11)

and

$$p(\xi, D_{\xi})\hat{f}(\xi) = [p(D_x, -x)f(x)]^{\wedge}(\xi).$$
(5.12)

Page: 31

#### 32 5 Fourier Transform

**Proof.** The identities  $(-D_{\xi})^{\alpha} e^{-ix\cdot\xi} = x^{\alpha}e^{-ix\cdot\xi}$  and  $D_x^{\alpha}e^{ix\cdot\xi} = \xi^{\alpha}e^{ix\cdot\xi}$  imply, for any polynomial function q on  $\mathbb{R}^d$ ,

$$q(-D_{\xi})e^{-ix\cdot\xi} = q(x)e^{-ix\cdot\xi} \text{ and } q(D_x)e^{ix\cdot\xi} = q(\xi)e^{ix\cdot\xi}.$$
(5.13)

Therefore using Eq. (5.13) repeatedly,

$$(p(x, D_x)f)^{\wedge}(\xi) = \int_{\mathbb{R}^d} \sum_{|\alpha| \le N} a_{\alpha}(x) D_x^{\alpha} f(x) \cdot e^{-ix \cdot \xi} \mathbf{d}\xi$$
$$= \int_{\mathbb{R}^d} \sum_{|\alpha| \le N} D_x^{\alpha} f(x) \cdot a_{\alpha}(-D_{\xi}) e^{-ix \cdot \xi} \mathbf{d}\xi$$
$$= \int_{\mathbb{R}^d} f(x) \sum_{|\alpha| \le N} (-D_x)^{\alpha} \left[ a_{\alpha}(-D_{\xi}) e^{-ix \cdot \xi} \right] \mathbf{d}\xi$$
$$= \int_{\mathbb{R}^d} f(x) \sum_{|\alpha| \le N} a_{\alpha}(-D_{\xi}) \left[ \xi^{\alpha} e^{-ix \cdot \xi} \right] \mathbf{d}\xi = p(-D_{\xi}, \xi) \hat{f}(\xi)$$

wherein the third inequality we have used Lemma 4.19 to do repeated integration by parts, the fact that mixed partial derivatives commute in the fourth, and in the last we have repeatedly used Corollary ?? to differentiate under the integral. The proof of Eq. (5.12) is similar:

$$\begin{split} p(\xi, D_{\xi})\hat{f}(\xi) &= p(\xi, D_{\xi}) \int_{\mathbb{R}^d} f(x) e^{-ix \cdot \xi} \mathbf{d}x = \int_{\mathbb{R}^d} f(x) p(\xi, -x) e^{-ix \cdot \xi} \mathbf{d}x \\ &= \sum_{|\alpha| \le N} \int_{\mathbb{R}^d} f(x) (-x)^{\alpha} a_{\alpha}(\xi) e^{-ix \cdot \xi} \mathbf{d}x \\ &= \sum_{|\alpha| \le N} \int_{\mathbb{R}^d} f(x) (-x)^{\alpha} a_{\alpha}(-D_x) e^{-ix \cdot \xi} \mathbf{d}x \\ &= \sum_{|\alpha| \le N} \int_{\mathbb{R}^d} e^{-ix \cdot \xi} a_{\alpha}(D_x) \left[ (-x)^{\alpha} f(x) \right] \mathbf{d}x \\ &= \left[ p(D_x, -x) f(x) \right]^{\wedge} (\xi) \,. \end{split}$$

**Corollary 5.16.** The Fourier transform preserves the space S, *i.e.*  $\mathcal{F}(S) \subset S$ .

**Proof.** Let  $p(x,\xi) = \Sigma_{|\alpha| \le N} a_{\alpha}(x) \xi^{\alpha}$  with each  $a_{\alpha}(x)$  being a polynomial function in x. If  $f \in S$  then  $p(D_x, -x)f \in S \subset L^1$  and so by Eq. (5.12),  $p(\xi, D_{\xi})\hat{f}(\xi)$  is bounded in  $\xi$ , i.e.

$$\sup_{\xi \in \mathbb{R}^d} |p(\xi, D_{\xi})\hat{f}(\xi)| \le C(p, f) < \infty.$$

Taking  $p(x,\xi) = (1+|x|^2)^N \xi^{\alpha}$  with  $N \in \mathbb{Z}_+$  in this estimate shows  $\hat{f}(\xi)$  and all of its derivatives have rapid decay, i.e.  $\hat{f}$  is in S.

In the next few exercises you are asked to compute the Fourier transform of a number of functions.

**Exercise 5.3.** In this problem let d = 1 so that  $x, \xi \in \mathbb{R} = \mathbb{R}^1$ . For any m > 0, show

$$\mathcal{F}\left[e^{-m|x|}\right](\xi) = \frac{2m}{\sqrt{2\pi}} \frac{1}{m^2 + \xi^2}$$

and

$$\mathcal{F}\left(\frac{1}{m^2+\xi^2}\right)(x) = \frac{\sqrt{2\pi}}{2m}e^{-m|x|}$$

More precisely these equations mean;

$$\mathcal{F}\left[x \to e^{-m|x|}\right](\xi) = \frac{2m}{\sqrt{2\pi}} \frac{1}{m^2 + \xi^2} \text{ and}$$
$$\mathcal{F}\left(\xi \to \frac{1}{m^2 + \xi^2}\right)(x) = \frac{\sqrt{2\pi}}{2m} e^{-m|x|}$$

or equivalently,

$$\mathcal{F}\left[e^{-m|\cdot|}\right] = \frac{2m}{\sqrt{2\pi}} \frac{1}{m^2 + (\cdot)^2} \text{ and } \mathcal{F}\left(\frac{1}{m^2 + (\cdot)^2}\right) = \frac{\sqrt{2\pi}}{2m} e^{-m|\cdot|}.$$

Exercise 5.4. Using the identity

$$\frac{1}{\xi^2 + 1} = \int_0^\infty e^{-s(\xi^2 + 1)} ds$$

along with Exercise 5.3 and the known Fourier transform of Gaussians to show

$$e^{-|x|} = \int_0^\infty ds \frac{1}{\sqrt{\pi s}} e^{-s} e^{-\frac{x^2}{4s}}$$
 for all  $x \in \mathbb{R}$ . (5.14)

Thus we have written  $e^{-|x|}$  as an average of Gaussians.

**Exercise 5.5.** Now let  $x \in \mathbb{R}^d$  and  $|x|^2 := \sum_{i=1}^d x_i^2$  be the standard Euclidean norm. Show for all m > 0 that

$$\mathcal{F}\left[e^{-m|x|}\right](\xi) = \frac{2^{d/2}}{\sqrt{\pi}} \Gamma\left(\frac{d+1}{2}\right) \frac{m}{\left(m^2 + |\xi|^2\right)^{\frac{d+1}{2}}},$$
(5.15)

where  $\Gamma(x)$  in the gamma function defined as
$$\Gamma\left(x\right) := \int_{0}^{\infty} t^{x} e^{-t} \frac{dt}{t}.$$

**Hint:** By Exercise 5.4 with x replaced by m|x| we know that

$$e^{-m|x|} = \int_0^\infty ds \frac{1}{\sqrt{\pi s}} e^{-s} e^{-\frac{m^2}{4s}|x|^2}$$
 for all  $x \in \mathbb{R}^d$ .

Remark 5.17. This result can be used to show,

$$e^{-m\sqrt{-\Delta}}f(x) = \int_{\mathbb{R}^d} Q_m(x-y)f(y) \, dy$$

where

$$Q_m(x) = 2^{d/2} \frac{\Gamma((d+1)/2)}{(2\pi)^{d/2} \sqrt{\pi}} \frac{m}{(m^2 + |x|^2)^{(d+1)/2}} = \frac{\Gamma((d+1)/2)}{\pi^{d/2} \sqrt{\pi}} \frac{m}{(m^2 + |x|^2)^{(d+1)/2}}$$
$$= \frac{\Gamma((d+1)/2)}{\pi^{(d+1)/2}} \frac{m}{(m^2 + |x|^2)^{(d+1)/2}}.$$

The extra factors of  $\sqrt{2\pi}$  come from the normalized convolution.

**Corollary 5.18 (Fourier Transform on**  $L^2$ ). By the B.L.T. Theorem 8.1, the maps  $\mathcal{F}|_{\mathcal{S}}$  and  $\mathcal{F}^{-1}|_{\mathcal{S}}$  extend to bounded linear maps  $\overline{\mathcal{F}}$  and  $\overline{\mathcal{F}}^{-1}$  from  $L^2 \to L^2$ . These maps satisfy the following properties:

- 1.  $\overline{\mathcal{F}}$  and  $\overline{\mathcal{F}}^{-1}$  are unitary and are inverses to one another as the notation suggests.
- 2. If  $f \in L^2$ , then  $\overline{\mathcal{F}}f$  is uniquely characterized as the function,  $G \in L^2$  such that

$$\langle G, \psi \rangle = \langle f, \hat{\psi} \rangle$$
 for all  $\psi \in C_c^{\infty}(\mathbb{R}^d)$ .

3. If  $f \in L^1 \cap L^2$ , then  $\overline{\mathcal{F}}f = \hat{f}$  a.e. 4. For  $f \in L^2$  we may compute  $\overline{\mathcal{F}}$  and  $\overline{\mathcal{F}}^{-1}$  by

$$\bar{\mathcal{F}}f\left(\xi\right) = L^{2} - \lim_{R \to \infty} \int_{|x| \le R} f\left(x\right) e^{-ix \cdot \xi} \mathbf{d}x \text{ and}$$
(5.16)

$$\bar{\mathcal{F}}^{-1}f\left(\xi\right) = L^2 - \lim_{R \to \infty} \int_{|x| \le R} f\left(x\right) e^{ix \cdot \xi} \mathbf{d}x.$$
(5.17)

5. We may further extend  $\overline{\mathcal{F}}$  to a map from  $L^1 + L^2 \to C_0 + L^2$  (still denote by  $\overline{\mathcal{F}}$ ) defined by  $\overline{\mathcal{F}}f = \hat{h} + \overline{\mathcal{F}}g$  where  $f = h + g \in L^1 + L^2$ . For  $f \in L^1 + L^2$ ,  $\overline{\mathcal{F}}f$  may be characterized as the unique function  $F \in L^1_{loc}(\mathbb{R}^d)$  such that

$$\langle F, \varphi \rangle = \langle f, \hat{\varphi} \rangle \text{ for all } \varphi \in C_c^{\infty}(\mathbb{R}^d).$$
 (5.18)

Moreover if Eq. (5.18) holds then  $F \in C_0 + L^2 \subset L^1_{loc}(\mathbb{R}^d)$  and Eq.(5.18) is valid for all  $\varphi \in S$ .

**Proof. 1. and 2.** If  $f \in L^2$  and  $\varphi_n \in S$  such that  $\varphi_n \to f$  in  $L^2$  (see Exercise 4.3), then  $\overline{\mathcal{F}}f := \lim_{n\to\infty} \hat{\varphi}_n$ . Since  $\hat{\varphi}_n \in S \subset L^1$ , we may concluded that  $\|\hat{\varphi}_n\|_2 = \|\varphi_n\|_2$  for all n. Thus

$$\left\|\bar{\mathcal{F}}f\right\|_{2} = \lim_{n \to \infty} \|\hat{\varphi}_{n}\|_{2} = \lim_{n \to \infty} \|\varphi_{n}\|_{2} = \|f\|_{2}$$

which shows that  $\bar{\mathcal{F}}$  is an isometry from  $L^2$  to  $L^2$  and similarly  $\bar{\mathcal{F}}^{-1}$  is an isometry. Since  $\bar{\mathcal{F}}^{-1}\bar{\mathcal{F}} = \mathcal{F}^{-1}\mathcal{F} = id$  on the dense set  $\mathcal{S}$ , it follows by continuity that  $\bar{\mathcal{F}}^{-1}\bar{\mathcal{F}} = id$  on all of  $L^2$ . Hence  $\bar{\mathcal{F}}\bar{\mathcal{F}}^{-1} = id$ , and thus  $\bar{\mathcal{F}}^{-1}$  is the inverse of  $\bar{\mathcal{F}}$ . This proves item 1. Moreover, if  $\psi \in C_c^{\infty}(\mathbb{R}^d)$ , then

$$\langle \bar{\mathcal{F}}f, \psi \rangle = \lim_{n \to \infty} \langle \hat{\varphi}_n, \psi \rangle = \lim_{n \to \infty} \langle \varphi_n, \hat{\psi} \rangle = \langle f, \psi \rangle$$
(5.19)

and this equation uniquely characterizes  $\overline{\mathcal{F}}f$  by Corollary 4.16. Notice that Eq. (5.19) also holds for all  $\psi \in \mathcal{S}$ .

**3.** If  $f \in L^1 \cap L^2$ , we have already seen that  $\hat{f} \in C_0(\mathbb{R}^d) \subset L^1_{loc}$  and that  $\langle \hat{f}, \psi \rangle = \langle f, \hat{\psi} \rangle$  for all  $\psi \in C_c^{\infty}(\mathbb{R}^d)$ . Combining this with item 2. shows  $\langle \hat{f} - \bar{\mathcal{F}}f, \psi \rangle = 0$  or all  $\psi \in C_c^{\infty}(\mathbb{R}^d)$  and so again by Corollary 4.16 we conclude that  $\hat{f} - \bar{\mathcal{F}}f = 0$  a.e.

**Alternatively** by Exercise 4.3, if  $f \in L^1 \cap L^2$ , there exists  $f_n \in C_c^{\infty}(\mathbb{R})$  such that  $\lim_{n\to\infty} \|f - f_n\|_p = 0$  for both p = 1 and p = 2. Therefore  $\hat{f} = \lim_{n\to\infty} \hat{f}_n$  in  $L^{\infty}$  while  $\lim_{n\to\infty} \hat{f}_n = \lim_{n\to\infty} \mathcal{F}f_n = \bar{\mathcal{F}}f$  in  $L^2$  which is enough to conclude that  $\hat{f} = \bar{\mathcal{F}}f$  a.e.

**4.** Let  $f \in L^2$  and  $R < \infty$  and set  $f_R(x) := f(x) \mathbf{1}_{|x| \leq R}$ . Then  $f_R \in L^1 \cap L^2$  and therefore  $\overline{\mathcal{F}} f_R = \widehat{f}_R$ . Since  $\overline{\mathcal{F}}$  is an isometry and (by the dominated convergence theorem)  $f_R \to f$  in  $L^2$ , it follows that

$$\bar{\mathcal{F}}f = L^2 - \lim_{R \to \infty} \bar{\mathcal{F}}f_R = L^2 - \lim_{R \to \infty} \hat{f}_R.$$

5. If  $f = h + g \in L^1 + L^2$  and  $\varphi \in S$ , then by Eq. (5.19) and item 4. of Theorem 5.9,

$$\langle \hat{h} + \bar{\mathcal{F}}g, \varphi \rangle = \langle h, \hat{\varphi} \rangle + \langle g, \hat{\varphi} \rangle = \langle h + g, \hat{\varphi} \rangle.$$
 (5.20)

In particular if h + g = 0 a.e., then  $\langle \hat{h} + \bar{\mathcal{F}}g, \varphi \rangle = 0$  for all  $\varphi \in \mathcal{S}$  and since  $\hat{h} + \bar{\mathcal{F}}g \in L^1_{loc}$  it follows from Corollary 4.16 that  $\hat{h} + \bar{\mathcal{F}}g = 0$  a.e. This shows that  $\bar{\mathcal{F}}f$  is well defined independent of how  $f \in L^1 + L^2$  is decomposed into the sum of an  $L^1$  and an  $L^2$  function. Moreover Eq. (5.20) shows Eq. (5.18) holds with  $F = \hat{h} + \bar{\mathcal{F}}g \in C_0 + L^2$  and  $\varphi \in \mathcal{S}$ . Now suppose  $G \in L^1_{loc}$  and  $\langle G, \varphi \rangle = \langle f, \hat{\varphi} \rangle$  for all  $\varphi \in C_c^{\infty} (\mathbb{R}^d)$ . Then by what we just proved,  $\langle G, \varphi \rangle = \langle F, \varphi \rangle$  for all  $\varphi \in C_c^{\infty} (\mathbb{R}^d)$  and so another application of Corollary 4.16 shows  $G = F \in C_0 + L^2$ .

#### 34 5 Fourier Transform

**Notation 5.19** Given the results of Corollary 5.18, there is little danger in writing  $\hat{f}$  or  $\mathcal{F}f$  for  $\bar{\mathcal{F}}f$  when  $f \in L^1 + L^2$ .

**Corollary 5.20.** If f and g are  $L^1$  functions such that  $\hat{f}, \hat{g} \in L^1$ , then

$$\mathcal{F}(fg) = \hat{f} \bigstar \hat{g} \text{ and } \mathcal{F}^{-1}(fg) = f^{\vee} \bigstar g^{\vee}.$$

Since S is closed under pointwise products and  $\mathcal{F} : S \to S$  is an isomorphism it follows that S is closed under convolution as well.

**Proof.** By Theorem 5.8,  $f, g, \hat{f}, \hat{g} \in L^1 \cap L^\infty$  and hence  $f \cdot g \in L^1 \cap L^\infty$  and  $\hat{f} \bigstar \hat{g} \in L^1 \cap L^\infty$ . Since

$$\mathcal{F}^{-1}\left(\hat{f}\bigstar\hat{g}\right) = \mathcal{F}^{-1}\left(\hat{f}\right) \cdot \mathcal{F}^{-1}\left(\hat{g}\right) = f \cdot g \in L^{1}$$

we may conclude from Theorem 5.8 that

$$\hat{f} \bigstar \hat{g} = \mathcal{F}\mathcal{F}^{-1}\left(\hat{f}\bigstar \hat{g}\right) = \mathcal{F}(f \cdot g).$$

Similarly one shows  $\mathcal{F}^{-1}(fg) = f^{\vee} \bigstar g^{\vee}$ .

**Corollary 5.21.** Let  $p(x,\xi)$  and  $p(x,D_x)$  be as in Notation 5.14 with each function  $a_{\alpha}(x)$  being a smooth function of  $x \in \mathbb{R}^d$ . Then for  $f \in S$ ,

$$p(x, D_x)f(x) = \int_{\mathbb{R}^d} p(x, \xi)\hat{f}(\xi) e^{ix\cdot\xi} \mathbf{d}\xi.$$
 (5.21)

**Proof.** For  $f \in \mathcal{S}$ , we have

$$p(x, D_x)f(x) = p(x, D_x) \left(\mathcal{F}^{-1}\hat{f}\right)(x) = p(x, D_x) \int_{\mathbb{R}^d} \hat{f}(\xi) e^{ix\cdot\xi} d\xi$$
$$= \int_{\mathbb{R}^d} \hat{f}(\xi) p(x, D_x) e^{ix\cdot\xi} d\xi = \int_{\mathbb{R}^d} \hat{f}(\xi) p(x, \xi) e^{ix\cdot\xi} d\xi.$$

Lemma 5.22 (Petree's inequalities). If  $x, y \in \mathbb{R}^n$ , then

$$(1+|x-y|)(1+|x|)^{-1} \le 1+|y| \tag{5.22}$$

and

$$(1+|x-y|)^{-1} \le (1+|x|)^{-1} (1+|y|).$$
(5.23)

**Proof.** For  $x, y \in \mathbb{R}^n$  we have the following simple estimate,

$$1 + |x - y| \le 1 + |x| + |y| \le (1 + |x|)(1 + |y|)$$

which is equivalent to Eq. (5.22). Moreover, Eq. (5.22) is equivalent to

$$(1+|x|)^{-1} \le (1+|x-y|)^{-1} (1+|y|).$$

Replacing  $x \to x + y$  and then  $y \to -y$  in this last inequality gives Eq. (5.23).

**Lemma 5.23.** If  $f, g \in S := S(\mathbb{R}^n)$ , then  $f * g \in S$  where

$$f * g(x) := \int_{\mathbb{R}^n} f(x - y) g(y) \, dy.$$

[The proof shows  $f * g \in S$  if  $f \in S$  and g is a measurable function such that  $\int_{\mathbb{R}^n} (1+|y|)^n |g(y)| dy < \infty$  for all  $n \in \mathbb{N}$ .]

**Proof.** For any  $k \in \mathbb{N}_0^n$  and  $m \in \mathbb{N}$ , we have  $|f^{(k)}(x)| \leq C_{m,k} (1+|x|)^{-m}$  for some  $C_{m,k} < \infty$  where

$$f^{(k)} := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j}\right)^{k_j} f.$$

Using

$$\left(f\ast g\right)^{\left(k\right)}\left(x\right):=\int_{\mathbb{R}^{n}}f^{\left(k\right)}\left(x-y\right)g\left(y\right)dy$$

and Petree's inequality in Eq. (5.23), we conclude,

$$(f * g)^{(k)} (x) \Big| \leq \int_{\mathbb{R}^n} \Big| f^{(k)} (x - y) \Big| |g(y)| dy \leq C_{m,k} \int_{\mathbb{R}^n} (1 + |x - y|)^{-m} |g(y)| dy \leq C_{m,k} \int_{\mathbb{R}^n} (1 + |x|)^{-m} (1 + |y|)^m |g(y)| dy \leq \tilde{C}_{m,k} (1 + |x|)^{-m}$$

where

$$\tilde{C}_{m,k} := C_{m,k} \cdot \int_{\mathbb{R}^n} \left(1 + |y|\right)^m |g\left(y\right)| \, dy < \infty.$$

Lemma 5.24 (Convolution and products in S). If  $f, g \in S$  then

$$\left(\frac{1}{2\pi}\right)^{n/2}\widehat{f*g} = \hat{f}\cdot\hat{g} \text{ and } \widehat{f\cdot g} = \left(\frac{1}{2\pi}\right)^{n/2}\hat{f}*\hat{g}.$$

**Proof.** The first equality is proved using Fubini's theorem and the translation invariance of Lebesgue measure;

Page: 34

$$\widehat{f * g}(k) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} (f * g)(x) e^{-ikx} dx$$
  
=  $(2\pi)^{-n/2} \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy f(x - y) g(y) e^{-ikx}$   
=  $(2\pi)^{-n/2} \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy f(x) g(y) e^{-ik(x+y)}$   
=  $(2\pi)^{-n/2} \int_{\mathbb{R}^n} dx \int_{\mathbb{R}^n} dy f(x) e^{-ikx} g(y) e^{-iky}$   
=  $(2\pi)^{-n/2} \cdot \widehat{f}(k) \widehat{g}(k)$ .

Similarly one shows,  $(f * g)^{\vee} = \sqrt{2\pi} f^{\vee} \cdot g^{\vee}$ . Replacing f by  $\hat{f}$  and g by  $\hat{g}$  in this equation then shows,

$$\left(\hat{f} * \hat{g}\right)^{\vee} = \left(2\pi\right)^{-n/2} f \cdot g$$

and then taking the Fourier transform of this result gives the second stated equation.  $\hfill\blacksquare$ 

If  $p(x,\xi)$  is a more general function of  $(x,\xi)$  then that given in Notation 5.14, the right member of Eq. (5.21) may still make sense, in which case we may use it as a definition of  $p(x, D_x)$ . A linear operator defined this way is called a **pseudo differential operator** and they turn out to be a useful class of operators to study when working with partial differential equations.

**Definition 5.25 (Weak Differentiability).** Let  $v \in \mathbb{R}^d$  and  $u \in L^p(\mathbb{R}^d)$ , then  $\partial_v u$  is said to **exist weakly** in  $L^p(\mathbb{R}^d)$  if there exists a function  $g \in L^p(\mathbb{R}^d)$  such that

$$\langle u, \partial_v \varphi \rangle = -\langle g, \varphi \rangle \text{ for all } \varphi \in C_c^\infty \left( \mathbb{R}^d \right)$$
 (5.24)

where

$$\langle u,v \rangle := \int_{\mathbb{R}^d} u\left(x\right) v\left(x\right) dx.$$

More generally if  $p(\xi) = \sum_{|\alpha| \le N} a_{\alpha} \xi^{\alpha}$  is a polynomial in  $\xi \in \mathbb{R}^d$  and  $p(\partial) := \sum_{|\alpha| \le N} a_{\alpha} \partial^{\alpha}$ , then we say  $p(\partial) u$  exists weakly in  $L^p(\mathbb{R}^d)$  if there exists a function  $g \in L^p(\mathbb{R}^d)$  such that

$$\langle u, p(-\partial) \varphi \rangle = \langle g, \varphi \rangle \text{ for all } \varphi \in C_c^{\infty}(\mathbb{R}^d).$$
 (5.25)

[This definition also makes sense if  $L^p(\mathbb{R}^d)$  is replaced by  $L^p_{loc}(\mathbb{R}^d)$  everywhere.]

**Proposition 5.26.** Suppose that  $f \in L^1_{loc}(\mathbb{R})$  such that  $\partial^{(w)}f = 0$  in  $L^1_{loc}(\mathbb{R})$ . Then there exists  $c \in \mathbb{C}$  such that f = c a.e. More generally, suppose  $F : C^{\infty}_c(\mathbb{R}) \to \mathbb{C}$  is a linear functional such that  $F(\varphi') = 0$  for all  $\varphi \in C^{\infty}_c(\mathbb{R})$ , where  $\varphi'(x) = \frac{d}{dx}\varphi(x)$ , then there exists  $c \in \mathbb{C}$  such that

$$F(\varphi) = \langle c, \varphi \rangle = \int_{\mathbb{R}} c\varphi(x) \, dx \text{ for all } \varphi \in C_c^{\infty}(\mathbb{R}) \,. \tag{5.26}$$

**Proof.** Before giving a proof of the second assertion, let us show it includes the first. Indeed, if  $F(\varphi) := \int_{\mathbb{R}} \varphi f dm$  and  $\partial^{(w)} f = 0$ , then  $F(\varphi') = 0$  for all  $\varphi \in C_c^{\infty}(\mathbb{R})$  and therefore there exists  $c \in \mathbb{C}$  such that

$$\int_{\mathbb{R}} \varphi f dm = F(\varphi) = c \langle \varphi, 1 \rangle = c \int_{\mathbb{R}} \varphi f dm.$$

But this implies f = c a.e. So it only remains to prove the second assertion.

Let  $\eta \in C_c^{\infty}(\mathbb{R})$  such that  $\int_{\mathbb{R}} \eta dm = 1$ . Given  $\varphi \in C_c^{\infty}(\mathbb{R}) \subset C_c^{\infty}(\mathbb{R})$ , let

$$\psi\left(x\right) = \int_{-\infty}^{x} \left(\varphi\left(y\right) - \eta\left(y\right)\left\langle\varphi, 1\right\rangle\right) dy$$

Then  $\psi'(x) = \varphi(x) - \eta(x) \langle \varphi, 1 \rangle$  and  $\psi \in C_c^{\infty}(\mathbb{R})$  as the reader should check. Therefore,

$$0 = F(\psi) = F(\varphi - \langle \varphi, \eta \rangle \eta) = F(\varphi) - \langle \varphi, 1 \rangle F(\eta)$$

which shows Eq. (5.26) holds with  $c = F(\eta)$ .

Alternative proof of first assertion. Suppose  $f \in L^1_{loc}(\mathbb{R})$  and  $\partial^{(w)} f = 0$ and  $f_m := f * \eta_m$  as is in the proof of Lemma ??. Then  $f'_m = \partial^{(w)} f * \eta_m = 0$ , so  $f_m = c_m$  for some constant  $c_m \in \mathbb{C}$ . By Theorem 4.6,  $f_m \to f$  in  $L^1_{loc}(\mathbb{R})$ and therefore if J = [a, b] is a compact subinterval of  $\mathbb{R}$ ,

$$|c_m - c_k| = \frac{1}{b-a} \int_J |f_m - f_k| \, dm \to 0 \text{ as } m, k \to \infty.$$

So  $\{c_m\}_{m=1}^{\infty}$  is a Cauchy sequence and therefore  $c := \lim_{m \to \infty} c_m$  exists and  $f = \lim_{m \to \infty} f_m = c$  a.e.

**Proposition 5.27.** Let  $f, g \in L^1_{loc}(\mathbb{R})$ , then f' = g weakly iff f has a continuous version  $\tilde{f}$  which is absolutely continuous on  $\mathbb{R}$  and satisfies  $\tilde{f}'(x) = g(x)$  for a.e. x.

**Proof.** If f is locally absolutely continuous and f' = g a.e., then by integration by parts for absolutely continuous functions,

$$\int_{\mathbb{R}} g\varphi dm = \int_{\mathbb{R}} f'\varphi dm = -\int_{\mathbb{R}} f\varphi' dm$$

#### 36 5 Fourier Transform

which shows that f' = g weakly. Conversely if f' = g weakly, let

$$F\left(x\right) := \int_{0}^{x} g\left(y\right) dy$$

which is absolutely continuous and satisfied F'(x) = g(x) for a.e. x. From what we just proved this implies F' = g weakly and therefore (F - f)' = 0 weakly and hence by Proposition 5.26,  $f = F + c =: \tilde{f}$  a.e. for some constant c.

**Exercise 5.6.** Suppose  $p(\xi) = \sum_{|\alpha| \le N} a_{\alpha} \xi^{\alpha}$  is a polynomial in  $\xi \in \mathbb{R}^d$ , and  $u \in L^2$  such that  $p(\partial) u = g \in L^2$  in the weak sense, i.e.

$$\langle u, p(-\partial) \varphi \rangle = \langle g, \varphi \rangle \text{ for all } \varphi \in C_c^{\infty}(\mathbb{R}^d).$$
 (5.27)

Show that Eq. (5.27) also holds for all  $\varphi \in \mathcal{S}(\mathbb{R}^d)$ . **Hints:** Let  $\psi \in C_c^{\infty}(\mathbb{R}^d, [0, 1])$  be chosen so that  $\psi(x) = 1$  for  $|x| \leq 1$  and for  $n \in \mathbb{N}$ , let  $\psi_n(x) := \psi(x/n)$ . Then for  $\varphi \in \mathcal{S}$ , consider  $\psi_n \cdot \varphi$ .

**Exercise 5.7 (F.T. and Weak derivatives).** Suppose  $p(\xi) = \sum_{|\alpha| \le N} a_{\alpha} \xi^{\alpha}$  is a polynomial in  $\xi \in \mathbb{R}^d$  and  $f, g \in L^2(m)$ . Show  $p(\partial) f = g$  weakly iff  $p(ik) \hat{f}(k) = \hat{g}(k)$  for a.e. k.

**Exercise 5.8.** Show for  $f \in \mathcal{S}(\mathbb{R})$  that;

1. For all  $x \in \mathbb{R}$ ,

$$|f(x)| \le \frac{1}{\sqrt{2\pi}} \left\| \hat{f} \right\|_1 = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left| \hat{f}(k) \right| dk$$

and

$$|f(x)| \le \frac{1}{\sqrt{2}} \left[ \int_{\mathbb{R}} \left| \hat{f}(k) \right|^2 (1+k^2) dk \right]^{1/2}.$$

2. Use the last displayed inequality and the basic properties of the Fourier transform to prove the "Sobolev inequality,"

$$|f(x)|^{2} \leq \frac{1}{2} \left[ ||f||_{2}^{2} + ||f'||_{2}^{2} \right] \text{ for all } x \in \mathbb{R},$$

where

$$||f||_{2}^{2} := \int_{\mathbb{R}} |f(x)|^{2} dx.$$

# 5.4 Summary of Basic Properties of $\mathcal{F}$ and $\mathcal{F}^{-1}$

The following table summarizes some of the basic properties of the Fourier transform and its inverse.

$$\begin{array}{rcl} f & \longleftrightarrow & \hat{f} \text{ or } f^{\vee} \\ \text{Smoothness} & \longleftrightarrow & \text{Decay at infinity} \\ \partial^{\alpha} & \longleftrightarrow & \text{Multiplication by } (\pm i\xi)^{\alpha} \\ \mathcal{S} & \longleftrightarrow & \mathcal{S} \\ L^2 \left( \mathbb{R}^d \right) & \longleftrightarrow & L^2 \left( \mathbb{R}^d \right) \\ \text{Convolution} & \longleftrightarrow & \text{Products.} \end{array}$$

# Constant Coefficient partial differential equations

Suppose that  $p(\xi) = \sum_{|\alpha| \le k} a_{\alpha} \xi^{\alpha}$  with  $a_{\alpha} \in \mathbb{C}$  and

$$L = p(D_x) := \Sigma_{|\alpha| \le N} a_{\alpha} D_x^{\alpha} = \Sigma_{|\alpha| \le N} a_{\alpha} \left(\frac{1}{i} \partial_x\right)^{\alpha}.$$
 (6.1)

Then for  $f \in \mathcal{S}$ 

$$\widehat{Lf}\left(\xi\right) = p\left(\xi\right)\widehat{f}\left(\xi\right)$$

that is to say the Fourier transform takes a constant coefficient partial differential operator to multiplication by a polynomial. This fact can often be used to solve constant coefficient partial differential equation. For example suppose  $g: \mathbb{R}^n \to \mathbb{C}$  is a given function and we want to find a solution to the equation Lf = g. Taking the Fourier transform of both sides of the equation Lf = gwould imply  $p(\xi) \hat{f}(\xi) = \hat{g}(\xi)$  and therefore  $\hat{f}(\xi) = \hat{g}(\xi) / p(\xi)$  provided  $p(\xi)$ is never zero. (We will discuss what happens when  $p(\xi)$  has zeros a bit more later on.) So we should expect

$$f(x) = \mathcal{F}^{-1}\left(\frac{1}{p(\xi)}\hat{g}(\xi)\right)(x) = \mathcal{F}^{-1}\left(\frac{1}{p(\xi)}\right) \bigstar g(x)$$

**Definition 6.1.** Let  $L = p(D_x)$  as in Eq. (6.1). Then we let  $\sigma(L) := \operatorname{Ran}(p) \subset \mathbb{C}$  and call  $\sigma(L)$  the **spectrum** of L. Given a measurable function  $G : \sigma(L) \to \mathbb{C}$ , we define (a possibly unbounded operator)  $G(L) : L^2(\mathbb{R}^n, m) \to L^2(\mathbb{R}^n, m)$  by

$$G(L)f := \mathcal{F}^{-1}M_{G \circ p}\mathcal{F}$$

where  $M_{G \circ p}$  denotes the operation on  $L^2(\mathbb{R}^n, m)$  of multiplication by  $G \circ p$ , i.e.

$$M_{G \circ p} f = (G \circ p) f$$

with domain given by those  $f \in L^2$  such that  $(G \circ p) f \in L^2$ .

At a formal level we expect

$$G(L)f = \mathcal{F}^{-1}\left(G \circ p\right) \bigstar g$$

# 6.1 Elliptic examples

As a specific example consider the equation

$$\left(-\Delta + m^2\right)f = g \tag{6.2}$$

where  $f, g : \mathbb{R}^n \to \mathbb{C}$  and  $\Delta = \sum_{i=1}^n \partial^2 / \partial x_i^2$  is the usual Laplacian on  $\mathbb{R}^n$ . By Corollary ?? (i.e. taking the Fourier transform of this equation), solving Eq. (6.2) with  $f, g \in L^2$  is equivalent to solving

$$\left(|\xi|^2 + m^2\right)\hat{f}(\xi) = \hat{g}(\xi).$$
 (6.3)

The unique solution to this latter equation is

$$\hat{f}(\xi) = \left( |\xi|^2 + m^2 \right)^{-1} \hat{g}(\xi)$$

and therefore,

$$f(x) = \mathcal{F}^{-1}\left(\left(\left|\xi\right|^2 + m^2\right)^{-1}\hat{g}\left(\xi\right)\right)(x) =: \left(-\Delta + m^2\right)^{-1}g(x)$$

We expect

$$\mathcal{F}^{-1}\left(\left(\left|\xi\right|^{2}+m^{2}\right)^{-1}\hat{g}\left(\xi\right)\right)(x)=G_{m}\star g(x)=\int_{\mathbb{R}^{n}}G_{m}(x-y)g\left(y\right)\mathrm{d}y,$$

where

$$G_m(x) := \mathcal{F}^{-1} \left( |\xi|^2 + m^2 \right)^{-1} (x) = \int_{\mathbb{R}^n} \frac{1}{m^2 + |\xi|^2} e^{i\xi \cdot x} \mathbf{d}\xi$$

At the moment  $\mathcal{F}^{-1}\left(|\xi|^2 + m^2\right)^{-1}$  only makes sense when n = 1 or 2 because only then is  $\left(|\xi|^2 + m^2\right)^{-1} \in L^2(\mathbb{R}^n).$ 

For now we will restrict our attention to the one dimensional case, n = 1, in which case

$$G_m(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{1}{(\xi + mi)(\xi - mi)} e^{i\xi x} d\xi.$$
(6.4)

#### 38 6 Constant Coefficient partial differential equations

The function  $G_m$  may be computed using standard complex variable contour integration methods to find, for  $x \ge 0$ ,

$$G_m(x) = \frac{1}{\sqrt{2\pi}} 2\pi i \frac{e^{i^2 m x}}{2im} = \frac{1}{2m} \sqrt{2\pi} e^{-mx}$$

and since  $G_m$  is an even function,

$$G_m(x) = \mathcal{F}^{-1} \left( |\xi|^2 + m^2 \right)^{-1} (x) = \frac{\sqrt{2\pi}}{2m} e^{-m|x|}.$$
 (6.5)

This result is easily verified to be correct, since

$$\mathcal{F}\left[\frac{\sqrt{2\pi}}{2m}e^{-m|x|}\right](\xi) = \frac{\sqrt{2\pi}}{2m} \int_{\mathbb{R}} e^{-m|x|} e^{-ix\cdot\xi} \mathbf{d}x$$
$$= \frac{1}{2m} \left( \int_{0}^{\infty} e^{-mx} e^{-ix\cdot\xi} dx + \int_{-\infty}^{0} e^{mx} e^{-ix\cdot\xi} dx \right)$$
$$= \frac{1}{2m} \left( \frac{1}{m+i\xi} + \frac{1}{m-i\xi} \right) = \frac{1}{m^2 + \xi^2}.$$

Hence in conclusion we find that  $(-\Delta + m^2) f = g$  has solution given by

$$f(x) = G_m \bigstar g(x) = \frac{\sqrt{2\pi}}{2m} \int_{\mathbb{R}} e^{-m|x-y|} g(y) \, \mathrm{d}y = \frac{1}{2m} \int_{\mathbb{R}} e^{-m|x-y|} g(y) \, \mathrm{d}y.$$

**Question.** Why do we get a unique answer here given that  $f(x) = A \sinh(x) + B \cosh(x)$  solves

$$\left(-\Delta + m^2\right)f = 0?$$

The answer is that such an f is not in  $L^2$  unless f = 0! More generally it is worth noting that  $A\sinh(x) + B\cosh(x)$  is not in  $\mathcal{P}$  unless A = B = 0.

What about when m = 0 in which case  $m^2 + \xi^2$  becomes  $\xi^2$  which has a zero at 0. Noting that constants are solutions to  $\Delta f = 0$ , we might look at

$$\lim_{m \downarrow 0} (G_m(x) - 1) = \lim_{m \downarrow 0} \frac{\sqrt{2\pi}}{2m} (e^{-m|x|} - 1) = -\frac{\sqrt{2\pi}}{2} |x|.$$

as a solution, i.e. we might conjecture that

$$f(x) := -\frac{1}{2} \int_{\mathbb{R}} |x - y| g(y) dy$$

solves the equation -f'' = g. To verify this we have

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

so that

$$\begin{aligned} f'(x) &= -\frac{1}{2} \int_{-\infty}^{x} g\left(y\right) dy + \frac{1}{2} \int_{x}^{\infty} g\left(y\right) dy \text{ and} \\ f''(x) &= -\frac{1}{2} g(x) - \frac{1}{2} g(x). \end{aligned}$$

### 6.2 Heat Equation on $\mathbb{R}^n$

The heat equation for a function  $u:\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{C}$  is the partial differential equation

$$\left(\partial_t - \frac{1}{2}\Delta\right)u = 0 \text{ with } u(0,x) = f(x), \tag{6.6}$$

where f is a given function on  $\mathbb{R}^n$ . By Fourier transforming Eq. (6.6) in the x – variables only, one finds that (6.6) implies that

$$\left(\partial_t + \frac{1}{2} |\xi|^2\right) \hat{u}(t,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi).$$
(6.7)

and hence that  $\hat{u}(t,\xi)=e^{-t|\xi|^2/2}\hat{f}\left(\xi\right).$  Inverting the Fourier transform then shows that

$$u(t,x) = \mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\hat{f}(\xi)\right)(x) = \left(\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right) \bigstar f\right)(x) =: e^{t\Delta/2}f(x).$$

From Corollary 5.7,

$$\mathcal{F}^{-1}\left(e^{-t|\xi|^2/2}\right)(x) = p_t(x) = t^{-n/2}e^{-\frac{1}{2t}|x|^2}$$

and therefore,

$$u(t,x) = \int_{\mathbb{R}^n} p_t(x-y)f(y) \mathrm{d}y.$$

This suggests the following theorem.

Theorem 6.2. Let

$$\rho(t, x, y) := (2\pi t)^{-n/2} e^{-|x-y|^2/2t}$$
(6.8)

be the **heat kernel** on  $\mathbb{R}^n$ . Then

$$\left(\partial_t - \frac{1}{2}\Delta_x\right)\rho(t, x, y) = 0 \text{ and } \lim_{t\downarrow 0}\rho(t, x, y) = \delta_x(y), \tag{6.9}$$

Page: 38

where  $\delta_x$  is the  $\delta$  – function at x in  $\mathbb{R}^n$ . More precisely, if f is a continuous bounded (can be relaxed considerably) function on  $\mathbb{R}^n$ , then

$$u(t,x) = \int_{\mathbb{R}^n} \rho(t,x,y) f(y) dy$$

is a solution to Eq. (6.6) where  $u(0,x) := \lim_{t\downarrow 0} u(t,x)$ .

**Proof.** Direct computations show that  $(\partial_t - \frac{1}{2}\Delta_x)\rho(t, x, y) = 0$  and an application of Theorem 4.6 shows  $\lim_{t\downarrow 0} \rho(t, x, y) = \delta_x(y)$  or equivalently that  $\lim_{t\downarrow 0} \int_{\mathbb{R}^n} \rho(t, x, y) f(y) dy = f(x)$  uniformly on compact subsets of  $\mathbb{R}^n$ . This shows that  $\lim_{t\downarrow 0} u(t, x) = f(x)$  uniformly on compact subsets of  $\mathbb{R}^n$ .

This notation suggests that we should be able to compute the solution to g to  $(\varDelta-m^2)g=f$  using

$$g(x) = \left(m^2 - \Delta\right)^{-1} f(x) = \int_0^\infty \left(e^{-\left(m^2 - \Delta\right)t} f\right)(x) dt$$
$$= \int_0^\infty \left(e^{-m^2 t} p_{2t} \bigstar f\right)(x) dt,$$

a fact which is easily verified using the Fourier transform. This gives us a method to compute  $G_m(x)$  from the previous section, namely

$$G_m(x) = \int_0^\infty e^{-m^2 t} p_{2t}(x) dt = \int_0^\infty (2t)^{-n/2} e^{-m^2 t - \frac{1}{4t}|x|^2} dt.$$

We make the change of variables,  $\lambda = |x|^2/4t$   $(t = |x|^2/4\lambda, dt = -\frac{|x|^2}{4\lambda^2}d\lambda)$  to find

$$G_m(x) = \int_0^\infty (2t)^{-n/2} e^{-m^2 t - \frac{1}{4t}|x|^2} dt = \int_0^\infty \left(\frac{|x|^2}{2\lambda}\right)^{-n/2} e^{-m^2|x|^2/4\lambda - \lambda} \frac{|x|^2}{(2\lambda)^2} d\lambda$$
$$= \frac{2^{(n/2-2)}}{|x|^{n-2}} \int_0^\infty \lambda^{n/2-2} e^{-\lambda} e^{-m^2|x|^2/4\lambda} d\lambda.$$
(6.10)

In case n = 3, Eq. (6.10) becomes

$$G_m(x) = \frac{\sqrt{\pi}}{\sqrt{2}|x|} \int_0^\infty \frac{1}{\sqrt{\pi\lambda}} e^{-\lambda} e^{-m^2|x|^2/4\lambda} d\lambda = \frac{\sqrt{\pi}}{\sqrt{2}|x|} e^{-m|x|}$$

where the last equality follows from Exercise 5.4. Hence when n = 3 we have found

$$\left[ (m^2 - \Delta)^{-1} f(x) = G_m \bigstar f(x) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} \frac{\sqrt{\pi}}{\sqrt{2} |x - y|} e^{-m|x - y|} f(y) dy = \int_{\mathbb{R}^3} \frac{1}{4\pi |x - y|} e^{-m|x - y|} f(y) dy.$$
 (6.11)

The function  $\frac{1}{4\pi|x|}e^{-m|x|}$  is called the Yukawa potential.

Let us work out  $G_m(x)$  for n odd. By differentiating Eq. (??) of Exercise 5.4 we find

$$\int_0^\infty d\lambda \lambda^{k-1/2} e^{-\frac{1}{4\lambda}x^2} e^{-\lambda m^2} = \int_0^\infty d\lambda \frac{1}{\sqrt{\lambda}} e^{-\frac{1}{4\lambda}x^2} \left(-\frac{d}{da}\right)^k e^{-\lambda a}|_{a=m^2}$$
$$= \left(-\frac{d}{da}\right)^k \frac{\sqrt{\pi}}{\sqrt{a}} e^{-\sqrt{a}x} = p_{m,k}(x) e^{-mx}$$

where  $p_{m,k}(x)$  is a polynomial in x with deg  $p_m = k$  with

$$p_{m,k}(0) = \sqrt{\pi} \left( -\frac{d}{da} \right)^k a^{-1/2} |_{a=m^2} = \sqrt{\pi} \left( \frac{1}{2} \frac{3}{2} \dots \frac{2k-1}{2} \right) m^{2k+1}$$
$$= m^{2k+1} \sqrt{\pi} 2^{-k} (2k-1) !!.$$

Letting k - 1/2 = n/2 - 2 and m = 1 we find  $k = \frac{n-1}{2} - 2 \in \mathbb{N}$  for  $n = 3, 5, \ldots$ and we find

$$\int_0^\infty \lambda^{n/2-2} e^{-\frac{1}{4\lambda}x^2} e^{-\lambda} d\lambda = p_{1,k}(x) e^{-x} \text{ for all } x > 0.$$

Therefore,

$$G_m(x) = \frac{2^{(n/2-2)}}{|x|^{n-2}} \int_0^\infty \lambda^{n/2-2} e^{-\lambda} e^{-m^2|x|^2/4\lambda} d\lambda = \frac{2^{(n/2-2)}}{|x|^{n-2}} p_{1,n/2-2}(m|x|) e^{-m|x|}.$$

Now for even m, I think we get Bessel functions in the answer. (BRUCE: look this up.) Let us at least work out the asymptotics of  $G_m(x)$  for  $x \to \infty$ . To this end let

$$\psi(y) := \int_0^\infty \lambda^{n/2-2} e^{-(\lambda+\lambda^{-1}y^2)} d\lambda = y^{n-2} \int_0^\infty \lambda^{n/2-2} e^{-(\lambda y^2+\lambda^{-1})} d\lambda$$

The function  $f_y(\lambda) := (y^2 \lambda + \lambda^{-1})$  satisfies,

$$f'_y(\lambda) = (y^2 - \lambda^{-2})$$
 and  $f''_y(\lambda) = 2\lambda^{-3}$  and  $f''_y(\lambda) = -6\lambda^{-4}$ 

so by Taylor's theorem with remainder we learn

$$f_y(\lambda) \cong 2y + y^3(\lambda - y^{-1})^2$$
 for all  $\lambda > 0$ ,

see Figure 6.1 below. So by the usual asymptotics arguments,



Fig. 6.1. Plot of  $f_4$  and its second order Taylor approximation.

$$\begin{split} \psi(y) &\cong y^{n-2} \int_{(-\varepsilon+y^{-1},y^{-1}+\varepsilon)} \lambda^{n/2-2} e^{-(\lambda y^2+\lambda^{-1})} d\lambda \\ &\cong y^{n-2} \int_{(-\varepsilon+y^{-1},y^{-1}+\varepsilon)} \lambda^{n/2-2} \exp\left(-2y - y^3(\lambda - y^{-1})^2\right) d\lambda \\ &\cong y^{n-2} e^{-2y} \int_{\mathbb{R}} \lambda^{n/2-2} \exp\left(-y^3(\lambda - y^{-1})^2\right) d\lambda \ (\text{let } \lambda \to \lambda y^{-1}) \\ &= e^{-2y} y^{n-2} y^{-n/2+1} \int_{\mathbb{R}} \lambda^{n/2-2} \exp\left(-y(\lambda - 1)^2\right) d\lambda \\ &= e^{-2y} y^{n-2} y^{-n/2+1} \int_{\mathbb{R}} (\lambda + 1)^{n/2-2} \exp\left(-y\lambda^2\right) d\lambda. \end{split}$$

The point is we are still going to get exponential decay at  $\infty$ .

When m = 0, Eq. (6.10) becomes

$$G_0(x) = \frac{2^{(n/2-2)}}{|x|^{n-2}} \int_0^\infty \lambda^{n/2-1} e^{-\lambda} \frac{d\lambda}{\lambda} = \frac{2^{(n/2-2)}}{|x|^{n-2}} \Gamma(n/2-1)$$

where  $\Gamma(x)$  in the gamma function defined in Eq. (??). Hence for "reasonable" functions f (and  $n \neq 2$ ) we expect that (see Proposition 6.3 below)

$$\begin{aligned} (-\Delta)^{-1}f(x) &= G_0 \bigstar f(x) = 2^{(n/2-2)} \Gamma(n/2-1) (2\pi)^{-n/2} \int_{\mathbb{R}^n} \frac{1}{|x-y|^{n-2}} f(y) dy \\ &= \frac{1}{4\pi^{n/2}} \Gamma(n/2-1) \int_{\mathbb{R}^n} \frac{1}{|x-y|^{n-2}} f(y) dy. \end{aligned}$$

The function

$$G(x) := \frac{1}{4\pi^{n/2}} \Gamma(n/2 - 1) \frac{1}{|x|^{n-2}}$$
(6.12)

is a "Green's function" for  $-\Delta$ . Recall from Exercise ?? that, for n = 2k,  $\Gamma(\frac{n}{2} - 1) = \Gamma(k - 1) = (k - 2)!$ , and for n = 2k + 1,

$$\begin{split} \Gamma(\frac{n}{2}-1) &= \Gamma(k-1/2) = \Gamma(k-1+1/2) = \sqrt{\pi} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3)}{2^{k-1}} \\ &= \sqrt{\pi} \frac{(2k-3)!!}{2^{k-1}} \text{ where } (-1)!! =: 1. \end{split}$$

Hence

$$G(x) = \frac{1}{4} \frac{1}{|x|^{n-2}} \begin{cases} \frac{1}{\pi^k} (k-2)! & \text{if } n = 2k \\ \frac{1}{\pi^k} \frac{(2k-3)!!}{2^{k-1}} & \text{if } n = 2k+1 \end{cases}$$

and in particular when n = 3,

$$G(x) = \frac{1}{4\pi} \frac{1}{|x|}$$

which is consistent with Eq. (6.11) with m = 0.

**Proposition 6.3.** Let  $n \geq 3$  and for  $x \in \mathbb{R}^n$ , let  $\rho_t(x) = \rho(t, x, 0) := \left(\frac{1}{2\pi t}\right)^{n/2} e^{-\frac{1}{2t}|x|^2}$  (see Eq. (6.8))and G(x) be as in Eq. (6.12) so that

$$G(x) := \frac{C_n}{|x|^{n-2}} = \frac{1}{2} \int_0^\infty \rho_t(x) \, dt \text{ for } x \neq 0.$$

Then

$$-\varDelta \left( G \ast u \right) = -G \ast \varDelta u = u$$

for all  $u \in C_c^2(\mathbb{R}^n)$ .

**Proof.** For  $f \in C_c(\mathbb{R}^n)$ ,

$$G * f(x) = C_n \int_{\mathbb{R}^n} f(x-y) \frac{1}{|y|^{n-2}} dy$$

is well defined, since

$$\int_{\mathbb{R}^{n}} |f(x-y)| \frac{1}{|y|^{n-2}} dy \le M \int_{|y| \le R+|x|} \frac{1}{|y|^{n-2}} dy < \infty$$

where M is a bound on f and supp  $(f) \subset B(0, R)$ . Similarly,  $|x| \leq r$ , we have

$$\sup_{|x| \le r} |f(x-y)| \frac{1}{|y|^{n-2}} \le M \mathbb{1}_{\{|y| \le R+r\}} \frac{1}{|y|^{n-2}} \in L^1(dy),$$

Page: 40

6.2 Heat Equation on  $\mathbb{R}^n$  41

from which it follows that G \* f is a continuous function. Similar arguments show if  $f \in C_c^2(\mathbb{R}^n)$ , then  $G * f \in C^2(\mathbb{R}^n)$  and  $\Delta(G * f) = G * \Delta f$ . So to finish the proof it suffices to show  $G * \Delta u = u$ .

For this we now write, making use of Fubini-Tonelli, integration by parts, the fact that  $\partial_t \rho_t(y) = \frac{1}{2} \Delta \rho_t(y)$  and the dominated convergence theorem,

$$\begin{split} G * \Delta u \left( x \right) &= \frac{1}{2} \int_{\mathbb{R}^n} \Delta u \left( x - y \right) \left( \int_0^\infty \rho_t \left( y \right) dt \right) dy \\ &= \frac{1}{2} \int_0^\infty dt \int_{\mathbb{R}^n} \Delta u \left( x - y \right) \rho_t \left( y \right) dy \\ &= \frac{1}{2} \int_0^\infty dt \int_{\mathbb{R}^n} \Delta_y u \left( x - y \right) \rho_t \left( y \right) dy \\ &= \frac{1}{2} \int_0^\infty dt \int_{\mathbb{R}^n} u \left( x - y \right) \Delta_y \rho_t \left( y \right) dy \\ &= \int_0^\infty dt \int_{\mathbb{R}^n} u \left( x - y \right) \frac{d}{dt} \rho_t \left( y \right) dy \\ &= \lim_{\varepsilon \downarrow 0} \int_{\varepsilon}^\infty dt \int_{\mathbb{R}^n} u \left( x - y \right) \frac{d}{dt} \rho_t \left( y \right) dy \\ &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^n} u \left( x - y \right) \left( \int_{\varepsilon}^\infty \frac{d}{dt} \rho_t \left( y \right) dt \right) dy \\ &= -\lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^n} u \left( x - y \right) \rho_\varepsilon \left( y \right) dy = u \left( x \right), \end{split}$$

where in the last equality we have used the fact that  $\rho_t$  is an approximate  $\delta$  – sequence.

Remark 6.4 (Computing the Green's function by the Fourier Transform). Green's function via the Fourier transform. We wish to solve  $\Delta u = \delta$  and so taking the Fourier transform of this equation suggests we solve

$$-\left|\xi\right|^{2} \hat{u}\left(\xi\right) = (2\pi)^{-d/2} \implies \hat{u}\left(\xi\right) = -(2\pi)^{-d/2} \left|\xi\right|^{-2}$$

Therefore,  $u = \lim_{M \to \infty} u_M$  where

$$u_M(x) := (2\pi)^{-d/2} \int_{|\xi| \le M} \hat{u}(\xi) e^{i\xi \cdot x} d\xi = -\left(\frac{1}{2\pi}\right)^d \int_{|\xi| \le M} \frac{1}{|\xi|^2} e^{i\xi \cdot x} d\xi.$$

We now let  $\xi = |x|^{-1} k$  in this last integral to find,

$$u_{M}(x) = -\left(\frac{1}{2\pi}\right)^{d} \frac{1}{|x|^{d-2}} \int_{|k| \le M|x|} \frac{1}{|k|^{2}} e^{ik \cdot \hat{x}} d\xi$$

If we let

$$C\left(M\right):=-\left(\frac{1}{2\pi}\right)^{d}\int_{\left|k\right|\leq M}\frac{1}{\left|k\right|^{2}}e^{ik\cdot e_{d}}d\xi$$

then we have shown

$$u_M(x) = C(M|x|) \frac{1}{|x|^{d-2}}$$

Working in polar coordinates it then follows that

$$C(M) = c(d) \int_0^M dr \ r^{d-3} \int_0^\pi d\varphi e^{ir\cos\varphi} \sin^{d-2}\varphi \ .$$

Letting  $y = \cos \theta$  this becomes,

$$C(M) = c(d) \int_0^M dr \ r^{d-3} \int_{-1}^1 dy e^{iry} \left(1 - y^2\right)^{\frac{d-3}{2}}.$$

The case d = 3 is well known how to handle and we find

$$C_3(M) = k \int_0^M \frac{\sin r}{r} dr \to \tilde{k} \text{ as } M \to \infty.$$

Let us consider the case where d = 5 so that

$$C(M) = c(5) \int_0^M dr \ r^2 \int_{-1}^1 dy e^{iry} \left(1 - y^2\right).$$

Now in this cae

$$\int_{-1}^{1} dy e^{iry} \left(1 - y^2\right) = \left(\frac{1}{ir}\right)^2 \int_{-1}^{1} dy \left(\partial_y^2 e^{iry}\right) \left(1 - y^2\right)$$
$$= \left(\frac{1}{ir}\right)^2 \int_{-1}^{1} \left(\partial_y e^{iry}\right) 2y dy$$
$$= \frac{-2}{r^2} \left[e^{iry} y|_{-1}^1 - \int_{-1}^{1} e^{iry} dy\right]$$

and so

$$C\left(M\right) = k \int_{0}^{M} dr \; \left[\cos r - \frac{\sin r}{r}\right]$$

The only new term to consider is

$$\int_0^M dr \ \cos r = -\sin M$$

and hence

Page: 41

42 6 Constant Coefficient partial differential equations

$$u_M(x) = k \frac{1}{|x|^3} \left[ \sin(M|x|) \pm \int_0^{M|x|} dr \; \frac{\sin r}{r} \right]$$

We now need to argue that for any  $f \in L^1(\mathbb{R}^5, m)$ 

$$\lim_{M \to \infty} \int_{\mathbb{R}^5} f(x) \sin(M |x|) \, dx = 0$$

by the Riemann Lebesgue Lemma again. This in fact follows from the one dimensional version after going to polar coordinates and integrating out all of the angular variables. Putting this all together we eventually learn that

$$u(x) = \lim_{M \to \infty} u_M(x) = c \frac{1}{|x|^3}$$
 for  $d = 5$ .

# 6.3 Poisson Semi-Group

Let us now consider the problems of finding a function  $(x_0, x) \in [0, \infty) \times \mathbb{R}^n \to u(x_0, x) \in \mathbb{C}$  such that

$$\left(\frac{\partial^2}{\partial x_0^2} + \Delta\right) u = 0 \text{ with } u(0, \cdot) = f \in L^2(\mathbb{R}^n).$$
(6.13)

Let  $\hat{u}(x_0,\xi) := \int_{\mathbb{R}^n} u(x_0,x) e^{-ix\cdot\xi} dx$  denote the Fourier transform of u in the  $x \in \mathbb{R}^n$  variable. Then Eq. (6.13) becomes

$$\left(\frac{\partial^2}{\partial x_0^2} - \left|\xi\right|^2\right)\hat{u}(x_0,\xi) = 0 \text{ with } \hat{u}(0,\xi) = \hat{f}(\xi)$$
(6.14)

and the general solution to this differential equation ignoring the initial condition is of the form

$$\hat{u}(x_0,\xi) = A(\xi) e^{-x_0|\xi|} + B(\xi) e^{x_0|\xi|}$$
(6.15)

for some function  $A(\xi)$  and  $B(\xi)$ . Let us now impose the extra condition that  $u(x_0, \cdot) \in L^2(\mathbb{R}^n)$  or equivalently that  $\hat{u}(x_0, \cdot) \in L^2(\mathbb{R}^n)$  for all  $x_0 \geq 0$ . The solution in Eq. (6.15) will not have this property unless  $B(\xi)$  decays very rapidly at  $\infty$ . The simplest way to achieve this is to assume B = 0 in which case we now get a unique solution to Eq. (6.14), namely

$$\hat{u}(x_0,\xi) = \hat{f}(\xi) e^{-x_0|\xi|}.$$

Applying the inverse Fourier transform gives

$$u(x_0, x) = \mathcal{F}^{-1}\left[\hat{f}(\xi) e^{-x_0|\xi|}\right](x) =: \left(e^{-x_0\sqrt{-\Delta}}f\right)(x)$$

and moreover

$$\left(e^{-x_0\sqrt{-\Delta}}f\right)(x) = P_{x_0} * f(x)$$

where  $P_{x_0}(x) = (2\pi)^{-n/2} \left( \mathcal{F}^{-1} e^{-x_0 |\xi|} \right)(x)$ . From Exercise 5.5,

$$P_{x_0}(x) = (2\pi)^{-n/2} \left( \mathcal{F}^{-1} e^{-x_0|\xi|} \right)(x) = c_n \frac{x_0}{(x_0^2 + |x|^2)^{(n+1)/2}}$$

where

$$c_n = (2\pi)^{-n/2} \frac{\Gamma((n+1)/2)}{\sqrt{\pi} 2^{n/2}} = \frac{\Gamma((n+1)/2)}{2^n \pi^{(n+1)/2}}.$$

Hence we have proved the following proposition.

**Proposition 6.5.** For  $f \in L^2(\mathbb{R}^n)$ ,

$$e^{-x_0\sqrt{-\Delta}}f = P_{x_0} * f \text{ for all } x_0 \ge 0$$

and the function  $u(x_0, x) := e^{-x_0\sqrt{-\Delta}} f(x)$  is  $C^{\infty}$  for  $(x_0, x) \in (0, \infty) \times \mathbb{R}^n$  and solves Eq. (6.13).

# 6.4 Addendum: convolutions and Fourier Transforms involving measures

**Notation 6.6** If  $\mu$  is a finite (could be complex) measure on  $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$  and  $f: \mathbb{R}^d \to \mathbb{C}$  is a measurable function, let

$$\hat{\mu}(k) := c_d \int_{\mathbb{R}^d} e^{-ik \cdot x} d\mu(x) \, dx$$

and

$$f * \mu(x) := \int_{\mathbb{R}^d} f(x - y) \, d\mu(y)$$
$$f \bigstar \mu(x) := c_d \cdot f * \mu(x) = c_d \int_{\mathbb{R}^d} f(x - y) \, d\mu(y)$$

when these integrals are defined. As usual we let  $c_d := (2\pi)^{-d/2}$  in all of these formula.

Remark 6.7. If  $\varphi \in \mathcal{S} = \mathcal{S}(\mathbb{R}^d)$ , then

$$\begin{split} \int_{\mathbb{R}^d} \hat{\mu} \left( k \right) \varphi \left( k \right) dk &= c_d \int_{\mathbb{R}^d} dk \int_{\mathbb{R}^d} d\mu \left( x \right) e^{-ik \cdot x} \varphi \left( k \right) \\ &= c_d \int_{\mathbb{R}^d} d\mu \left( x \right) \int_{\mathbb{R}^d} dk e^{-ik \cdot x} \varphi \left( k \right) \\ &= \int_{\mathbb{R}^d} \hat{\varphi} \left( x \right) d\mu \left( x \right) \end{split}$$

which shows that  $\hat{\mu}$  is the Fourier transform of  $\mu$  in the sense of tempered distributions, i.e.  $T_{\hat{\mu}} = \hat{T}_{\mu}$ .

**Lemma 6.8.** If  $f \in L^1(\mathbb{R}^d, m)$ , then  $f \bigstar \mu \in L^1(\mathbb{R}^d)$  with

$$\|f \bigstar \mu\|_1 \le c_d \|f\|_1 |\mu| \left(\mathbb{R}^d\right)$$

and  $\widehat{f \star \mu} = \hat{\mu} \cdot \hat{f}$ .

**Proof.** Let  $d\mu = gd |\mu|$  where  $g : \mathbb{R}^d \to S^1$  is a measurable function and  $|\mu|$  is the total variation measure of  $\mu$ . Then

$$f \bigstar \mu(x) := c_d \int_{\mathbb{R}^d} f(x-y) g(y) d|\mu|(y)$$

and by Hölder's inequality for integrals (or by direct calculation) we find  $f \bigstar \mu(x)$  is well defined for *m*-a.e. *x* and

$$\|f \bigstar \mu\|_{1} \leq c_{d} \int_{\mathbb{R}^{d}} \|f(\cdot - y)\|_{1} |g(y)| d |\mu|(y) = c_{d} \|f\|_{1} |\mu|(\mathbb{R}^{d}).$$

For the last assertion, we compute using Fubini-Tonelli and the translation invariance of Lebesgue measure that

$$\begin{split} \widehat{\mathbf{f} \star \mu} \left( k \right) &= c_d^2 \int_{\mathbb{R}^d} dx e^{-ik \cdot x} \int_{\mathbb{R}^d} d\mu \left( y \right) f \left( x - y \right) \\ &= c_d^2 \int_{\mathbb{R}^d} d\mu \left( y \right) \int_{\mathbb{R}^d} dx e^{-ik \cdot x} f \left( x - y \right) \\ &= c_d^2 \int_{\mathbb{R}^d} d\mu \left( y \right) \int_{\mathbb{R}^d} dx e^{-ik \cdot (x+y)} f \left( x \right) \\ &= \hat{\mu} \left( k \right) \hat{f} \left( k \right). \end{split}$$

**Corollary 6.9.** Suppose  $\mu$  is a complex measure on  $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$  and  $f \in L^1(m)$  is such that  $\hat{f} \in L^1(m)$ . If  $f_0 \in C_0(\mathbb{R}^d)$  is the continuous version of f, then  $f_0 \bigstar \mu \in C_0(\mathbb{R}^d)$  and

$$(f_0 \bigstar \mu)(x) = c_d \int_{\mathbb{R}^d} \hat{\mu}(k) \, \hat{f}(k) \, e^{ik \cdot x} dk \text{ for all } x \in \mathbb{R}^d.$$
(6.16)

**Proof.** Since  $\hat{f}$  is assumed to be in  $L^1(\mathbb{R}^d)$  and  $\hat{\mu}$  is bounded it follows that  $\hat{\mu} \cdot \hat{f} \in L^1(\mathbb{R}^d)$  and hence by our basic Fourier inversion formula,

$$(f \bigstar \mu)(x) = \left(\hat{\mu} \cdot \hat{f}\right)^{\vee}(x) = c_d \int_{\mathbb{R}^d} \hat{\mu}(k) \, \hat{f}(k) \, e^{ik \cdot x} dk \text{ for } m\text{-a.e. } x$$

and in particular  $(f \bigstar \mu)(x)$  has a continuous version. Since  $f_0 \bigstar \mu$  is continuous and  $f_0 \bigstar \mu = f \bigstar \mu$  a.e. (can you prove these statements?) we conclude that Eq. (6.16) holds.

# 6.5 Wave Equation on $\mathbb{R}^n$

Let us now consider the wave equation on  $\mathbb{R}^n$ ,

$$0 = \left(\partial_t^2 - \Delta\right) u(t, x) \text{ with}$$
  

$$u(0, x) = f(x) \text{ and } u_t(0, x) = g(x).$$
(6.17)

Taking the Fourier transform in the x variables gives the following equation

$$0 = \hat{u}_{t\,t}(t,\xi) + |\xi|^2 \,\hat{u}(t,\xi) \text{ with}$$
$$\hat{u}(0,\xi) = \hat{f}(\xi) \text{ and } \hat{u}_t(0,\xi) = \hat{g}(\xi).$$
(6.18)

The solution to these equations is

$$\hat{u}(t,\xi) = \hat{f}(\xi)\cos(t|\xi|) + \hat{g}(\xi)\frac{\sin t|\xi|}{|\xi|}$$

and hence we should have

$$u(t,x) = \mathcal{F}^{-1}\left(\hat{f}\left(\xi\right)\cos\left(t\left|\xi\right|\right) + \hat{g}\left(\xi\right)\frac{\sin t\left|\xi\right|}{\left|\xi\right|}\right)(x)$$

$$= \mathcal{F}^{-1}\cos\left(t\left|\xi\right|\right) \bigstar f(x) + \mathcal{F}^{-1}\frac{\sin t\left|\xi\right|}{\left|\xi\right|} \bigstar g(x)$$

$$= \frac{d}{dt}\mathcal{F}^{-1}\left[\frac{\sin t\left|\xi\right|}{\left|\xi\right|}\right] \bigstar f(x) + \mathcal{F}^{-1}\left[\frac{\sin t\left|\xi\right|}{\left|\xi\right|}\right] \bigstar g(x).$$
(6.19)
(6.19)

job: Supplements

#### 44 6 Constant Coefficient partial differential equations

**Theorem 6.10 (One** d wave equation). If  $f, g \in L^1(\mathbb{R})$  such that  $\hat{f}, \hat{g} \in L^1(\mathbb{R})$ , then u(t,x) defined by Eq. (6.19) is given by

$$u(t,x) = \frac{1}{2} \left( f(x-t) + f(x+t) \right) + \frac{1}{2} \int_{x-t}^{x+t} g(y) dy.$$

Here we assume that f and g are already chosen to be the continuous version of f and g.

**Proof.** For  $t \in \mathbb{R}$ ,

$$\hat{\delta}_t \left( k \right) = c_1 \cdot e^{-ikt} = \frac{1}{\sqrt{2\pi}} e^{-ikt}$$

and hence

$$(\delta_t + \delta_{-t})^{-}(k) = 2c_1 \cos kt.$$

Thus we may conclude that

$$\mathcal{F}^{-1}\left(\hat{f}\left(\cdot\right)\cos\left(t\left(\cdot\right)\right)\right) = \mathcal{F}^{-1}\left(\hat{f}\left(\cdot\right)\frac{1}{2c_{1}}\left(\delta_{t}+\delta_{-t}\right)^{2}\right)$$
$$= \frac{1}{2c_{1}}f\bigstar\left(\delta_{t}+\delta_{-t}\right)(x)$$
$$= \frac{1}{2}\left[f\left(x+t\right)+f\left(x-t\right)\right].$$

Similarly for t > 0 if we let  $d\mu_t(x) = 1_{[-t,t]}(x) dm(x)$ , then

$$\hat{\mu}_t \left( k \right) = c_1 \int_{-t}^t e^{-ikx} dx = c_1 \frac{e^{-ikx}}{-ik} |_{x=-t}^t = 2c_1 \frac{\sin tk}{k}.$$

Hence we may conclude that

$$\mathcal{F}^{-1}\left(k \to \hat{g}(k) \,\frac{\sin(tk)}{k}\right)(x) = \frac{1}{2c_1} \mathcal{F}^{-1}\left[\hat{g}\hat{\mu}_t\right](x) \\ = \frac{1}{2}\left(g * \mu_t\right)(x) = \frac{1}{2} \int_{-t}^t g\left(x - y\right) dm\left(y\right).$$

By making the change of variable, z = x - y we find

$$\frac{1}{2} \int_{-t}^{t} g(x-y) \, dm(y) = \frac{1}{2} \int_{x-t}^{x+t} g(z) \, dz.$$

Combining all of these results gives the desired conclusion.

 $Remark\ 6.11.$  We can arrive at this same solution by more elementary means as follows. We first note in the one dimensional case that wave operator factors, namely

$$0 = \left(\partial_t^2 - \partial_x^2\right) u(t, x) = \left(\partial_t - \partial_x\right) \left(\partial_t + \partial_x\right) u(t, x).$$

Let  $U(t,x) := (\partial_t + \partial_x) u(t,x)$ , then the wave equation states  $(\partial_t - \partial_x) U = 0$ and hence by the chain rule  $\frac{d}{dt}U(t,x-t) = 0$ . So

$$U(t, x - t) = U(0, x) = g(x) + f'(x)$$

and replacing x by x + t in this equation shows

$$\left(\partial_t + \partial_x\right)u(t, x) = U(t, x) = g(x+t) + f'(x+t).$$

Working similarly, we learn that

$$\frac{d}{dt}u(t,x+t) = g(x+2t) + f'(x+2t)$$

which upon integration implies

$$u(t, x + t) = u(0, x) + \int_0^t \left\{ g(x + 2\tau) + f'(x + 2\tau) \right\} d\tau$$
  
=  $f(x) + \int_0^t g(x + 2\tau) d\tau + \frac{1}{2} f(x + 2\tau) |_0^t$   
=  $\frac{1}{2} \left( f(x) + f(x + 2t) \right) + \int_0^t g(x + 2\tau) d\tau.$ 

Replacing  $x \to x - t$  in this equation gives

$$u(t,x) = \frac{1}{2} \left( f(x-t) + f(x+t) \right) + \int_0^t g(x-t+2\tau) d\tau$$

and then letting  $y = x - t + 2\tau$  in the last integral shows again that

$$u(t,x) = \frac{1}{2} \left( f(x-t) + f(x+t) \right) + \frac{1}{2} \int_{x-t}^{x+t} g(y) dy.$$

Our next goal is to solve the wave equation in dimension 3. As we will see shortly it will be very useful to first compute that Fourier transform of the surface measure on three spheres in  $\mathbb{R}^3$ .

**Lemma 6.12.** If  $\sigma_t$  is the surface measure on the sphere  $S_t \subset \mathbb{R}^3$  of radius t centered at zero in  $\mathbb{R}^3$ , then

$$\hat{\sigma}_t\left(\xi\right) := c_3 \int_{S_t} e^{-ix \cdot \xi} d\sigma_t\left(x\right) = c_3 \cdot 4\pi t \frac{\sin t \left|\xi\right|}{\left|\xi\right|}.$$

**Proof.** To show this we work in spherical coordinates where

$$\int_{S_t} f(x) \, d\sigma_t(x) = \int_0^{2\pi} d\theta \int_0^{\pi} d\varphi t^2 \sin \varphi \, f(t \sin \varphi \cos \theta, t \sin \varphi \sin \theta, t \cos \varphi) \, .$$

By rotation invariance or  $\sigma_{t}(\cdot)$  we know that

$$\begin{split} \frac{1}{c_3} \hat{\sigma}_t \left( \xi \right) &= \int_{S_t} e^{-i|\xi| e_3 \cdot x} d\sigma_t \left( x \right) = t^2 \int_0^{2\pi} d\theta \int_0^{\pi} d\varphi \sin \varphi \, e^{-i|\xi| t \cos \varphi} \\ &= 2\pi t^2 \cdot \int_0^{\pi} d\varphi \sin \varphi \, e^{-i|\xi| t \cos \varphi}. \end{split}$$

Let  $u = \cos \varphi$  so that  $du = -\sin \varphi d\varphi$ , to find,

$$\frac{1}{c_3}\hat{\sigma}_t\left(\xi\right) = 2\pi t^2 \int_{-1}^1 e^{-itu|\xi|} du = 2\pi t^2 \frac{1}{-it|\xi|} e^{-itu|\xi|} \Big|_{u=-1}^{u=1} = 4\pi t^2 \frac{\sin t |\xi|}{t|\xi|}.$$

**Theorem 6.13 (Three-***d* wave equation). If f = 0 and  $g \in L^1(\mathbb{R}^3)$  such that  $\hat{g} \in L^1(\mathbb{R}^3)$ , then u(t,x) defined by Eq. (6.19) is given by

$$u(t,x) = \frac{1}{4\pi t} \int_{S_t} g(x-y) \, d\sigma_t(y)$$
$$= t \int_{S_t} g(x-y) \, d\bar{\sigma}_t(y)$$
$$= t \int_{S_1} g(x+t\omega) d\bar{\sigma}_1(\omega)$$

where  $\bar{\sigma}_t := \frac{1}{4\pi t^2} \sigma_t$  is the normalized surface measure on  $S_t$ . Here we assume that g has already been chosen to be its continuous version. More generally if  $f \neq 0$  and f is sufficiently nice, then

$$u(t,x) = \frac{d}{dt} \left[ t \int_{S_1} f(x+t\omega) d\bar{\sigma}_1(\omega) \right] + t \int_{S_1} g(x+t\omega) d\bar{\sigma}_1(\omega) \,. \tag{6.21}$$

If we further assume f

**Proof.** Since

$$\frac{\sin\left(t\,|k|\right)}{|k|} = \frac{1}{c_3 \cdot 4\pi t}\hat{\sigma}_t$$

we may conclude that

$$\mathcal{F}^{-1}\left(k \to \hat{g}\left(k\right) \frac{\sin\left(t\left|k\right|\right)}{\left|k\right|}\right)\left(x\right) = \frac{1}{c_3 \cdot 4\pi t} \mathcal{F}^{-1}\left[\hat{g}\hat{\sigma}_t\right]\left(x\right)$$
$$= \frac{1}{4\pi t} \left(g * \sigma_t\right)\left(x\right) = \frac{1}{4\pi t} \int_{S_t} g\left(x - y\right) d\sigma_t\left(y\right).$$

**Proposition 6.14.** Suppose  $f \in C^3(\mathbb{R}^3)$  and  $g \in C^2(\mathbb{R}^3)$ , then u(t,x) defined by Eq. (6.21) is in  $C^2(\mathbb{R} \times \mathbb{R}^3)$  and is a classical solution of the wave equation in Eq. (6.17).

**Proof.** The fact that  $u \in C^2(\mathbb{R} \times \mathbb{R}^3)$  follows by the usual differentiation under the integral arguments. Suppose we can prove the proposition in the special case that  $f \equiv 0$ . Then for  $f \in C^3(\mathbb{R}^3)$ , the function  $v(t, x) = +t \int_{S_1} g(x + t\omega) d\bar{\sigma}_1(\omega)$  solves the wave equation  $0 = (\partial_t^2 - \Delta) v(t, x)$  with v(0, x) = 0 and  $v_t(0, x) = g(x)$ . Differentiating the wave equation in t shows  $u = v_t$  also solves the wave equation with u(0, x) = g(x) and  $u_t(0, x) = v_{tt}(0, x) = -\Delta_x v(0, x) =$ 0. These remarks reduced the problems to showing u in Eq. (6.21) with  $f \equiv 0$ solves the wave equation. So let

$$u(t,x) := t \int_{S_1} g(x+t\omega) d\bar{\sigma}_1(\omega) \,. \tag{6.22}$$

We now give two proofs the u solves the wave equation.

**Proof 1.** Since solving the wave equation is a local statement and u(t, x) only depends on the values of g in B(x, t) we it suffices to consider the case where  $g \in C_c^2(\mathbb{R}^3)$ . Taking the Fourier transform of Eq. (6.22) in the x variable shows

$$\begin{split} \hat{u}(t,\xi) &= t \int_{S_1} d\bar{\sigma}_1\left(\omega\right) \int_{\mathbb{R}^3} g(x+t\omega) e^{-i\xi \cdot x} \mathbf{d}x \\ &= t \int_{S_1} d\bar{\sigma}_1\left(\omega\right) \int_{\mathbb{R}^3} g(x) e^{-i\xi \cdot x} e^{it\omega \cdot \xi} \mathbf{d}x = \hat{g}\left(\xi\right) t \int_{S_1} e^{it\omega \cdot \xi} d\bar{\sigma}_1\left(\omega\right) \\ &= \hat{g}\left(\xi\right) t \frac{\sin|tk|}{|tk|} = \hat{g}\left(\xi\right) \frac{\sin\left(t\left|\xi\right|\right)}{|\xi|} \end{split}$$

wherein we have made use of Example ??. This completes the proof since  $\hat{u}(t,\xi)$  solves Eq. (6.18) as desired.

**Proof 2.** Differentiating

$$S(t,x) := \int_{S_1} g(x+t\omega) d\bar{\sigma}_1\left(\omega\right)$$

in t gives

46 6 Constant Coefficient partial differential equations

$$\begin{split} S_t(t,x) &= \frac{1}{4\pi} \int_{S_1} \nabla g(x+t\omega) \cdot \omega d\sigma \left(\omega\right) \\ &= \frac{1}{4\pi} \int_{B(0,1)} \nabla_\omega \cdot \nabla g(x+t\omega) dm \left(\omega\right) \\ &= \frac{t}{4\pi} \int_{B(0,1)} \Delta g(x+t\omega) dm \left(\omega\right) \\ &= \frac{1}{4\pi t^2} \int_{B(0,t)} \Delta g(x+y) dm(y) \\ &= \frac{1}{4\pi t^2} \int_0^t dr \ r^2 \int_{|y|=r} \Delta g(x+y) d\sigma(y) \end{split}$$

where we have used the divergence theorem, made the change of variables  $y = t\omega$ and used the disintegration formula in Eq. (??),

$$\int_{\mathbb{R}^d} f(x) dm(x) = \int_{[0,\infty) \times S^{n-1}} f(r \ \omega) \ d\sigma(\omega) \ r^{n-1} dr = \int_0^\infty dr \int_{|y|=r} f(y) d\sigma(y) d\sigma(y)$$

Since u(t, x) = tS(t, x) if follows that

$$\begin{split} u_{tt}(t,x) &= \frac{\partial}{\partial t} \left[ S(t,x) + tS_t(t,x) \right] \\ &= S_t(t,x) + \frac{\partial}{\partial t} \left[ \frac{1}{4\pi t} \int_0^t dr \ r^2 \int_{|y|=r} \Delta g(x+y) d\sigma(y) \right] \\ &= S_t(t,x) - \frac{1}{4\pi t^2} \int_0^t dr \int_{|y|=r} \Delta g(x+y) d\sigma(y) \\ &\quad + \frac{1}{4\pi t} \int_{|y|=t} \Delta g(x+y) d\sigma(y) \\ &= S_t(t,x) - S_t(t,x) + \frac{t}{4\pi t^2} \int_{|y|=1} \Delta g(x+t\omega) d\sigma(\omega) \\ &= t \Delta u(t,x) \end{split}$$

as required.

The solution in Eq. (6.21) exhibits a basic property of wave equations, namely finite propagation speed. To exhibit the finite propagation speed, suppose that f = 0 (for simplicity) and g has compact support near the origin, for example think of  $g = \delta_0(x)$ . Then x + tw = 0 for some w iff |x| = t. Hence the "wave front" propagates at unit speed and the wave front is sharp. See Figure 6.2 below.

The solution of the two dimensional wave equation may be found using "Hadamard's method of decent" which we now describe. Suppose now that f



Fig. 6.2. The geometry of the solution to the wave equation in three dimensions. The observer sees a flash at t = 0 and x = 0 only at time t = |x|. The wave progates sharply with speed 1.

and g are functions on  $\mathbb{R}^2$  which we may view as functions on  $\mathbb{R}^3$  which happen not to depend on the third coordinate. We now go ahead and solve the three dimensional wave equation using Eq. (6.21) and f and g as initial conditions. It is easily seen that the solution u(t, x, y, z) is again independent of z and hence is a solution to the two dimensional wave equation. See figure 6.3 below.

Notice that we still have finite speed of propagation but no longer sharp propagation. The explicit formula for u is given in the next proposition.

**Proposition 6.15.** Suppose  $f \in C^3(\mathbb{R}^2)$  and  $g \in C^2(\mathbb{R}^2)$ , then

$$\begin{split} u(t,x) &:= \frac{\partial}{\partial t} \left[ \frac{t}{2\pi} \iint_{D_1} \frac{f(x+tw)}{\sqrt{1-|w|^2}} dm\left(w\right) \right] \\ &+ \frac{t}{2\pi} \iint_{D_1} \frac{g(x+tw)}{\sqrt{1-|w|^2}} dm\left(w\right) \end{split}$$

is in  $C^2(\mathbb{R} \times \mathbb{R}^2)$  and solves the wave equation in Eq. (6.17).

**Proof.** As usual it suffices to consider the case where  $f \equiv 0$ . By symmetry u may be written as

$$u(t,x) = 2t \int_{S_t^+} g(x-y) d\bar{\sigma}_t(y) = 2t \int_{S_t^+} g(x+y) d\bar{\sigma}_t(y)$$

where  $S_t^+$  is the portion of  $S_t$  with  $z \ge 0$ . The surface  $S_t^+$  may be parametrized by  $R(u, v) = (u, v, \sqrt{t^2 - u^2 - v^2})$  with  $(u, v) \in D_t := \{(u, v) : u^2 + v^2 \le t^2\}$ . In these coordinates we have



$$\begin{split} u(t,x) &= \frac{1}{2\pi} \operatorname{sgn}(t) \iint_{D_t} \frac{g(x+w)}{\sqrt{t^2 - |w|^2}} dm \left( w \right) \\ &= \frac{1}{2\pi} \operatorname{sgn}(t) \frac{t^2}{|t|} \iint_{D_1} \frac{g(x+tw)}{\sqrt{1 - |w|^2}} dm \left( w \right) \\ &= \frac{1}{2\pi} t \iint_{D_1} \frac{g(x+tw)}{\sqrt{1 - |w|^2}} dm \left( w \right). \end{split}$$

**Fig. 6.3.** The geometry of the solution to the wave equation in two dimensions. A flash at  $0 \in \mathbb{R}^2$  looks like a line of flashes to the fictitious 3 – d observer and hence she sees the effect of the flash for  $t \ge |x|$ . The wave still propagates with speed 1. However there is no longer sharp propagation of the wave front, similar to water waves.

$$\begin{aligned} 4\pi t^2 d\bar{\sigma}_t &= \left| \left( -\partial_u \sqrt{t^2 - u^2 - v^2}, -\partial_v \sqrt{t^2 - u^2 - v^2}, 1 \right) \right| du dv \\ &= \left| \left( \frac{u}{\sqrt{t^2 - u^2 - v^2}}, \frac{v}{\sqrt{t^2 - u^2 - v^2}}, 1 \right) \right| du dv \\ &= \sqrt{\frac{u^2 + v^2}{t^2 - u^2 - v^2}} + 1 du dv = \frac{|t|}{\sqrt{t^2 - u^2 - v^2}} du dv \end{aligned}$$

and therefore,

$$\begin{split} u(t,x) &= \frac{2t}{4\pi t^2} \int_{D_t} g(x + (u,v,\sqrt{t^2 - u^2 - v^2})) \frac{|t|}{\sqrt{t^2 - u^2 - v^2}} du dv \\ &= \frac{1}{2\pi} \text{sgn}(t) \int_{D_t} \frac{g(x + (u,v))}{\sqrt{t^2 - u^2 - v^2}} du dv. \end{split}$$

This may be written as

# Radon Measures and the Dual of $C_0(X)$

In this chapter let X be a locally compact Hausdorff space and  $\mathcal{B} = \mathcal{B}_X$  is the Borel  $\sigma$ -algebra on X. Open subsets of  $\mathbb{R}^d$  and locally compact separable metric spaces are examples of such spaces. In this chapter will only state and discuss the Riesz-Markov theorem which associates measures to linear functionals on  $C_c(X)$  and  $C_0(X)$ . We will give a number of examples of using this theorem.

# 7.1 The Riesz-Markov Theorem

**Definition 7.1.** A linear functional I on  $C_c(X)$  is **positive** if  $I(f) \ge 0$  for all  $f \in C_c(X, [0, \infty))$ .

If I is a positive linear functional on  $C_{c}(X)$  and  $f \in C_{c}(X, \mathbb{R})$ , then

$$I(f) = I(f_{+}) - I(f_{-}) \in \mathbb{R}$$

where  $f = f_+ - f_-$  and  $f_{\pm} = \max(0, \pm f) \ge 0$ . That is positive linear functionals are real on real functions.

**Proposition 7.2.** If I is a positive linear functional on  $C_c(X)$  and K is a compact subset of X, then there exists  $C_K < \infty$  such that  $|I(f)| \leq C_K ||f||_{\infty}$  for all  $f \in C_c(X)$  with supp  $(f) \subset K$ .

**Proof.** By Urysohn's Lemma ??, there exists  $\varphi \in C_c(X, [0, 1])$  such that  $\varphi = 1$  on K. Then for all  $f \in C_c(X, \mathbb{R})$  such that  $\operatorname{supp}(f) \subset K$ ,  $|f| \leq ||f||_{\infty} \varphi$  or equivalently  $||f||_{\infty} \varphi \pm f \geq 0$ . Hence  $||f||_{\infty} I(\varphi) \pm I(f) \geq 0$  or equivalently which is to say  $|I(f)| \leq ||f||_{\infty} I(\varphi)$ . Letting  $C_K := I(\varphi)$ , we have shown that  $|I(f)| \leq C_K ||f||_{\infty}$  for all  $f \in C_c(X, \mathbb{R})$  with  $\operatorname{supp}(f) \subset K$ . For general  $f \in C_c(X, \mathbb{C})$  with  $\operatorname{supp}(f) \subset K$ , choose  $|\alpha| = 1$  such that  $\alpha I(f) \geq 0$ . Then

$$I(f)| = \alpha I(f) = I(\alpha f) = I(\operatorname{Re}(\alpha f)) \le C_K ||\operatorname{Re}(\alpha f)||_{\infty} \le C_K ||f||_{\infty}.$$

*Example 7.3.* Let  $\mu$  be a K-finite measure on  $(X, \mathcal{B}_X)$ , i.e.  $\mu(K) < \infty$  for all compact subsets of X. Then

$$I_{\mu}(f) = \int_{X} f d\mu \; \forall f \in C_{c}(X)$$

defines a positive linear functional on  $C_c(X)$ . In the future, we will often simply write  $\mu(f)$  for  $I_{\mu}(f)$ .

The Riesz-Markov Theorem 7.11 below asserts that every positive linear functional on  $C_c(X)$  comes from a K-finite measure  $\mu$ .

*Example 7.4.* Let  $X = \mathbb{R}$  and  $\tau = \tau_d = 2^X$  be the discrete topology on X. Now let  $\mu(A) = 0$  if A is countable and  $\mu(A) = \infty$  otherwise. Since  $K \subset X$  is compact iff  $\#(K) < \infty$ ,  $\mu$  is a K-finite measure on X and

$$I_{\mu}(f) = \int_{X} f d\mu = 0 \text{ for all } f \in C_{c}(X).$$

This shows that the correspondence  $\mu \to I_{\mu}$  from K-finite measures to positive linear functionals on  $C_c(X)$  is not injective without further restriction.

**Definition 7.5.** Suppose that  $\mu$  is a Borel measure on X and  $B \in \mathcal{B}_X$ . We say  $\mu$  is inner regular on B if

$$\mu(B) = \sup\{\mu(K) : K \sqsubset B\}$$
(7.1)

and  $\mu$  is outer regular on B if

$$\mu(B) = \inf\{\mu(U) : B \subset U \subset_o X\}.$$
(7.2)

The measure  $\mu$  is said to be a **regular Borel measure** on X, if it is both inner and outer regular on all Borel measurable subsets of X.

**Definition 7.6.** A measure  $\mu : \mathcal{B}_X \to [0, \infty]$  is a **Radon measure** on X if  $\mu$  is a K-finite measure which is inner regular on all open subsets of X and outer regular on all Borel subsets of X. In full detail;

- 1.  $\mu(K) < \infty$  for all compact subsets  $K \subset X$ , i.e.  $\mu$  is K-finite.
- 2.  $\mu(V) = \sup \{\mu(K) : K \sqsubset V\}$  for all  $V \in \tau$ , i.e.  $\mu$  is inner regular on open sets.
- 3.  $\mu(B) = \inf \{ \mu(V) : B \subset V \in \tau \}$  for all  $B \in \mathcal{B}_X$ , i.e.  $\mu$  is outer regular. [Clearly in verifying this property it suffices to assume  $\mu(B) < \infty$ .]

The measure in Example 7.4 is an example of a K-finite measure on X which is not a Radon measure on X.

#### 50 7 Radon Measures and the Dual of $C_0(X)$

*Example 7.7.* If the topology on a set, X, is the discrete topology, then a measure  $\mu$  on  $\mathcal{B}_X$  is a Radon measure iff  $\mu$  is of the form

$$\mu = \sum_{x \in X} \mu_x \delta_x \tag{7.3}$$

where  $\mu_x \in [0, \infty)$  for all  $x \in X$ . To verify this first notice that  $\mathcal{B}_X = \tau_X = 2^X$ and hence every measure on  $\mathcal{B}_X$  is necessarily outer regular on all subsets of X. The measure  $\mu$  is K-finite iff  $\mu_x := \mu(\{x\}) < \infty$  for all  $x \in X$ . If  $\mu$  is a Radon measure, then for  $A \subset X$  we have, by inner regularity,

$$\mu(A) = \sup \left\{ \mu(A) : A \subset_f A \right\} = \sup \left\{ \sum_{x \in A} \mu_x : A \subset_f A \right\} = \sum_{x \in A} \mu_x.$$

On the other hand if  $\mu$  is given by Eq. (7.3) and  $A \subset X$ , then

$$\mu(A) = \sum_{x \in A} \mu_x = \sup \left\{ \mu(A) = \sum_{x \in A} \mu_x : A \subset_f A \right\}$$

showing  $\mu$  is inner regular on all (open) subsets of X.

Example 7.8. Let X be an uncountable set and  $\tau = 2^X$  be the discrete topology on X. If we let  $\mu$  be counting measure on X, and  $\nu$  be the measure defined by  $\nu(A) = 0$  if A is a finite or countable set and  $\nu(A) = \infty$  if A is un-countable, then  $\mu$  is a Radon measure and  $\nu \leq \mu$ , yet  $\nu$  is not a Radon measure. Thus being dominated by a radon measure is not sufficient to imply a measure is Radon.

**Exercise 7.1.** Suppose that  $(X, \tau)$  is a LCH and  $\mu$  and  $\nu$  are two positive measures on  $(X, \mathcal{B}_X)$ .

- 1. If  $\nu \leq \mu$  and  $\mu$  is a finite Radon measure, then  $\nu$  is a finite Radon measure.
- 2. If both  $\mu$  and  $\nu$  are Radon measures then  $\mu + \nu$  is also a Radon measure. [This does not hold for countable sums of Radon measures as such a sum may not even be K - finite.]
- 3. If there exists constants  $A, B \in (0, \infty)$  such that  $\mu \leq A\nu$  and  $\nu \leq B\mu$ , then  $\mu$  is a Radon measure iff  $\nu$  is a Radon measure.

#### Exercise 7.2.

#### Example 7.9. Exercise 7.3.

Recall from Definition ?? that if U is an open subset of X, we write  $f \prec U$  to mean that  $f \in C_c(X, [0, 1])$  with  $\operatorname{supp}(f) := \overline{\{f \neq 0\}} \subset U$ .

**Notation 7.10** Given a positive linear functional, I, on  $C_c(X)$  define  $\mu = \mu_I$  on  $\mathcal{B}_X$  by

$$\mu(U) = \sup\{I(f) : f \prec U\}$$
(7.4)

for all  $U \subset_o X$  and then define

$$\mu(B) = \inf\{\mu(U) : B \subset U \text{ and } U \text{ is open}\}.$$
(7.5)

**Theorem 7.11 (Riesz-Markov Theorem).** The map  $\mu \to I_{\mu}$  taking Radon measures on X to positive linear functionals on  $C_c(X)$  is bijective. Moreover if I is a positive linear functional on  $C_c(X)$ , the function  $\mu := \mu_I$  defined in Notation 7.10 has the following properties.

- 1.  $\mu$  is a Radon measure on X and the map  $I \to \mu_I$  is the inverse to the map  $\mu \to I_{\mu}$ .
- 2. For all compact subsets  $K \subset X$ ,

$$\mu(K) = \inf\{I(f) : 1_K \le f \prec X\}.$$
(7.6)

3. If  $||I_{\mu}||$  denotes the dual norm of  $I = I_{\mu}$  on  $C_c(X, \mathbb{R})^*$ , then  $||I|| = \mu(X)$ . In particular, the linear functional,  $I_{\mu}$ , is bounded iff  $\mu(X) < \infty$ .

**Proof.** (Also see Theorem ?? and related material about the Daniel integral.) The proof of the surjectivity of the map  $\mu \to I_{\mu}$  and the assertion in item 1. is the content of Theorem ?? below.

**Injectivity of**  $\mu \to I_{\mu}$ . Suppose that  $\mu$  is a sa Radon measure on X. To each open subset  $U \subset X$  let

$$\mu_0(U) := \sup\{I_\mu(f) : f \prec U\}.$$
(7.7)

It is evident that  $\mu_0(U) \leq \mu(U)$  because  $f \prec U$  implies  $f \leq 1_U$ . Given a compact subset  $K \subset U$ , Urysohn's Lemma ?? implies there exists  $f \prec U$  such that f = 1 on K. Therefore,

$$\mu(K) \le \int_X f d\mu \le \mu_0(U) \le \mu(U) \tag{7.8}$$

By assumption  $\mu$  is inner regular on open sets, and therefore taking the supremum of Eq. (7.8) over compact subsets, K, of U shows

$$\mu(U) = \mu_0(U) = \sup\{I_\mu(f) : f \prec U\}.$$
(7.9)

If  $\mu$  and  $\nu$  are two Radon measures such that  $I_{\mu} = I_{\nu}$ . Then by Eq. (7.9) it follows that  $\mu = \nu$  on all open sets. Then by outer regularity,  $\mu = \nu$  on  $\mathcal{B}_X$  and this shows the map  $\mu \to I_{\mu}$  is injective.

Item 2. Let  $K \subset X$  be a compact set, then by monotonicity of the integral,

$$\mu(K) \le \inf\{I_{\mu}(f) : f \in C_{c}(X) \text{ with } f \ge 1_{K}\}.$$
(7.10)

To prove the reverse inequality, choose, by outer regularity,  $U \subset_o X$  such that  $K \subset U$  and  $\mu(U \setminus K) < \varepsilon$ . By Urysohn's Lemma ?? there exists  $f \prec U$  such that f = 1 on K and hence,

$$I_{\mu}(f) = \int_{X} f \ d\mu = \mu(K) + \int_{U \setminus K} f \ d\mu \le \mu(K) + \mu(U \setminus K) < \mu(K) + \varepsilon.$$

Consequently,

$$\inf\{I_{\mu}(f): f \in C_{c}(X) \text{ with } f \ge 1_{K}\} < \mu(K) + \varepsilon$$

and because  $\varepsilon > 0$  was arbitrary, the reverse inequality in Eq. (7.10) holds and Eq. (7.6) is verified.

**Item 3.** If  $f \in C_c(X)$ , then

i

$$|I_{\mu}(f)| \leq \int_{X} |f| \, d\mu = \int_{\mathrm{supp}(f)} |f| \, d\mu \leq ||f||_{\infty} \, \mu(\mathrm{supp}\,(f)) \leq ||f||_{\infty} \, \mu(X)$$
(7.11)

and thus  $||I_{\mu}|| \leq \mu(X)$ . For the reverse inequality let K be a compact subset of X and use Urysohn's Lemma ?? again to find a function  $f \prec X$  such that f = 1 on K. By Eq. (7.8) we have

$$\mu(K) \le \int_{X} f d\mu = I_{\mu}(f) \le \|I_{\mu}\| \|f\|_{\infty} = \|I_{\mu}\|,$$

which by the inner regularity of  $\mu$  on open sets implies

$$\mu(X) = \sup\{\mu(K) : K \sqsubset X\} \le \|I_{\mu}\|$$

Example 7.12 (Discrete Version of Theorem 7.11). Suppose X is a set,  $\tau = 2^X$  is the discrete topology on X and for  $x \in X$ , let  $e_x \in C_c(X)$  be defined by  $e_x(y) = 1_{\{x\}}(y)$ . Let I be positive linear functional on  $C_c(X)$  and define a Radon measure,  $\mu$ , on X by

$$\mu(A) := \sum_{x \in A} I(e_x) \text{ for all } A \subset X.$$

Then for  $f \in C_c(X)$  (so f is a complex valued function on X supported on a finite set),

$$\int_{X} f d\mu = \sum_{x \in X} f(x) I(e_x) = I\left(\sum_{x \in X} f(x) e_x\right) = I(f),$$

so that  $I = I_{\mu}$ . It is easy to see in this example that  $\mu$  defined above is the unique regular radon measure on X such that  $I = I_{\mu}$  while example Example 7.4 shows the uniqueness is lost if the regularity assumption is dropped.

# 7.2 Classifying Radon Measures on $\mathbb R$

Throughout this section, let  $X = \mathbb{R}$ ,  $\mathcal{E}$  be the elementary class

$$\mathcal{E} = \{ (a, b] \cap \mathbb{R} : -\infty \le a \le b \le \infty \}, \tag{7.12}$$

and  $\mathcal{A} = \mathcal{A}(\mathcal{E})$  be the algebra formed by taking finite disjoint unions of elements from  $\mathcal{E}$ , see Proposition ??. The aim of this section is to prove again the following theorem.

**Theorem 7.13.** The collection of K-finite measure on  $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$  are in one to one correspondence with a right continuous non-decreasing functions,  $F : \mathbb{R} \to \mathbb{R}$ , with F(0) = 0. The correspondence is as follows. If F is a right continuous non-decreasing function  $F : \mathbb{R} \to \mathbb{R}$ , then there exists a unique measure,  $\mu_F$ , on  $\mathcal{B}_{\mathbb{R}}$  such that

$$\mu_F((a,b]) = F(b) - F(a) \quad \forall -\infty < a \le b < \infty$$

and this measure may be defined by

$$u_F(A) = \inf\left\{\sum_{i=1}^{\infty} (F(b_i) - F(a_i)) : A \subset \bigcup_{i=1}^{\infty} (a_i, b_i]\right\}$$
$$= \inf\left\{\sum_{i=1}^{\infty} (F(b_i) - F(a_i)) : A \subset \sum_{i=1}^{\infty} (a_i, b_i]\right\}$$
(7.13)

for all  $A \in \mathcal{B}_{\mathbb{R}}$ . Conversely if  $\mu$  is K-finite measure on  $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ , then

$$F(x) := \begin{cases} -\mu((x,0]) & \text{if } x \le 0\\ \mu((0,x]) & \text{if } x \ge 0 \end{cases}$$
(7.14)

is a right continuous non-decreasing function and this map is the inverse to the map,  $F \rightarrow \mu_F$ .

There are three aspects to this theorem; namely the existence of the map  $F \to \mu_F$ , the surjectivity of the map and the injectivity of this map. Assuming the map  $F \to \mu_F$  exists, the surjectivity follows from Eq. (7.14) and the injectivity is an easy consequence of Theorem ??. The rest of this section is devoted to giving two proofs for the existence of the map  $F \to \mu_F$ .

**Exercise 7.4.** Show by direct means any measure  $\mu = \mu_F$  satisfying Eq. (7.13) is outer regular on all Borel sets. **Hint:** it suffices to show if  $B := \sum_{i=1}^{\infty} (a_i, b_i]$ , then there exists  $V \subset_o \mathbb{R}$  such that  $\mu(V \setminus B)$  is as small as you please.

#### 52 7 Radon Measures and the Dual of $C_0(X)$

# 7.3 Classifying Radon Measures on $\mathbb{R}$ using Theorem 7.11

**Notation 7.14** Given an increasing function  $F : \mathbb{R} \to \mathbb{R}$ , let  $F(x-) = \lim_{y \uparrow x} F(y)$ ,  $F(x+) = \lim_{y \downarrow x} F(y)$  and  $F(\pm \infty) = \lim_{x \to \pm \infty} F(x) \in \mathbb{R}$ . Since F is increasing all of theses limits exists.

Let  $\mathcal{A}$  be the algebra of subsets of  $\mathbb{R}$  generated by the elementary class,

$$\mathcal{E} = \{ (a, b] \cap \mathbb{R} : -\infty \le a < b \le \infty \}.$$

Given any increasing (i.e. non-decreasing) function,  $F : \mathbb{R} \to \mathbb{R}$ , there exists a unique finitely additive measure,  $\nu_F : \mathcal{A} \to [0, \infty]$  such that

$$\nu_F\left((a,b] \cap \mathbb{R}\right) = F\left(b\right) - F\left(a\right)$$

where  $F(\infty) := \lim_{x \uparrow \infty} F(x)$  and  $F(-\infty) := \lim_{x \downarrow -\infty} F(x)$ . Let S denote the  $\mathcal{A}$ -simple functions with compact support which we express as

$$f = \sum_{i=1}^{n} c_i \mathbf{1}_{(a_i, b_i]} \in \mathbb{S}$$

where  $\{(a_i, b_i)\}_{i=1}^n$  are pairwise disjoint sets and  $c_i \in \mathbb{C}$ . For such an f we have

$$\int_{\mathbb{R}} f d\nu_F = \sum_{i=1}^{n} c_i \nu_F \left( (a_i, b_i] \right) = \sum_{i=1}^{n} c_i \left[ F(b_i) - F(a_i) \right]$$

and if  $a = \min \{a_i : 1 \le i \le n\}$  and  $b = \max \{a_i : 1 \le i \le n\}$ , then

$$\left| \int_{\mathbb{R}} f d\nu_F \right| \le \sum_{i=1}^n |c_i| |F(b_i) - F(a_i)| \le ||f||_u \cdot \nu_F((a, b]) \le ||f||_u \cdot [F(b) - F(a)].$$

Let  $I_F$  denote the extension of  $\mathbb{S} \ni f \to \int_{\mathbb{R}} f d\nu_F$  to the closure  $\tilde{\mathbb{S}}$  of  $\mathbb{S}$  in the uniform norm sense with all functions in a sequence being supported in a fixed compact integral.

Remark 7.15. A few remarks are now in order.

1. If  $f \in \tilde{\mathbb{S}}$  and  $f \ge 0$ , there exists  $f_n \in \mathbb{S}$  such that  $||f - f_n||_u \to 0$ . Since  $f_n \lor 0 \in \mathbb{S}$  and

$$\|f - f_n \vee 0\|_u \le \|f - f_n\|_u \to 0 \text{ as } n \to \infty,$$

we may assume that  $f_n \ge 0$  for all n. Therefore,

$$I_F(f) = \lim_{n \to \infty} \int_{\mathbb{R}} f_n d\nu_F \ge 0.$$

i.e.  $I_F$  is still positive.

2.  $C_c(\mathbb{R}, \mathbb{R}) \subset \tilde{\mathbb{S}}$ . Indeed, if  $f \in C_c(\mathbb{R}, \mathbb{R})$  and choose a < b such that supp  $(f) \subset (a, b)$  and suppose that

$$\pi_n = \{ a = a_0^n < a_1^n < \dots < a_{N_n}^n = b \},\$$

for  $n = 1, 2, 3, \ldots$ , is a sequence of refining partitions such that  $\operatorname{mesh}(\pi_n) \to 0$  as  $n \to \infty$ . Then define  $f_n \in \mathbb{S}$  by

$$f_n(x) = \sum_{l=0}^{N_n - 1} \min\left\{f(x) : a_l^n \le x \le a_{l+1}^n\right\} \mathbf{1}_{(a_l^n, a_{l+1}^n]}(x).$$

Since f is continuous and compactly supported it is uniformly continuous on  $\mathbb{R}$  and hence  $||f - f_n||_u \to 0$  as  $n \to \infty$ . Thus by the BLT theorem

$$I_F(f) := \lim_{n \to \infty} \int_{\mathbb{R}} f_n d\nu_F$$
  
= 
$$\lim_{n \to \infty} \sum_{l=0}^{N_n - 1} \min \left\{ f(x) : a_l^n \le x \le a_{l+1}^n \right\} \left[ F\left(a_{l+1}^n\right) - F\left(a_l^n\right) \right].$$

Consequently, λ<sub>F</sub> := I<sub>F</sub>|<sub>C<sub>c</sub>(ℝ,ℝ)</sub> is a positive linear functional on C<sub>c</sub>(ℝ, ℝ).
 By the Riesz-Markov Theorem 7.11, there exists a unique Radon measure, μ, on (ℝ, B<sub>ℝ</sub>) such that

$$\lambda_F(f) = \mu(f)$$
 for all  $f \in C_c(\mathbb{R}, \mathbb{R})$ .

**Theorem 7.16.** The measure  $\mu$  constructed above is the unique measure on  $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$  such that

$$\mu((a, b]) = F(b+) - F(a+) \text{ for all } -\infty < a < b < \infty.$$

[In general  $\mu$  and  $\nu_F$  need not agree on  $\mathcal{A}$  unless F is right continuous!]

**Proof.** Let  $-\infty < a < b < \infty$ ,  $\varepsilon > 0$  be small and  $\chi_{\varepsilon}(x)$  be the function defined in Figure 7.1.



**Fig. 7.1.** Approximating the characteristic function,  $1_{(a,b]}$ .

Since  $\chi_{\varepsilon} \to 1_{(a,b]}$  boundedly and having supports ins a fixed compact set, it follows by the dominated convergence theorem that

$$\mu((a,b]) = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \chi_{\varepsilon} d\mu = \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} \chi_{\varepsilon} d\nu_F.$$
(7.15)

On the other hand we have

$$1_{(a+2\varepsilon,b+\varepsilon]} \le \chi_{\varepsilon} \le 1_{(a+\varepsilon,b+2\varepsilon]},$$

and therefore applying  $I_F$  to this equation gives the inequalities;

$$F(b+\varepsilon) - F(a+2\varepsilon) = I_F \left( \mathbf{1}_{(a+2\varepsilon,b+\varepsilon]} \right)$$
  
$$\leq \int_{\mathbb{R}} \chi_{\varepsilon} d\nu_F$$
  
$$\leq I_F \left( \mathbf{1}_{(a+\varepsilon,b+2\varepsilon]} \right) = F(b+2\varepsilon) - F(a+\varepsilon)$$

Letting  $\varepsilon \downarrow 0$  in this equation and using Eq. (7.15) shows

$$F(b+) - F(a+) \le \mu((a,b]) \le F(b+) - F(a+).$$

# 7.4 Kolmogorov's Existence of Measure on Products Spaces

Throughout this section, let  $\{(X_{\alpha}, \tau_{\alpha})\}_{\alpha \in A}$  be second countable locally compact Hausdorff spaces and let  $X := \prod_{\alpha \in A} X_{\alpha}$  be equipped with the product topology,  $\tau := \bigotimes_{\alpha \in A} \tau_{\alpha}$ . More generally for  $\Lambda \subset A$ , let  $X_{\Lambda} := \prod_{\alpha \in \Lambda} X_{\alpha}$  and  $\tau_{\Lambda} := \bigotimes_{\alpha \in \Lambda} \tau_{\alpha}$ and  $\Lambda \subset \Gamma \subset A$ , let  $\pi_{\Lambda,\Gamma} : X_{\Gamma} \to X_{\Lambda}$  be the projection map;  $\pi_{\Lambda,\Gamma}(x) = x|_{\Lambda}$  for  $x \in X_{\Gamma}$ . We will simply write  $\pi_{\Lambda}$  for  $\pi_{\Lambda,A} : X \to X_{\Lambda}$ . (Notice that if  $\Lambda$  is a finite subset of A then  $(X_{\Lambda}, \tau_{\Lambda})$  is still second countable as the reader should verify.) Let  $\mathcal{M} = \bigotimes_{\alpha \in A} \mathcal{B}_{\alpha}$  be the product  $\sigma$ -algebra on  $X = X_{A}$  and  $\mathcal{B}_{\Lambda} = \sigma(\tau_{\Lambda})$  be the Borel  $\sigma$ -algebra on  $X_{\Lambda}$ .

**Theorem 7.17 (Kolmogorov's Existence Theorem).** Suppose  $\{\mu_A : A \subset_f A\}$  are probability measures on  $(X_A, \mathcal{B}_A)$  satisfying the following compatibility condition:

•  $(\pi_{\Lambda,\Gamma})_* \mu_{\Gamma} = \mu_{\Lambda}$  whenever  $\Lambda \subset \Gamma \subset_f \Lambda$ .

Then there exists a unique probability measure,  $\mu$ , on  $(X, \mathcal{M})$  such that  $(\pi_A)_* \mu = \mu_A$  whenever  $\Lambda \subset_f A$ . Recall, see Exercise ??, that the condition  $(\pi_A)_* \mu = \mu_A$  is equivalent to the statement;

$$\int_{X} F(\pi_{\Lambda}(x)) d\mu(x) = \int_{X_{\Lambda}} F(y) d\mu_{\Lambda}(y)$$
(7.16)

for all  $\Lambda \subset_f A$  and  $F: X_\Lambda \to \mathbb{R}$  bounded a measurable.

We will first prove the theorem in the following special case. The full proof will be given after Exercise 7.5 below.

**Theorem 7.18.** Theorem 7.17 holds under the additional assumption that each of the spaces,  $\{(X_{\alpha}, \tau_{\alpha})\}_{\alpha \in A}$ , are **compact** second countable and Hausdorff and A is countable.

**Proof.** Recall from Theorem ?? that the Borel  $\sigma$ -algebra,  $\mathcal{B}_{\Lambda} = \sigma(\tau_{\Lambda})$ , and the product  $\sigma$ -algebra,  $\bigotimes_{\alpha \in \Lambda} \mathcal{B}_{\alpha}$ , are the same for any  $\Lambda \subset A$ . By Tychonoff's Theorem ?? and Proposition ??, X and  $X_{\Lambda}$  for any  $\Lambda \subset A$  are still compact Hausdorff spaces which are second countable if  $\Lambda$  is finite. By the Stone Weierstrass Theorem ??,

$$\mathbb{D} := \{ f \in C(X) : f = F \circ \pi_A \text{ with } F \in C(X_A) \text{ and } \Lambda \subset_f A \}$$

is a dense subspace of C(X). For  $f = F \circ \pi_A \in \mathbb{D}$ , let

$$I(f) = \int_{X_A} F \circ \pi_A(x) \, d\mu_A(x) \,. \tag{7.17}$$

Let us verify that I is well defined. Suppose that f may also be expressed as  $f = F' \circ \pi_{A'}$  with  $A' \subset_f A$  and  $F' \in C(X_{A'})$ . Let  $\Gamma := A \cup A'$  and define  $G \in C(X_{\Gamma})$  by  $G := F \circ \pi_{A,\Gamma}$ . Hence, using Exercise ??,

$$\int_{X_{\Gamma}} G d\mu_{\Gamma} = \int_{X_{\Gamma}} F \circ \pi_{\Lambda,\Gamma} \ d\mu_{\Gamma} = \int_{X_{\Lambda}} F \ d\left[ \left( \pi_{\Lambda,\Gamma} \right)_{*} \mu_{\Gamma} \right] = \int_{X_{\Lambda}} F \ d\mu_{\Lambda}$$

wherein we have used the compatibility condition in the last equality. Similarly, using  $G = F' \circ \pi_{A',\Gamma}$  (as the reader should verify), one shows

$$\int_{X_{\varGamma}} G d\mu_{\varGamma} = \int_{X_{\Lambda'}} F' \ d\mu_{\Lambda'}.$$

Therefore

$$\int_{X_{\Lambda'}} F' \ d\mu_{\Lambda'} = \int_{X_{\Gamma}} G d\mu_{\Gamma} = \int_{X_{\Lambda}} F \ d\mu_{\Lambda},$$

which shows I in Eq. (7.17) is well defined.

Since  $|I(f)| \leq ||f||_{\infty}$ , the B.L.T. Theorem ?? allows us to extend I from the dense subspace,  $\mathbb{D}$ , to a continuous linear functional,  $\bar{I}$ , on C(X). Because I was positive on  $\mathbb{D}$ , it is easy to check that  $\bar{I}$  is still positive on C(X). So by the Riesz-Markov Theorem 7.11, there exists a Radon measure on  $\mathcal{B} = \mathcal{M}$  such that  $\bar{I}(f) = \int_{X} f d\mu$  for all  $f \in C(X)$ . By the definition of  $\bar{I}$  in now follows that

$$\int_{X^{\Lambda}} Fd(\pi_{\Lambda})_{*} \mu = \int_{X^{\Lambda}} F \circ \pi_{\Lambda} d\mu = \bar{I}(F \circ \pi_{\Lambda}) = \int_{X^{\Lambda}} Fd\mu_{\Lambda}$$

Page: 53

#### 54 7 Radon Measures and the Dual of $C_0(X)$

for all  $F \in C(X^A)$  and  $A \subset_f A$ . Since  $X_A$  is a second countable locally compact Hausdorff space, this identity implies, see Theorem ??<sup>1</sup>, that  $(\pi_A)_* \mu = \mu_A$ . The uniqueness assertion of the theorem follows from the fact that the measure  $\mu$  is determined uniquely by its values on the algebra  $\mathcal{A} := \bigcup_{A \subset_f A} \pi_A^{-1}(\mathcal{B}_{X_A})$  which generates  $\mathcal{B} = \mathcal{M}$ , see Theorem ??.

**Exercise 7.5.** Let  $(Y, \tau)$  be a locally compact Hausdorff space and  $(Y^* = Y \cup \{\infty\}, \tau^*)$  be the one point compactification of Y. Then

$$\mathcal{B}_{Y^*} := \sigma(\tau^*) = \{ A \subset Y^* : A \cap Y \in \mathcal{B}_Y = \sigma(\tau) \}$$

or equivalently put

$$\mathcal{B}_{Y^*} = \mathcal{B}_Y \cup \{A \cup \{\infty\} : A \in \mathcal{B}_Y\}.$$

Also shows that  $(Y^* = Y \cup \{\infty\}, \tau^*)$  is second countable if  $(Y, \tau)$  was second countable.

#### Proof. Proof of Theorem 7.17.

**Case 1;** A is a countable. Let  $(X_{\alpha}^* = X_{\alpha} \cup \{\infty_{\alpha}\}, \tau_{\alpha}^*)$  be the one point compactification of  $(X_{\alpha}, \tau_{\alpha})$ . For  $\Lambda \subset A$ , let  $X_{\Lambda}^* := \prod_{\alpha \in \Lambda} X_{\alpha}^*$  equipped with the product topology and Borel  $\sigma$ -algebra,  $\mathcal{B}_{\Lambda}^*$ . Since  $\Lambda$  is at most countable, the set,

$$X_A := \bigcap_{\alpha \in A} \left\{ \pi_\alpha = \infty_\alpha \right\},\,$$

is a measurable subset of  $X_{\Lambda}^*$ . Therefore for each  $\Lambda \subset_f \Lambda$ , we may extend  $\mu_{\Lambda}$  to a measure,  $\bar{\mu}_{\Lambda}$ , on  $(X_{\Lambda}^*, \mathcal{B}_{\Lambda}^*)$  using the formula,

$$\bar{\mu}_{\Lambda}(B) = \mu_{\Lambda}(B \cap X_{\Lambda}) \text{ for all } B \in X_{\Lambda}^*.$$

An application of Theorem 7.18 shows there exists a unique probability measure,  $\bar{\mu}$ , on  $X^* := X^*_A$  such that  $(\pi_A)_* \bar{\mu} = \bar{\mu}_A$  for all  $A \subset_f A$ . Since

$$X^* \setminus X = \bigcup_{\alpha \in A} \{ \pi_\alpha = \infty_\alpha \}$$

and  $\bar{\mu}(\{\pi_{\alpha} = \infty\}) = \bar{\mu}_{\{\alpha\}}(\{\infty_{\alpha}\}) = 0$ , it follows that  $\bar{\mu}(X^* \setminus X) = 0$ . Hence  $\mu := \bar{\mu}|_{\mathcal{B}_X}$  is a probability measure on  $(X, \mathcal{B}_X)$ . Finally if  $B \in \mathcal{B}_X \subset \mathcal{B}_{X^*}$ ,

$$\mu_{\Lambda}(B) = \bar{\mu}_{\Lambda}(B) = (\pi_{\Lambda})_{*} \bar{\mu}(B) = \bar{\mu}(\pi_{\Lambda}^{-1}(B))$$
$$= \bar{\mu}(\pi_{\Lambda}^{-1}(B) \cap X) = \mu(\pi_{\Lambda}|_{X}^{-1}(B))$$

which shows  $\mu$  is the required probability measure on  $\mathcal{B}_X$ .

**Case 2;** A is uncountable. By case 1. for each countable or finite subset  $\Gamma \subset A$  there is a measure  $\mu_{\Gamma}$  on  $(X_{\Gamma}, \mathcal{B}_{\Gamma})$  such that  $(\pi_{\Lambda,\Gamma})_* \mu_{\Gamma} = \mu_{\Lambda}$  for all  $\Lambda \subset_f \Gamma$ . By Exercise ??,

$$\mathcal{M} = \bigcup \left\{ \pi_{\Gamma}^{-1} \left( \mathcal{B}_{\Gamma} \right) : \Gamma \text{ is a countable subset of } A \right\},\$$

i.e. every  $B \in \mathcal{M}$  may be written in the form  $B = \pi_{\Gamma}^{-1}(C)$  for some countable subset,  $\Gamma \subset A$ , and  $C \in \mathcal{B}_{\Gamma}$ . For such a B we define  $\mu(B) := \mu_{\Gamma}(C)$ . It is left to the reader to check that  $\mu$  is well defined and that  $\mu$  is a measure on  $\mathcal{M}$ . (Keep in mind the countable union of countable sets is countable.) If  $\Lambda \subset_f A$ and  $C \in \mathcal{B}_A$ , then

$$[(\pi_{\Lambda})_{*} \mu](C) = \mu(\pi_{\Lambda}^{-1}(C)) := \mu_{\Lambda}(C),$$

i.e.  $(\pi_A)_* \mu = \mu_A$  as desired.

**Corollary 7.19.** Suppose that  $\{\mu_{\alpha}\}_{\alpha \in A}$  are probability measure on  $(X_{\alpha}, \mathcal{B}_{\alpha})$ for all  $\alpha \in A$  and if  $\Lambda \subset_f A$  let  $\mu_{\Lambda} := \otimes_{\alpha \in \Lambda} \mu_{\alpha}$  be the product measure on  $(X_{\Lambda}, \mathcal{B}_{\Lambda} = \otimes_{\alpha \in \Lambda} \mathcal{B}_{\alpha})$ . Then there exists a unique probability measure,  $\mu$ , on  $(X, \mathcal{M})$  such that  $(\pi_{\Lambda})_* \mu = \mu_{\Lambda}$  for all  $\Lambda \subset_f A$ . (It is possible remove the topology from this corollary, see Theorem ?? below.)

**Exercise 7.6.** Prove Corollary 7.19 by showing the measures  $\mu_{\Lambda} := \bigotimes_{\alpha \in \Lambda} \mu_{\alpha}$  satisfy the compatibility condition in Theorem 7.17.

# 7.5 The dual of $C_0(X)$

**Definition 7.20.** Let  $(X, \tau)$  be a locally compact Hausdorff space and  $\mathcal{B} = \sigma(\tau)$ be the Borel  $\sigma$ -algebra. A signed Radon measure is a signed measure  $\mu$  on  $\mathcal{B}$ such that the measures,  $\mu_{\pm}$ , in the Jordan decomposition of  $\mu$  are both Radon measures. A complex Radon measure is a complex measure  $\mu$  on  $\mathcal{B}$  such that  $\operatorname{Re} \mu$  and  $\operatorname{Im} \mu$  are signed radon measures.

**Exercise 7.7.** If  $(X, \tau)$  is a LCH and  $\mu$  is a finite signed measure on  $(X, \mathcal{B}_X)$  then  $\mu_{\pm}$  are Radon measures iff  $|\mu|$  is a Radon measure.

**Exercise 7.8.** If  $(X, \tau)$  is a LCH and  $\mu$  is a complex measure on  $(X, \mathcal{B}_X)$  then the following are equivalent;

1.  $|\mu|$  is a Radon measure,

- 2.  $|\operatorname{Re} \mu|$  and  $|\operatorname{Im} \mu|$  are Radon measures, and
- 3.  $(\operatorname{Re} \mu)_+$  and  $(\operatorname{Im} \mu)_+$  are Radon measures.

<sup>&</sup>lt;sup>1</sup> Alternatively, use Theorems ?? and the uniqueess assertion in Markov-Riesz Theorem 7.11 to conclude  $(\pi_A)_* \mu = \mu_A$ .

Thus any one of the above conditions may be used as the definition of a complex measure,  $\mu$ , being a Radon measure.

*Example 7.21.* Every complex measure  $\mu$  on  $\mathcal{B}_{\mathbb{R}^d}$  is a Radon measure. BRUCE: add some more examples and perhaps some exercises here.

**Proposition 7.22.** Suppose  $(X, \tau)$  is a topological space and  $I \in C_0(X, \mathbb{R})^*$ . Then we may write  $I = I_+ - I_-$  where  $I_{\pm} \in C_0(X, \mathbb{R})^*$  are positive linear functionals.

**Proof.** For  $f \in C_0(X, [0, \infty))$ , let  $I_+(f) := \sup \{ I(g) : g \in C_0(X, [0, \infty)) \text{ and } g \le f \}$ 

and notice that  $|I_{+}(f)| \leq ||I|| ||f||$ . If c > 0, then  $I_{+}(cf) = cI_{+}(f)$ . Suppose that  $f_{1}, f_{2} \in C_{0}(X, [0, \infty))$  and  $g_{i} \in C_{0}(X, [0, \infty))$  such that  $g_{i} \leq f_{i}$ , then  $g_{1} + g_{2} \leq f_{1} + f_{2}$  so that

$$I(g_1) + I(g_2) = I(g_1 + g_2) \le I_+(f_1 + f_2)$$

and therefore

$$I_{+}(f_{1}) + I_{+}(f_{2}) \le I_{+}(f_{1} + f_{2}).$$
(7.18)

Moreover, if  $g \in C_0(X, [0, \infty))$  and  $g \leq f_1 + f_2$ , let  $g_1 = \min(f_1, g)$ , so that

 $0 \le g_2 := g - g_1 \le f_1 - g_1 + f_2 \le f_2.$ 

Hence  $I(g) = I(g_1) + I(g_2) \le I_+(f_1) + I_+(f_2)$  for all such g and therefore,

$$I_{+}(f_{1}+f_{2}) \leq I_{+}(f_{1}) + I_{+}(f_{2}).$$
(7.19)

Combining Eqs. (7.18) and (7.19) shows that  $I_+(f_1+f_2) = I_+(f_1)+I_+(f_2)$ . For general  $f \in C_0(X, \mathbb{R})$ , let  $I_+(f) = I_+(f_+) - I_+(f_-)$  where  $f_+ = \max(f, 0)$  and  $f_- = -\min(f, 0)$ . (Notice that  $f = f_+ - f_-$ .) If f = h - g with  $h, g \in C_0(X, \mathbb{R})$ , then  $g + f_+ = h + f_-$  and therefore,

$$I_{+}(g) + I_{+}(f_{+}) = I_{+}(h) + I_{+}(f_{-})$$

and hence  $I_{+}(f) = I_{+}(h) - I_{+}(g)$ . In particular,

$$I_{+}(-f) = I_{+}(f_{-} - f_{+}) = I_{+}(f_{-}) - I_{+}(f_{+}) = -I_{+}(f)$$

so that  $I_+(cf) = cI_+(f)$  for all  $c \in \mathbb{R}$ . Also,

$$\begin{split} I_{+}\left(f+g\right) &= I_{+}(f_{+}+g_{+}-(f_{-}+g_{-})) = I_{+}(f_{+}+g_{+}) - I_{+}(f_{-}+g_{-}) \\ &= I_{+}(f_{+}) + I_{+}(g_{+}) - I_{+}(f_{-}) - I_{+}(g_{-}) \\ &= I_{+}\left(f\right) + I_{+}\left(g\right). \end{split}$$

Therefore  $I_+$  is linear. Moreover,

$$I_{+}(f)| \le \max\left(|I_{+}(f_{+})|, |I_{+}(f_{-})|\right) \le \|I\| \max\left(\|f_{+}\|, \|f_{-}\|\right) = \|I\| \|f\|$$

which shows that  $||I_+|| \leq ||I||$ . Let  $I_- = I_+ - I \in C_0(X, \mathbb{R})^*$ , then for  $f \geq 0$ ,

$$I_{-}(f) = I_{+}(f) - I(f) \ge 0$$

by definition of  $I_+$ , so  $I_- \ge 0$  as well.

Remark 7.23. The above proof works for functionals on linear spaces of bounded functions which are closed under taking  $f \wedge g$  and  $f \vee g$ . As an example, let  $\lambda(f) = \int_0^1 f(x) dx$  for all bounded measurable functions  $f: [0,1] \to \mathbb{R}$ . By the Hahn Banach Theorem ?? (or Corollary ??) below, we may extend  $\lambda$  to a linear functional  $\Lambda$  on all bounded functions on [0,1] in such a way that  $||\Lambda|| = 1$ . Let  $\Lambda_+$  be as above, then  $\Lambda_+ = \lambda$  on bounded measurable functions and  $||\Lambda_+|| = 1$ . Define  $\mu(\Lambda) := \Lambda(1_\Lambda)$  for all  $\Lambda \subset [0,1]$  and notice that if  $\Lambda$  is measurable, the  $\mu(\Lambda) = m(\Lambda)$ . So  $\mu$  is a finitely additive extension of m to **all** subsets of [0,1].

**Exercise 7.9.** Suppose that  $\mu$  is a signed Radon measure and  $I = I_{\mu}$ . Let  $\mu_+$  and  $\mu_-$  be the Radon measures associated to  $I_{\pm}$  with  $I_{\pm}$  being constructed as in the proof of Proposition 7.22. Show that  $\mu = \mu_+ - \mu_-$  is the Jordan decomposition of  $\mu$ .

**Theorem 7.24 (Dual of**  $C_0(X)$ ). Let X be a locally compact Hausdorff space, M(X) be the space of complex Radon measures on X and for  $\mu \in M(X)$  let  $\|\mu\| = |\mu|(X)$ . Then the map

$$\mu \in M(X) \to I_{\mu} \in C_0(X)^*$$

is an isometric isomorphism. Here again  $I_{\mu}(f) := \int_{X} f \ d\mu$ .

**Proof.** To show that the map  $M(X) \to C_0(X)^*$  is surjective, let  $I \in C_0(X)^*$  and then write  $I = I^{re} + iI^{im}$  be the decomposition into real and imaginary parts. Then further decompose these into there plus and minus parts so

$$I = I_{+}^{re} - I_{-}^{re} + i \left( I_{+}^{im} - I_{-}^{im} \right)$$

and let  $\mu_{\pm}^{re}$  and  $\mu_{\pm}^{im}$  be the corresponding positive Radon measures associated to  $I_{\pm}^{re}$  and  $I_{\pm}^{im}$ . Then  $I = I_{\mu}$  where

$$\mu = \mu_{+}^{re} - \mu_{-}^{re} + i \left( \mu_{+}^{im} - \mu_{-}^{im} \right).$$

To finish the proof it suffices to show  $||I_{\mu}||_{C_0(X)^*} = ||\mu|| = |\mu|(X)$ . We have

$$\begin{split} \|I_{\mu}\|_{C_0(X)^*} &= \sup\left\{\left|\int_X f d\mu\right| : f \in C_0(X) \ni \|f\|_{\infty} \le 1\right\} \\ &\leq \sup\left\{\left|\int_X f d\mu\right| : f \text{ measurable and } \|f\|_{\infty} \le 1\right\} = \|\mu\|. \end{split}$$

To prove the opposite inequality, write  $d\mu = gd |\mu|$  with g a complex measurable function such that |g| = 1. By Proposition ??, there exist  $f_n \in C_c(X)$  such that  $f_n \to g$  in  $L^1(|\mu|)$  as  $n \to \infty$ . Let  $g_n = \varphi(f_n)$  where  $\varphi : \mathbb{C} \to \mathbb{C}$  is the continuous function defined by  $\varphi(z) = z$  if  $|z| \leq 1$  and  $\varphi(z) = z/|z|$  if  $|z| \geq 1$ . Then  $|g_n| \leq 1$  and making use of the Lemma 2.14,  $g_n \to g$  in  $L^1(\mu)$  Thus

$$\|\mu\| = |\mu|(X) = \int_X d|\mu| = \int_X \bar{g} d\mu = \lim_{n \to \infty} \int_X \bar{g}_n d\mu \le \|I_\mu\|_{C_0(X)^*}.$$

# Homework #6 (Spring 2018)

For this last homework assignment:

Hand in Exercises 8.2 and 8.3.

Look at Exercise 8.1.

The next elementary theorem (referred to as the bounded linear transformation theorem, or B.L.T. theorem for short) is often useful when constructing bounded linear transformations.

**Theorem 8.1 (B. L. T. Theorem).** Suppose that Z is a normed space, X is a Banach space, and  $S \subset Z$  is a dense linear subspace of Z. If  $T : S \to X$  is a bounded linear transformation (i.e. there exists  $C < \infty$  such that  $||Tz|| \leq C ||z||$  for all  $z \in S$ ), then T has a unique extension to an element  $\overline{T} \in L(Z, X)$  and this extension still satisfies

$$\|\bar{T}z\| \leq C \|z\|$$
 for all  $z \in \bar{S}$ .

Exercise 8.1. Prove the B.L.T. Theorem 8.1.

**Exercise 8.2 (Dini's Theorem).** Let X be a compact topological space and  $f_n : X \to [0, \infty)$  be a sequence of continuous functions such that  $f_n(x) \downarrow 0$  as  $n \to \infty$  for each  $x \in X$ . Show that in fact  $f_n \downarrow 0$  uniformly in x, i.e.  $\sup_{x \in X} f_n(x) \downarrow 0$  as  $n \to \infty$ .<sup>1</sup>

**Hint:** Given  $\varepsilon > 0$ , consider the open sets  $V_n := \{x \in X : f_n(x) < \varepsilon\}$ .

**Theorem 8.2 (Riesz Markov Theorem for an Interval).** Let X = [0, 1]and  $\lambda \in C(X)^*$  be a positive linear functional. Then there exists a unique Borel measure,  $\mu$ , on  $\mathcal{B}_X$  such that  $\lambda(f) = \mu(f)$  for all  $f \in C(X, \mathbb{R})$  where

$$\mu\left(f\right) := \int_{X} f d\mu. \tag{8.1}$$

The following notations will be used in Exercise 8.3 below where you are asked prove the existence part of the Riesz Markov Theorem 8.2 on [0, 1].

Notation 8.3 For  $0 \le a \le b \le 1$ , let

$$\mathcal{V}\left([a,b]\right) := \min\left\{\mathcal{N}\left(f\right) : \mathbf{1}_{[a,b]} \leq f \in \mathcal{O}\left(\mathcal{N},\mathbb{R}\right)\right\}$$

 $u([a, b]) := \inf \{\lambda(f) : 1, \dots \leq f \in C(X \mathbb{R})\}$ 

and

$$F(b) := \nu([0,b]) := \inf \left\{ \lambda(f) : 1_{[0,b]} \le f \in C(X,\mathbb{R}) \right\}.$$
(8.3)

(8.2)

**Notation 8.4** For  $0 \leq \alpha < a < b < \beta \leq 1$ , let  $\chi_{\alpha,a,b,\beta} \in C([0,1],[0,1])$  be the piecewise linear function on [0,1] which is 0 on  $[0,\alpha]$ , linearly interpolates from 0 to 1 on  $[\alpha, a]$ , is 1 on [a, b], linearly interpolates from 1 to 0 on  $[b,\beta]$ , and is 0 again on  $[\beta, 1]$ . Also for  $0 \leq b \leq \beta \leq 1$ , let  $\theta_{b,\beta} \in C([0,1],[0,1])$  be the piecewise linear function on [0,1] which is 1 on [0,b], linearly interpolates from 1 to 0 on  $[b,\beta]$ , and is 0 again on  $[\beta,1]$ , see Figure 8.1.



**Fig. 8.1.** The graphs of smooth approximations to  $1_{[a,b]}$  and  $1_{[0,b]}$  as continuous functions on [0,1].

Exercise 8.3 (Riesz Markov Theorem for an Interval). Show there exists a finite Borel measure,  $\mu$ , on  $(X = [0, 1], \mathcal{B} = \mathcal{B}_X)$  satisfying Eq. (8.1) of Theorem 8.2. One way to prove this result is to prove the results listed below.

1. If  $0 \leq a \leq b \leq 1$ , show  $\nu([a,b]) = \lim_{n\to\infty} \lambda(\varphi_n)$  for any sequence,  $\{\varphi_n\}_{n=1}^{\infty} \subset C(X,[0,1])$ , such that  $\varphi_n(x) \downarrow 1_{[a,b]}(x)$  for all  $x \in [0,1]$ . Suggestions: given  $1_{[a,b]} \leq f \in C(X,[0,1])$  notice that

<sup>&</sup>lt;sup>1</sup> More generally, if  $g_n, g: X \to \mathbb{R}$  are continuous functions such that  $g_n(x) \downarrow g(x)$  as  $n \to \infty$  for each  $x \in X$ , then  $g_n(x) \to g(x)$  uniformly in x. Indeed, apply what you have proved to  $f_n := g_n - g$ .

#### 58 8 Homework #6 (Spring 2018)

a)  $\nu([a,b]) \leq \lambda(\varphi_n) \leq \lambda(\varphi_n \lor f)$  where  $\varphi_n \lor f := \max(\varphi_n, f)$ , and b)  $\varphi_n \lor f \downarrow f$  uniformly on [0,1] by Dini's theorem.

2. Show  $F(b) = \nu([0, b])$  is right continuous in b. **Hint:** if  $\{b_n\} \subset (0, 1]$  is a strictly decreasing sequence such that  $b_n \downarrow b$ , then

$$F(b) \le F(b_{+}) \le F(b_{n}) \le \lambda \left(\theta_{b_{n},b_{n-1}}\right). \tag{8.4}$$

Let  $\mu$  be the unique Borel measure on [0, 1] such that  $\mu([0, b]) = F(b)$  for all  $b \in [0, 1]$ . The goal is to show that this measure  $\mu$  satisfies Eq. (8.1). 3. Show  $\nu([a, b]) \leq \mu([a, b])$  for all  $0 \leq a \leq b \leq 1$ . **Hints**:

- a) if a = 0 there is nothing to prove so assume that  $0 < a \le b \le 1$ .
- b) Choose  $\{a_n\}_{n=1}^{\infty} \subset (0, a)$  so that  $a_n$  strictly increases to a as  $n \to \infty$  and let  $\varphi_n := \theta_{a_n, a}$  and  $\psi_n := \chi_{a_n, a, b, (b+\frac{1}{n}) \wedge 1}$  and observe that  $\varphi_n + \psi_n = \theta_{b, (b+\frac{1}{n}) \wedge 1}$  and hence

$$F(a_n) + \lambda(\psi_n) \le \lambda(\varphi_n) + \lambda(\psi_n) = \lambda\left(\theta_{b,(b+\frac{1}{n})\wedge 1}\right).$$
(8.5)

c) Pass to the limit as  $n \to \infty$  in the previous inequality.

4. Suppose that  $f \in C(X, [0, \infty))$  and  $\pi = \{0 = t_0 < t_1 < \dots < t_{n-1} < t_n = 1\}$  is a partition of [0, 1]. Let

$$c_i := \max \{ f(t) : t \in [t_{i-1}, t_i] \}$$
 for  $1 \le i \le n$ 

and set

$$f_{\pi} = c_1 \mathbf{1}_{[0,t_1]} + c_2 \mathbf{1}_{(t_1,t_2]} + \dots + c_n \mathbf{1}_{(t_{n-1},t_n]}.$$

Show

$$\lambda(f) \le \sum_{i=1}^{n} c_i \nu([t_{i-1}, t_i]) \le \sum_{i=1}^{n} c_i \mu([t_{i-1}, t_i]).$$
(8.6)

**Hint**: If  $f_i \in C(X, [0, 1])$  satisfy  $1_{[t_{i-1}, t_i]} \leq f_i$  for  $1 \leq i \leq n$ , then

$$f \le f_{\pi} \le \sum_{i=1}^{n} c_i f_i. \tag{8.7}$$

5. Recall that

$$\sum_{x\in [0,1]} \mu\left(\{x\}\right) \leq \mu\left([0,1]\right) = F\left(1\right) = \lambda\left(1\right) < \infty$$

and hence if  $E := \{x \in X : \mu(\{x\}) > 0\}$ , then E is at most countable. We now suppose that all partitions,  $\pi$ , we use have now been chosen so that  $t_j \notin E$  for 0 < j < n. Under this assumption, show Eq. (8.6) implies

$$\lambda\left(f\right) \leq \int_{X} f_{\pi} d\mu = \mu\left(f_{\pi}\right).$$

Since  $f_{\pi} \to f$  boundedly (in fact uniformly) as mesh $(\pi) := \max\{|t_i - t_{i-1}| : 1 \le i \le n\} \to 0$ , conclude that  $\lambda(f) \le \mu(f)$ .

6. Using  $\lambda(1) = \mu(1)$ , show  $\lambda(f) \le \mu(f)$  for all  $f \in C(X, \mathbb{R})$  then apply this result with f replaced by -f to complete the proof.

# Spectral Theorem (Compact Operator Case)

Before giving the general spectral theorem for bounded self-adjoint operators in the next chapter, we pause to consider the special case of "compact" operators. The theory in this setting looks very much like the finite dimensional matrix case.

# 9.1 Basics of Compact Operators

**Definition 9.1 (Compact Operator).** Let  $A : X \to Y$  be a bounded operator between two Banach spaces. Then A is **compact** if  $A[B_X(0,1)]$  is precompact in Y or equivalently for any  $\{x_n\}_{n=1}^{\infty} \subset X$  such that  $||x_n|| \leq 1$  for all n the sequence  $y_n := Ax_n \in Y$  has a convergent subsequence.

Remark 9.2. It is sometimes useful to note that A is compact iff A takes bounded sets to precompact sets. Indeed if  $B \subset X$  is a bounded set, then there exists  $R < \infty$  such that  $\left\|\frac{x}{R}\right\| \leq 1$  for all  $x \in B$  and hence if  $\{x_n\}_{n=1}^{\infty} \subset B$  we know there exists  $\{x_{n_k}\} \prec \{x_n\}$  such that  $\lim_{k\to\infty} A \frac{x_{n_k}}{R} = y$  and hence  $\lim_{k\to\infty} A x_{n_k} = Ry$  also exists.

**Definition 9.3.** A bounded operator  $A : X \to Y$  is said to have finite rank if  $\operatorname{Ran}(A) \subset Y$  is finite dimensional.

The following result is a simple consequence of the fact that closed bounded sets are compact in finite dimensional normed spaces.

**Corollary 9.4.** If  $A : X \to Y$  is a finite rank operator, then A is compact. In particular if either  $\dim(X) < \infty$  or  $\dim(Y) < \infty$  then any bounded operator  $A : X \to Y$  is finite rank and hence compact.

**Lemma 9.5.** If  $X \xrightarrow{A} Y \xrightarrow{B} Z$  are bounded operators between Banach spaces such the either A or B is compact then the composition  $BA : X \to Z$  is also compact. In particular if dim  $X = \infty$  and  $A \in L(X, Y)$  is an invertible operator such that  $A^{-1} \in L(Y, X)$ , then A is **not** compact. **Proof.** Let  $B_X(0,1)$  be the open unit ball in X. If A is compact and B is bounded, then  $BA(B_X(0,1)) \subset B(\overline{AB_X(0,1)})$  which is compact since the image of compact sets under continuous maps are compact. Hence we conclude that  $\overline{BA(B_X(0,1))}$  is compact, being the closed subset of the compact set  $B(\overline{AB_X(0,1)})$ . If A is continuous and B is compact, then  $A(B_X(0,1))$  is a bounded set and so by the compactness of B,  $BA(B_X(0,1))$  is a precompact subset of Z, i.e. BA is compact.

Alternatively: Suppose that  $\{x_n\}_{n=1}^{\infty} \subset X$  is a bounded sequence. If A is compact, then  $y_n := Ax_n$  has a convergent subsequence,  $\{y_{n_k}\}_{k=1}^{\infty}$ . Since B is continuous it follows that  $z_{n_k} := By_{n_k} = BAx_{n_k}$  is a convergent subsequence of  $\{BAx_n\}_{n=1}^{\infty}$ . Similarly if A is bounded and B is compact then  $y_n = Ax_n$  defines a bounded sequence inside of Y. By compactness of B, there is a subsequence  $\{y_{n_k}\}_{k=1}^{\infty}$  for which  $\{BAx_{n_k} = By_{n_k}\}_{k=1}^{\infty}$  is convergent in Z.

For the second statement, if A were compact then  $I_X := A^{-1}A$  would be compact as well. As  $I_X$  takes the unit ball to the unit ball, the identity is compact iff dim  $X < \infty$ .

**Corollary 9.6.** Let X be a Banach space and  $\mathcal{K}(X) := \mathcal{K}(X, X)$ . Then  $\mathcal{K}(X)$  is a norm-closed ideal of L(X) which contains  $I_X$  iff dim  $X < \infty$ .

In order to give some more interesting examples of compact operators, let us recall that Ascoli-Arzela theorem for which we recall the following definition.

**Definition 9.7.** Let X be a topological space and  $\mathbb{F}$  be either  $\mathbb{R}$  or  $\mathbb{C}$  and suppose that  $\mathcal{F} \subset C(X, Y)$ .

- 1.  $\mathcal{F}$  is equicontinuous at  $x \in X$  iff  $\lim_{\xi \to x} \sup_{f \in \mathcal{F}} |f(\xi) f(x)| = 0.^2$
- 2.  $\mathcal{F}$  is equicontinuous if  $\mathcal{F}$  is equicontinuous at all points  $x \in X$ .
- 3.  $\mathcal{F}$  is **pointwise bounded** if  $\sup\{|f(x)|: f \in \mathcal{F}\} < \infty$  for all  $x \in X$ .

**Theorem 9.8 (Ascoli-Arzela Theorem).** Let  $(X, \tau)$  be a compact topological space and  $\mathcal{F} \subset C(X, \mathbb{F}^d)$ . Then  $\mathcal{F}$  is precompact in  $C(X, \mathbb{F}^d)$  iff  $\mathcal{F}$  is equicontinuous and point-wise bounded.

<sup>2</sup> This should be compared with  $f : X \to Y$  being continuous at x iff  $\lim_{\xi \to x} |f(\xi) - f(x)| = 0.$ 

<sup>&</sup>lt;sup>1</sup> Later we will see that A being one to one and onto automatically implies that  $A^{-1}$  is bounded by the open mapping Theorem ??.

**Proposition 9.9.** Let X be a compact topological space (metric space is fine), Z be a Banach space (normed space), and  $X \ni x \to k_x \in Z^*$  be a continuous map. Then  $K: Z \to C(X)$  defined by

$$(Kf)(x) = \langle f, k_x \rangle = \hat{f}(k_x)$$

is a compact operator where  $\langle f, k_x \rangle := k_x (f) =: \hat{f}(k_x)$ . [Note that  $||K||_{op} \leq \max_x ||k_x||_{Z^*} < \infty$  since  $X \ni x \to ||k_x||_{Z^*} \in \mathbb{R}$  is a continuous function on a compact set.]

**Proof.** The map  $Z \ni f \to \hat{f} \in Z^{**}$  is continuous and therefore  $Kf : X \to \mathbb{C}$  is continuous being the composition of two continuous maps,  $X \xrightarrow{k_{(\bullet)}} Z^* \xrightarrow{\hat{f}} \mathbb{C}$ . Let

$$\mathcal{F} := KB_Z(0,1) = \{Kf : f \in Z \text{ with } \|f\|_Z \le 1\}.$$

As

$$\sup_{\|f\| \le 1} |Kf(x)| = \sup_{\|f\| \le 1} |k_x(f)| = \|k_x\|_{Z^*} < \infty$$

it follows that  $\mathcal{F}$  is pointwise bounded. Also if  $\tilde{x}, x \in X$ ,

$$\sup_{\|f\| \le 1} |Kf(\tilde{x}) - Kf(x)| = \sup_{\|f\| \le 1} |\langle f, k_{\tilde{x}} - k_x \rangle| = \|k_{\tilde{x}} - k_x\|_{Z^*} \to 0 \text{ as } \tilde{x} \to x$$

since  $x \to k_x$  is continuous. This shows that  $\mathcal{F}$  is equicontinuous and hence  $\mathcal{F} = KB_Z(0,1)$  is precompact in  $C(X,\mathbb{C})$ .

**Corollary 9.10 (Integral operators).** Suppose that X be a compact metric space,  $(\Omega, \mathcal{F}, \mu)$  is a measure space,  $1 , <math>0 \leq g \in L^{p^*}(\mu)$ , and  $k : X \times \Omega \to \mathbb{C}$  is a jointly measurable function such that  $X \ni x \to k(x, \omega) \in \mathbb{C}$  is continuous, and  $|k(x, \omega)| \leq g(\omega)$  for all (or  $\mu$ -a.e.)  $\omega \in \Omega$ . For  $f \in Z := L^p(\mu)$  and  $x \in X$ , let

$$(Kf)(x) := \int_{\Omega} k(x,\omega) f(\omega) d\mu(\omega),$$

then  $K: L^{p}(\mu) \rightarrow C(X)$  is a compact operator.

**Proof.** To prove this let  $k_x := k(x, \cdot) \in L^{p^*}(\mu) \cong L^p(\mu)^*$  and observe that  $(Kf)(x) = \langle f, k_x \rangle$  and

$$\lim_{\tilde{x} \to x} \|k_{\tilde{x}} - k_{x}\|_{p^{*}}^{p^{*}} = \lim_{\tilde{x} \to x} \int_{\Omega} |k(\tilde{x}, \omega) - k(x, \omega)|^{p^{*}} d\mu(\omega)$$
$$= \int_{\Omega} \lim_{\tilde{x} \to x} |k(\tilde{x}, \omega) - k(x, \omega)|^{p^{*}} d\mu(\omega) = 0$$

wherein we have used the dominated convergence theorem with dominating function being  $2^{p^*}g^p \in L^1(\mu)$ . Technically, we take the limits along arbitrary sequences,  $\tilde{x} = x_n \to x$  as  $n \to \infty$  in order to apply DCT.

*Example 9.11 (Integral operators).* Suppose that X is a compact metric space,  $\mathcal{B}$ , is the Borel  $\sigma$ -algebra on X,  $\mu$  is a finite measure on  $(X, \mathcal{B})$ , and  $k: X \times X \to \mathbb{C}$  is a jointly continuous function. Then for any  $1 and <math>f \in L^p(\mu)$ , let

$$Kf(x) := \int_{X} k(x, y) f(y) d\mu(y).$$

Then  $K: L^{p}(X, \mu) \to C(X)$  is a compact operator and since  $C(X) \ni f \to f \in L^{p}(X, \mu)$  is a continuos injection we may further conclude that  $K: C(X) \to C(X)$  is a compact operator.

**Theorem 9.12.** Let X and Y be Banach spaces and  $\mathcal{K} := \mathcal{K}(X,Y)$  denote the compact operators from X to Y. Then  $\mathcal{K}(X,Y)$  is a norm-closed subspace of B(X,Y). In particular, operator norm limits of finite rank operators are compact.

**Proof.** Using the sequential definition of compactness it is easily seen that  $\mathcal{K}$  is a vector subspace of B(X,Y). To finish the proof, we must show that  $K \in B(X,Y)$  is compact if there exists  $K_n \in \mathcal{K}(X,Y)$  such that  $\lim_{n\to\infty} ||K_n - K||_{op} = 0$ .

**First Proof.** Let  $U := B_0(1)$  be the unit ball in X. Given  $\varepsilon > 0$ , choose  $N = N(\varepsilon)$  such that  $||K_N - K|| \le \varepsilon$ . Using the fact that  $K_N U$  is precompact, choose a finite subset  $\Lambda \subset U$  such that  $K_N U \subset \bigcup_{\sigma \in \Lambda} B_{K_N \sigma}(\varepsilon)$ . Then given  $y = Kx \in KU$  we have  $K_N x \in B_{K_N \sigma}(\varepsilon)$  for some  $\sigma \in \Lambda$  and for this  $\sigma$ ;

$$\begin{aligned} \|y - K_N \sigma\| &= \|Kx - K_N \sigma\| \\ &\leq \|Kx - K_N x\| + \|K_N x - K_N \sigma\| < \varepsilon \, \|x\| + \varepsilon < 2\varepsilon. \end{aligned}$$

This shows  $KU \subset \bigcup_{\sigma \in \Lambda} B_{K_N \sigma}(2\varepsilon)$  and therefore is KU is  $2\varepsilon$  – bounded for all  $\varepsilon > 0$ , i.e. KU is totally bounded and hence precompact.

**Second Proof.** Suppose  $\{x_n\}_{n=1}^{\infty}$  is a bounded sequence in X. By compactness, there is a subsequence  $\{x_n^1\}_{n=1}^{\infty}$  of  $\{x_n\}_{n=1}^{\infty}$  such that  $\{K_1x_n^1\}_{n=1}^{\infty}$  is convergent in Y. Working inductively, we may construct subsequences

$$\{x_n\}_{n=1}^{\infty} \supset \{x_n^1\}_{n=1}^{\infty} \supset \{x_n^2\}_{n=1}^{\infty} \cdots \supset \{x_n^m\}_{n=1}^{\infty} \supset \dots$$

such that  $\{K_m x_n^m\}_{n=1}^{\infty}$  is convergent in Y for each m. By the usual Cantor's diagonalization procedure, let  $\xi_n := x_n^n$ , then  $\{\xi_n\}_{n=1}^{\infty}$  is a subsequence of  $\{x_n\}_{n=1}^{\infty}$  such that  $\{K_m \xi_n\}_{n=1}^{\infty}$  is convergent for all m. Since

$$\|K\xi_n - K\xi_l\| \le \|(K - K_m)\xi_n\| + \|K_m(\xi_n - \xi_l)\| + \|(K_m - K)\xi_l\|$$
  
$$\le 2 \|K - K_m\| + \|K_m(\xi_n - \xi_l)\|,$$
  
$$\lim \sup_{n,l \to \infty} \|K\xi_n - K\xi_l\| \le 2 \|K - K_m\| \to 0 \text{ as } m \to \infty,$$

which shows  $\{K\xi_n\}_{n=1}^{\infty}$  is Cauchy and hence convergent.

*Example 9.13.* Let  $X = \ell^2 = Y$  and  $\lambda_n \in \mathbb{C}$  such that  $\lim_{n \to \infty} \lambda_n = 0$ , then  $A: X \to Y$  defined by  $(Ax)(n) = \lambda_n x(n)$  is compact. To verify this claim, for each  $m \in \mathbb{N}$  let  $(A_m x)(n) = \lambda_n x(n) \mathbb{1}_{n \leq m}$ . In matrix language,

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \text{ and } A_m = \begin{pmatrix} \lambda_1 & 0 & \cdots & & \\ 0 & \lambda_2 & 0 & \cdots & \\ \vdots & \ddots & \ddots & \ddots & \\ & 0 & \lambda_m & 0 & \cdots \\ & & & 0 & 0 & \ddots \\ & & & & \ddots & \ddots \end{pmatrix}.$$

Then  $A_m$  is finite rank and  $||A - A_m||_{op} = \max_{n>m} |\lambda_n| \to 0$  as  $m \to \infty$ . The claim now follows from Theorem 9.12.

*Example 9.14.* Let  $(X, \mathcal{B}, \mu)$  be a  $\sigma$  - finite measure spaces whose  $\sigma$  - algebra is countably generated by sets of finite measure. If  $k \in L^2(X \times X, \mu \otimes \mu)$ , then  $K : L^2(\mu) \to L^2(\mu)$  defined by

$$Kf(x) := \int_{X} k(x, y) f(y) d\mu(y)$$

is a compact operator.

**Proof.** First observe that

$$|Kf(x)|^{2} \le ||f||^{2} \int_{X} |k(x,y)|^{2} d\mu(y)$$

and hence

$$|Kf||^{2} \le ||f||^{2} \int_{X \times X} |k(x, y)|^{2} d\mu(x) d\mu(y)$$

from which it follows that  $||K||_{op} \leq ||k||_{L^2(\mu \otimes \mu)}$ .

Now let  $\{\psi_n\}_{n=1}^{\infty}$  be an orthonormal basis for  $L^2(X,\mu)$  and let

$$k_{N}(x,y) := \sum_{m,n=1}^{N} \langle k, \psi_{m} \otimes \psi_{n} \rangle \psi_{m} \otimes \psi_{n}$$

where  $f \otimes g(x, y) := f(x) g(y)$ . Then

$$K_N f(x) := \int_X k_N(x, y) f(y) d\mu(y) = \sum_{m, n=1}^N \langle k, \psi_m \otimes \psi_n \rangle \langle f, \bar{\psi}_n \rangle \psi_m$$

is a finite rank and hence compact operator. Since

 $||K - K_N||_{op} \le ||k - k_N||_{L^2(\mu \otimes \mu)} \to 0 \text{ as } N \to \infty$ 

it follows that K is compact as well.

We will see more examples of compact operators below in Section 9.4 and Exercise **??** below.

#### 9.2 Compact Operators on Hilbert spaces

(This section is not absolutely necessary as the results may be deduced from results from the following Spectral Theorem Section 9.3.)

**Lemma 9.15.** Suppose that  $T, T_n \in L(X, Y)$  for  $n \in \mathbb{N}$  where X and Y are normed spaces. If  $T_n \xrightarrow{s} T$ ,  $M = \sup_n ||T_n|| < \infty$ ,<sup>3</sup> and  $x_n \to x$  in X as  $n \to \infty$ , then  $T_n x_n \to Tx$  in Y as  $n \to \infty$ . Moreover if  $K \subset X$  is a compact set then

$$\lim_{n \to \infty} \sup_{x \in K} \|Tx - T_n x\| = 0.$$

$$(9.1)$$

Proof. 1. We have,

$$\begin{aligned} \|Tx - T_n x_n\| &\leq \|Tx - T_n x\| + \|T_n x - T_n x_n\| \\ &\leq \|Tx - T_n x\| + M \|x - x_n\| \to 0 \text{ as } n \to \infty. \end{aligned}$$

2. For sake of contradiction, suppose that

$$\limsup_{n \to \infty} \sup_{x \in K} \|Tx - T_n x\| = \varepsilon > 0.$$

In this case we can find  $\{n_k\}_{k=1}^{\infty} \subset \mathbb{N}$  and  $x_{n_k} \in K$  such that  $||Tx_{n_k} - T_{n_k}x_{n_k}|| \geq \varepsilon/2$ . Since K is compact, by passing to a subsequence if necessary, we may assume  $\lim_{k\to\infty} x_{n_k} = x$  exists in K. On the other hand by part **1**. we know that

$$\lim_{k \to \infty} \|Tx_{n_k} - T_{n_k}x_{n_k}\| = \left\|\lim_{k \to \infty} Tx_{n_k} - \lim_{k \to \infty} T_{n_k}x_{n_k}\right\| = \|Tx - Tx\| = 0.$$

**2 alternate proof.** Given  $\varepsilon > 0$ , there exists  $\{x_1, \ldots, x_N\} \subset K$  such that  $K \subset \bigcup_{l=1}^N B_{x_l}(\varepsilon)$ . If  $x \in K$ , choose l such that  $x \in B_{x_l}(\varepsilon)$  in which case,

$$||Tx - T_n x|| \le ||Tx - Tx_l|| + ||Tx_l - T_n x_l|| + ||T_n x_l - T_n x||$$
  
$$\le (||T||_{op} + M) \varepsilon + ||Tx_l - T_n x_l||$$

<sup>&</sup>lt;sup>3</sup> If X and Y are Banach spaces, the uniform boundedness principle shows that  $T_n \xrightarrow{s} T$  automatically implies  $\sup_n ||T_n|| < \infty$ .

#### 62 9 Spectral Theorem (Compact Operator Case)

and therefore it follows that

$$\sup_{x \in K} \|Tx - T_n x\| \le \left( \|T\|_{op} + M \right) \varepsilon + \max_{1 \le l \le N} \|Tx_l - T_n x_l\|$$

and therefore,

$$\limsup_{n \to \infty} \sup_{x \in K} \|Tx - T_n x\| \le \left( \|T\|_{op} + M \right) \varepsilon.$$

As  $\varepsilon > 0$  was arbitrary we conclude that Eq. (9.1) holds.

For the rest of this section, let H and B be Hilbert spaces and  $U := \{x \in H : ||x|| < 1\}$  be the **open unit ball** in H.

**Proposition 9.16.** A bounded operator  $K : H \to B$  is compact iff there exists finite rank operators,  $K_n : H \to B$ , such that  $||K - K_n|| \to 0$  as  $n \to \infty$ .

**Proof.** Suppose that  $K : H \to B$ . Since  $\overline{K(U)}$  is compact it contains a countable dense subset and from this it follows that  $\overline{K(H)}$  is a separable subspace of B. Let  $\{\varphi_{\ell}\}$  be an orthonormal basis for  $\overline{K(H)} \subset B$  and

$$P_n y = \sum_{\ell=1}^n \langle y, \varphi_\ell \rangle \varphi_\ell$$

be the orthogonal projection of y onto  $\operatorname{span}\{\varphi_\ell\}_{\ell=1}^n$ . Then  $\lim_{n\to\infty} ||P_n y - y|| = 0$  for all  $y \in \overline{K(H)}$ . Define  $K_n := P_n K$  – a finite rank operator on H. It then follows that

$$\begin{split} \limsup_{n \to \infty} \|K - K_n\| &= \limsup_{n \to \infty} \sup_{x \in U} \|Kx - K_n x\| \\ &= \limsup_{n \to \infty} \sup_{x \in U} \|(I - P_n) K x\| \\ &\leq \limsup_{n \to \infty} \sup_{y \in \overline{K(U)}} \|(I - P_n) y\| = 0 \end{split}$$

by Lemma 9.15 along with the facts that  $\overline{K(U)}$  is compact and  $P_n \xrightarrow{s} I$ . The converse direction follows from Corollary 9.4 and Theorem 9.12.

**Corollary 9.17.** If K is compact then so is  $K^*$ .

**Proof. First Proof.** Let  $K_n = P_n K$  be as in the proof of Proposition 9.16, then  $K_n^* = K^* P_n$  is still finite rank. Furthermore, using Proposition ??,

$$||K^* - K_n^*|| = ||K - K_n|| \to 0 \text{ as } n \to \infty$$

showing  $K^*$  is a limit of finite rank operators and hence compact.

**Second Proof.** Let  $\{x_n\}_{n=1}^{\infty}$  be a bounded sequence in *B*, then

$$\|K^*x_n - K^*x_m\|^2 = \langle x_n - x_m, KK^*(x_n - x_m) \rangle \le 2C \|KK^*(x_n - x_m)\|$$
(9.2)

where C is a bound on the norms of the  $x_n$ . Since  $\{K^*x_n\}_{n=1}^{\infty}$  is also a bounded sequence, by the compactness of K there is a subsequence  $\{x'_n\}$  of the  $\{x_n\}$  such that  $KK^*x'_n$  is convergent and hence by Eq. (9.2), so is the sequence  $\{K^*x'_n\}$ .

# 9.3 The Spectral Theorem for Self Adjoint Compact Operators

For the rest of this section,  $K \in \mathcal{K}(H) := \mathcal{K}(H, H)$  will be a self-adjoint compact operator or **S.A.C.O.** for short. Because of Proposition 9.16, we might expect compact operators to behave very much like finite dimensional matrices. This is typically the case as we will see below.

Example 9.18 (Model S.A.C.O.). Let  $H = \ell_2$  and K be the diagonal matrix

$$K = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix},$$

where  $\lim_{n\to\infty} |\lambda_n| = 0$  and  $\lambda_n \in \mathbb{R}$ . Then K is a self-adjoint compact operator. This assertion was proved in Example 9.13.

The main theorem (Theorem 9.24) of this subsection states that up to unitary equivalence, Example 9.18 is essentially the most general example of an S.A.C.O. Before stating and proving this theorem we will require the following results.

**Lemma 9.19.** Let  $Q : H \times H \to \mathbb{C}$  be a symmetric sesquilinear form on H where Q is symmetric means  $\overline{Q(h,k)} = Q(k,h)$  for all  $h,k \in H$ . Letting Q(h) := Q(h,h), then for all  $h,k \in H$ ,

$$Q(h+k) = Q(h) + Q(k) + 2 \operatorname{Re} Q(h,k), \qquad (9.3)$$

$$Q(h+k) + Q(h-k) = 2Q(h) + 2Q(k)$$
, and (9.4)

$$Q(h+k) - Q(h-k) = 4 \operatorname{Re} Q(h,k).$$
(9.5)

**Proof.** The simple proof is left as an exercise to the reader.

**Theorem 9.20 (Rayleigh quotient).** Suppose  $T \in B(H)$  is a bounded selfadjoint operator, then

$$M := \sup_{f \neq 0} \frac{|\langle Tf, f \rangle|}{\|f\|^2} = \|T\| \left( = \sup_{f \neq 0} \frac{\|Tf\|}{\|f\|} \right).$$

Moreover, if there exists a non-zero element  $f \in H \setminus \{0\}$  such that  $|\langle Tf, f \rangle| / ||f||^2 = ||T||$ , then f is an eigenvector of T with  $Tf = \lambda f$  and  $\lambda \in \{\pm ||T||\}$ .

**Proof. First proof.** Applying Eq. (9.5) with  $Q(f,g) = \langle Tf,g \rangle$  and Eq. (9.4) with  $Q(f,g) = \langle f,g \rangle$  along with the Cauchy-Schwarz inequality implies,

$$4\operatorname{Re} \langle Tf,g \rangle = \langle T(f+g), (f+g) \rangle - \langle T(f-g), (f-g) \rangle$$
$$\leq M \left[ \|f+g\|^2 + \|f-g\|^2 \right] = 2M \left[ \|f\|^2 + \|g\|^2 \right].$$

Replacing f by  $e^{i\theta}f$  where  $\theta$  is chosen so that  $e^{i\theta}\langle Tf,g\rangle = |\langle Tf,g\rangle|$  then shows

$$4 |\langle Tf, g \rangle| \le 2M \left[ ||f||^2 + ||g||^2 \right]$$

and therefore,

$$||T|| = \sup_{||f|| = ||g|| = 1} |\langle f, Tg \rangle| \le M$$

and since it is clear  $M \leq ||T||$  we have shown M = ||T||.

If  $f \in H \setminus \{0\}$  and  $||T|| = |\langle Tf, f \rangle| / ||f||^2$  then, using Schwarz's inequality,

$$||T|| = \frac{|\langle Tf, f \rangle|}{||f||^2} \le \frac{||Tf||}{||f||} \le ||T||.$$
(9.6)

This implies  $|\langle Tf, f \rangle| = ||Tf|| ||f||$  and forces equality in Schwarz's inequality. So by Theorem ??, Tf and f are linearly dependent, i.e.  $Tf = \lambda f$  for some  $\lambda \in \mathbb{C}$ . Substituting this into (9.6) shows that  $|\lambda| = ||T||$ . Since T is self-adjoint,

$$\lambda \left\| f \right\|^{2} = \left\langle \lambda f, f \right\rangle = \left\langle T f, f \right\rangle = \left\langle f, T f \right\rangle = \left\langle f, \lambda f \right\rangle = \bar{\lambda} \left\langle f, f \right\rangle = \bar{\lambda} \left\| f \right\|^{2},$$

which implies that  $\lambda \in \mathbb{R}$  and therefore,  $\lambda \in \{\pm \|T\|\}$ .

**Exercise 9.1 (This may be skipped).** Suppose that  $A : H \to H$  is a bounded self-adjoint operator on H. Show;

1.  $f(x) := \langle Ax, x \rangle \in \mathbb{R}$  for all  $x \in H$ .

2. If there exists  $x_0 \in H$  with  $||x_0|| = 1$  such that

$$\lambda_0 := \sup_{\|x\|=1} \langle Ax, x \rangle = \langle Ax_0, x_0 \rangle$$

then  $Ax_0 = \lambda_0 x_0$ . **Hint:** Given  $y \in H$  let  $c(t) := \frac{x_0 + ty}{\|x_0 + ty\|_H}$  for t near 0. Then apply the first derivative test to the function  $g(t) = \langle Ac(t), c(t) \rangle$ .

3. If we further assume that A is compact, then A has at least one eigenvector.

**Proposition 9.21.** Let K be a S.A.C.O., then either  $\lambda = ||K||$  or  $\lambda = -||K||$  is an eigenvalue of K.

**Proof.** (For those who have done Exercise 9.1, that exercise along with Theorem 9.20 constitutes a proof.) Without loss of generality we may assume that K is non-zero since otherwise the result is trivial. By Theorem 9.20, there exists  $u_n \in H$  such that  $||u_n|| = 1$  and

$$\frac{|\langle u_n, Ku_n \rangle|}{\|u_n\|^2} = |\langle u_n, Ku_n \rangle| \longrightarrow \|K\| \text{ as } n \to \infty.$$
(9.7)

By passing to a subsequence if necessary, we may assume that  $\lambda := \lim_{n\to\infty} \langle u_n, Ku_n \rangle$  exists and  $\lambda \in \{\pm \|K\|\}$ . By passing to a further subsequence if necessary, we may assume, using the compactness of K, that  $Ku_n$  is convergent as well. We now compute:

$$0 \le ||Ku_n - \lambda u_n||^2 = ||Ku_n||^2 - 2\lambda \langle Ku_n, u_n \rangle + \lambda^2$$
  
$$\le \lambda^2 - 2\lambda \langle Ku_n, u_n \rangle + \lambda^2$$
  
$$\to \lambda^2 - 2\lambda^2 + \lambda^2 = 0 \text{ as } n \to \infty.$$

Hence

$$Ku_n - \lambda u_n \to 0 \text{ as } n \to \infty$$
 (9.8)

and therefore

$$u := \lim_{n \to \infty} u_n = \frac{1}{\lambda} \lim_{n \to \infty} K u_n$$

exists. By the continuity of the inner product,  $||u|| = 1 \neq 0$ . By passing to the limit in Eq. (9.8) we find that  $Ku = \lambda u$ .

**Lemma 9.22.** If H and K be Hilbert spaces and  $A \in L(H, K)$ , then;

- 1.  $\underline{\operatorname{Nul}(A^*)} = \operatorname{Ran}(A)^{\perp}$ , and 2.  $\overline{\operatorname{Ran}(A)} = \operatorname{Nul}(A^*)^{\perp}$ , 3. If we further assume that K = H
- 3. If we further assume that K = H, and  $V \subset H$  is an A invariant subspace (i.e.  $A(V) \subset V$ ), then  $V^{\perp}$  is  $A^*$  invariant.

#### 64 9 Spectral Theorem (Compact Operator Case)

**Proof.** 1. We have  $y \in Nul(A^*) \iff A^*y = 0 \iff \langle y, Ah \rangle = \langle 0, h \rangle = 0$ for all  $h \in H \iff y \in \operatorname{Ran}(A)^{\perp}$ .

2. By Exercise ??,  $\overline{\operatorname{Ran}(A)} = \operatorname{Ran}(A)^{\perp \perp}$ , and so  $\overline{\operatorname{Ran}(A)} = \operatorname{Ran}(A)^{\perp \perp} =$  $\operatorname{Nul}(A^*)^{\perp}$ .

3. Now suppose that K = H and  $AV \subset V$ . If  $y \in V^{\perp}$  and  $x \in V$ , then

$$\langle A^*y, x \rangle = \langle y, Ax \rangle = 0$$
 for all  $x \in V \implies A^*y \in V^{\perp}$ .

**Definition 9.23 (Spectrum of an operator).** If X is a Banach space and  $A: X \to X$  is a bounded operator we define  $\lambda \in \sigma(A)$  iff  $(A - \lambda I)$  is not invertible. The subset,  $\sigma(A) \subset \mathbb{F}$ , is referred to as the spectrum of A.

Theorem 9.24 (Compact Operator Spectral Theorem). Suppose that  $K: H \to H$  is a non-zero S.A.C.O., then

- 1. there exists at least one eigenvalue  $\lambda \in \{\pm \|K\|\}$ .
- 2. There are at most countably many **non-zero** eigenvalues,  $\{\lambda_n\}_{n=1}^N$ , where  $N = \infty$  is allowed. (Unless K is finite rank (i.e. dim Ran  $(K) < \infty$ ), N will be infinite.)
- 3. The  $\lambda_n$ 's (including multiplicities) may be arranged so that  $|\lambda_n| \geq |\lambda_{n+1}|$ for all n. If  $N = \infty$  then  $\lim_{n \to \infty} |\lambda_n| = 0$ . (In particular any eigenspace for K with **non-zero** eigenvalue is finite dimensional.)
- 4. The eigenvectors  $\{\varphi_n\}_{n=1}^N$  can be chosen to be an O.N. set such that H = $\overline{span\{\varphi_n\}} \stackrel{\perp}{\oplus} \operatorname{Nul}(K).$

5. Using the  $\{\varphi_n\}_{n=1}^{N}$  above,

$$Kf = \sum_{n=1}^{N} \lambda_n \langle f, \varphi_n \rangle \varphi_n \text{ for all } f \in H.$$
(9.9)

6. The spectrum of K is  $\sigma(K) = \{0\} \cup \{\lambda_n : n < N+1\}$  if dim  $H = \infty$ , otherwise  $\sigma(K) = \{\lambda_n : n \leq N\}$  with  $N \leq \dim H$ .

**Proof.** We will find  $\lambda_n$ 's and  $\varphi_n$ 's recursively. Let  $\lambda_1 \in \{\pm \|K\|\}$  and  $\varphi_1 \in H$ such that  $K\varphi_1 = \lambda_1 \varphi_1$  as in Proposition 9.21.

Take  $M_1 = \operatorname{span}(\varphi_1)$  so  $K(M_1) \subset M_1$ . By Lemma 9.22,  $KM_1^{\perp} \subset M_1^{\perp}$ . Define  $K_1: M_1^{\perp} \to M_1^{\perp}$  via  $K_1 = K|_{M_1^{\perp}}$ . Then  $K_1$  is again a compact operator. If  $K_1 = 0$ , we are done. If  $K_1 \neq 0$ , by Proposition 9.21 there exists  $\lambda_2 \in$  $\{\pm \|K_1\|\}$  and  $\varphi_2 \in M_1^{\perp}$  such that  $\|\varphi_2\| = 1$  and  $K_1\varphi_2 = K\varphi_2 = \lambda_2\varphi_2$ . Let  $M_2 := \operatorname{span}(\varphi_1, \varphi_2).$ 

Again  $K(M_2) \subset M_2$  and hence  $K_2 := K|_{M_2^{\perp}} : M_2^{\perp} \to M_2^{\perp}$  is compact and if  $K_2 = 0$  we are done. When  $K_2 \neq 0$ , we apply Proposition 9.21 again to find  $\lambda_3 \in \{\pm \|K\|_2\}$  and  $\varphi_3 \in M_2^{\perp}$  such that  $\|\varphi_3\| = 1$  and  $K_2\varphi_3 = K\varphi_3 = \lambda_3\varphi_3$ .

Continuing this way indefinitely or until we reach a point where  $K_n = 0$ , we construct a sequence  $\{\lambda_n\}_{n=1}^N$  of eigenvalues and orthonormal eigenvectors  $\{\varphi_n\}_{n=1}^N$  such that  $|\lambda_n| \ge |\lambda_{n+1}|$  with the further property that

$$|\lambda_n| = \sup_{\varphi \perp \{\varphi_1, \varphi_2, \dots, \varphi_{n-1}\}} \frac{\|K\varphi\|}{\|\varphi\|}.$$
(9.10)

When  $N < \infty$ , the remaining results in the theorem are easily verified. So from now on let us assume that  $N = \infty$ .

If  $\varepsilon := \lim_{n \to \infty} |\lambda_n| > 0$ , then  $\{\lambda_n^{-1} \varphi_n\}_{n=1}^{\infty}$  is a bounded sequence in *H*. Hence, by the compactness of K, there exists a subsequence  $\{n_k : k \in \mathbb{N}\}$  of  $\mathbb{N}$  such that  $\{\varphi_{n_k} = \lambda_{n_k}^{-1} K \varphi_{n_k}\}_{k=1}^{\infty}$  is a convergent. However, since  $\{\varphi_{n_k}\}_{k=1}^{\infty}$ is an orthonormal set, this is impossible and hence we must conclude that  $\varepsilon := \lim_{n \to \infty} |\lambda_n| = 0.$ 

Let  $M := \operatorname{span}\{\varphi_n\}_{n=1}^{\infty}$ . Then  $K(M) \subset M$  and hence, by Lemma 9.22,  $K(M^{\perp}) \subset M^{\perp}$ . Using Eq. (9.10),

$$||K|_{M^{\perp}}|| \le ||K|_{M_n^{\perp}}|| = |\lambda_n| \longrightarrow 0 \text{ as } n \to \infty$$

showing  $K|M^{\perp} \equiv 0$ . Define  $P_0$  to be orthogonal projection onto  $M^{\perp}$ . Then for  $f \in H$ ,

$$f = P_0 f + (1 - P_0) f = P_0 f + \sum_{n=1}^{\infty} \langle f, \varphi_n \rangle \varphi_n$$

and

$$Kf = KP_0f + K\sum_{n=1}^{\infty} \langle f, \varphi_n \rangle \varphi_n = \sum_{n=1}^{\infty} \lambda_n \langle f, \varphi_n \rangle \varphi_n$$

which proves Eq. (9.9).

Since  $\{\lambda_n\}_{n=1}^{\infty} \subset \sigma(K)$  and  $\sigma(K)$  is closed, it follows that  $0 \in \sigma(K)$  and hence  $\{\lambda_n\}_{n=1}^{\infty} \cup \{0\} \subset \sigma(K)$ . Suppose that  $z \notin \{\lambda_n\}_{n=1}^{\infty} \cup \{0\}$  and let d be the distance between z and  $\{\lambda_n\}_{n=1}^{\infty} \cup \{0\}$ . Notice that d > 0 because  $\lim_{n \to \infty} \lambda_n = 0.$ 

A few simple computations show that:

$$(K-zI)f = \sum_{n=1}^{\infty} \langle f, \varphi_n \rangle (\lambda_n - z)\varphi_n - zP_0f,$$

 $(K-z)^{-1}$  exists,

$$(K-zI)^{-1}f = \sum_{n=1}^{\infty} \langle f, \varphi_n \rangle (\lambda_n - z)^{-1} \varphi_n - z^{-1} P_0 f,$$

and

macro: svmonob.cls

$$\begin{split} |(K-zI)^{-1}f||^2 &= \sum_{n=1}^{\infty} |\langle f,\varphi_n\rangle|^2 \frac{1}{|\lambda_n-z|^2} + \frac{1}{|z|^2} ||P_0f||^2 \\ &\leq \left(\frac{1}{d}\right)^2 \left(\sum_{n=1}^{\infty} |\langle f,\varphi_n\rangle|^2 + ||P_0f||^2\right) = \frac{1}{d^2} ||f||^2 \,. \end{split}$$

We have thus shown that  $(K - zI)^{-1}$  exists,  $||(K - zI)^{-1}|| \le d^{-1} < \infty$  and hence  $z \notin \sigma(K)$ .

**Theorem 9.25 (Structure of Compact Operators).** Let  $K : H \to B$ be a compact operator. Then there exists  $N \in \mathbb{N} \cup \{\infty\}$ , orthonormal subsets  $\{\varphi_n\}_{n=1}^N \subset H$  and  $\{\psi_n\}_{n=1}^N \subset B$  and a sequence  $\{\alpha_n\}_{n=1}^N \subset \mathbb{R}_+$  such that  $\alpha_1 \geq \alpha_2 \geq \ldots$  (with  $\lim_{n\to\infty} \alpha_n = 0$  if  $N = \infty$ ),  $\|\psi_n\| \leq 1$  for all n and

$$Kf = \sum_{n=1}^{N} \alpha_n \langle f, \varphi_n \rangle \psi_n \text{ for all } f \in H.$$
(9.11)

**Proof.** Since  $K^*K$  is a self-adjoint compact operator, Theorem 9.24 implies there exists an orthonormal set  $\{\varphi_n\}_{n=1}^N \subset H$  and positive numbers  $\{\lambda_n\}_{n=1}^N$  such that

$$K^*K\psi = \sum_{n=1}^N \lambda_n \langle \psi, \varphi_n \rangle \varphi_n \text{ for all } \psi \in H.$$

Let A be the positive square root of  $K^*K$  defined by

$$A\psi := \sum_{n=1}^{N} \sqrt{\lambda_n} \langle \psi, \varphi_n \rangle \varphi_n \text{ for all } \psi \in H.$$

A simple computation shows,  $A^2 = K^* K$ , and therefore,

$$\begin{split} \left\| A\psi \right\|^2 &= \left\langle A\psi, A\psi \right\rangle = \left\langle \psi, A^2\psi \right\rangle \\ &= \left\langle \psi, K^*K\psi \right\rangle = \left\langle K\psi, K\psi \right\rangle = \left\| K\psi \right\|^2 \end{split}$$

for all  $\psi \in H$ . Hence we may define a unitary operator,  $u : \overline{\operatorname{Ran}(A)} \to \overline{\operatorname{Ran}(K)}$  by the formula

$$uA\psi = K\psi$$
 for all  $\psi \in H$ .

We then have

$$K\psi = uA\psi = \sum_{n=1}^{N} \sqrt{\lambda_n} \langle \psi, \varphi_n \rangle u\varphi_n \tag{9.12}$$

which proves the result with  $\psi_n := u\varphi_n$  and  $\alpha_n = \sqrt{\lambda_n}$ .

It is instructive to find  $\psi_n$  explicitly and to verify Eq. (9.12) by brute force. Since  $\varphi_n = \lambda_n^{-1/2} A \varphi_n$ ,

$$\psi_n = \lambda_n^{-1/2} u A \varphi_n = \lambda_n^{-1/2} K \varphi_n$$

and

$$\langle K\varphi_n, K\varphi_m \rangle = \langle \varphi_n, K^* K\varphi_m \rangle = \lambda_n \delta_{mn}.$$

This verifies that  $\{\psi_n\}_{n=1}^N$  is an orthonormal set. Moreover,

$$\begin{split} \sum_{n=1}^{N} \sqrt{\lambda_n} \langle \psi, \varphi_n \rangle \psi_n &= \sum_{n=1}^{N} \sqrt{\lambda_n} \langle \psi, \varphi_n \rangle \lambda_n^{-1/2} K \varphi_n \\ &= K \sum_{n=1}^{N} \langle \psi, \varphi_n \rangle \varphi_n = K \psi \end{split}$$

since  $\sum_{n=1}^{N} \langle \psi, \varphi_n \rangle \varphi_n = P \psi$  where P is orthogonal projection onto  $\operatorname{Nul}(K)^{\perp}$ . Second Proof. Let K = u |K| be the polar decomposition of K. Then |K|

**Second Proof.** Let K = u |K| be the polar decomposition of K. Then |K|is self-adjoint and compact, by Corollary ?? below, and hence by Theorem 9.24 there exists an orthonormal basis  $\{\varphi_n\}_{n=1}^N$  for  $\operatorname{Nul}(|K|)^{\perp} = \operatorname{Nul}(K)^{\perp}$  such that  $|K|\varphi_n = \lambda_n\varphi_n, \lambda_1 \ge \lambda_2 \ge \ldots$  and  $\lim_{n\to\infty} \lambda_n = 0$  if  $N = \infty$ . For  $f \in H$ ,

$$Kf = u |K| \sum_{n=1}^{N} \langle f, \varphi_n \rangle \varphi_n = \sum_{n=1}^{N} \langle f, \varphi_n \rangle u |K| \varphi_n = \sum_{n=1}^{N} \lambda_n \langle f, \varphi_n \rangle u \varphi_n$$

which is Eq. (9.11) with  $\psi_n := u\varphi_n$ .

**Exercise 9.2 (Continuation of Example ??).** Let  $H := L^2([0,1],m)$ ,  $k(x,y) := \min(x,y)$  for  $x, y \in [0,1]$  and define  $K : H \to H$  by

$$Kf(x) = \int_{0}^{1} k(x, y) f(y) \, dy.$$

From Example 9.14 we know that K is a compact operator<sup>4</sup> on H. Since k is real and symmetric, it is easily seen that K is self-adjoint. Show:

- 1. If  $g \in C^2([0,1])$  with g(0) = 0 = g'(1), then Kg'' = -g. Use this to conclude  $\langle Kf|g'' \rangle = -\langle f|g \rangle$  for all  $g \in C_c^{\infty}((0,1))$  and consequently that  $\operatorname{Nul}(K) = \{0\}$ .
- 2. Now suppose that  $f \in H$  is an eigenvector of K with eigenvalue  $\lambda \neq 0$ . Show that there is a version<sup>5</sup> of f which is in  $C([0,1]) \cap C^2((0,1))$  and this version, still denoted by f, solves

$$\lambda f'' = -f \text{ with } f(0) = f'(1) = 0.$$
(9.13)

where  $f'(1) := \lim_{x \uparrow 1} f'(x)$ .

 $<sup>^4</sup>$  See Exercise 9.3 from which it will follow that K is a Hilbert Schmidt operator and hence compact.

<sup>&</sup>lt;sup>5</sup> A measurable function g is called a version of f iff g = f a.e..

#### 66 9 Spectral Theorem (Compact Operator Case)

3. Use Eq. (9.13) to find all the eigenvalues and eigenfunctions of K. 4. Use the results above along with the spectral Theorem 9.24, to show

$$\left\{\sqrt{2}\sin\left(\left(n+\frac{1}{2}\right)\pi x\right):n\in\mathbb{N}_0\right\}$$

is an orthonormal basis for  $L^2([0,1],m)$  with  $\lambda_n = \left[\left(n+\frac{1}{2}\right)\pi\right]^{-2}$ .

5. Repeat this problem in the case that  $k(x, y) = \min(x, y) - xy$ . In this case you should find that Eq. (9.13) is replaced by

$$\lambda f'' = -f$$
 with  $f(0) = f(1) = 0$ 

from which one finds;

$$\left\{f_n := \sqrt{2}\sin\left(n\pi x\right) : n \in \mathbb{N}\right\}$$

is an orthonormal basis of eigenvectors of K with corresponding eigenvalues;  $\lambda_n = (n\pi)^{-2}$ .

6. Use the result of the last part to show,

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

Hint: First show

$$k(x,y) = \sum_{n=1}^{\infty} \lambda_n f_n(x) f_n(y) \text{ for a.e. } (x,y).$$

Then argue the above equation holds for every  $(x, y) \in [0, 1]^2$ . Finally take y = x in the above equation and integrate to arrive at the desired result.

**Note:** for a wide reaching generalization of this exercise the reader should consult Conway [?, Section II.6 (p.49-54)].

# 9.4 Hilbert Schmidt Operators

In this section H and B will be Hilbert spaces.

**Proposition 9.26.** Let H and B be a separable Hilbert spaces,  $K : H \to B$  be a bounded linear operator,  $\{e_n\}_{n=1}^{\infty}$  and  $\{u_m\}_{m=1}^{\infty}$  be orthonormal basis for H and B respectively. Then:

1.  $\sum_{n=1}^{\infty} ||Ke_n||^2 = \sum_{m=1}^{\infty} ||K^*u_m||^2$  allowing for the possibility that the sums are infinite. In particular the **Hilbert Schmidt norm** of K,

$$||K||_{HS}^2 := \sum_{n=1}^{\infty} ||Ke_n||^2,$$

is well defined independent of the choice of orthonormal basis  $\{e_n\}_{n=1}^{\infty}$ . We say  $K : H \to B$  is a **Hilbert Schmidt operator** if  $||K||_{HS} < \infty$  and let HS(H,B) denote the space of Hilbert Schmidt operators from H to B. 2. For all  $K \in L(H,B)$ ,  $||K||_{HS} = ||K^*||_{HS}$  and

$$||K||_{HS} \ge ||K||_{op} := \sup \{ ||Kh|| : h \in H \text{ such that } ||h|| = 1 \}.$$

3. The set HS(H, B) is a subspace of L(H, B) (the bounded operators from  $H \to B$ ),  $\|\cdot\|_{HS}$  is a norm on HS(H, B) for which  $(HS(H, B), \|\cdot\|_{HS})$  is a Hilbert space, and the corresponding inner product is given by

$$\langle K_1 | K_2 \rangle_{HS} = \sum_{n=1}^{\infty} \langle K_1 e_n | K_2 e_n \rangle \,. \tag{9.14}$$

4. If  $K : H \to B$  is a bounded finite rank operator, then K is Hilbert Schmidt. 5. Let  $P_N x := \sum_{n=1}^{N} \langle x | e_n \rangle e_n$  be orthogonal projection onto  $\operatorname{span}\{e_n : n \leq N\} \subset H$  and for  $K \in HS(H, B)$ , let  $K_N := KP_N$ . Then

$$||K - K_N||_{op}^2 \le ||K - K_N||_{HS}^2 \to 0 \text{ as } N \to \infty,$$

which shows that finite rank operators are dense in  $(HS(H, B), \|\cdot\|_{HS})$ . In particular of  $HS(H, B) \subset \mathcal{K}(H, B)$  – the space of compact operators from  $H \to B$ .

6. If Y is another Hilbert space and  $A: Y \to H$  and  $C: B \to Y$  are bounded operators, then

$$\|KA\|_{HS} \leq \|K\|_{HS} \, \|A\|_{op} \ \text{ and } \ \|CK\|_{HS} \leq \|K\|_{HS} \, \|C\|_{op} \, ,$$

in particular HS(H, H) is an ideal in L(H).

**Proof. Items 1. and 2.** By Parseval's equality and Fubini's theorem for sums,

$$\sum_{n=1}^{\infty} \|Ke_n\|^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\langle Ke_n | u_m \rangle|^2$$
$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\langle e_n | K^* u_m \rangle|^2 = \sum_{m=1}^{\infty} \|K^* u_m\|^2$$

This proves  $||K||_{HS}$  is well defined independent of basis and that  $||K||_{HS} =$  $||K^*||_{HS}$ . For  $x \in H \setminus \{0\}$ , x/||x|| may be taken to be the first element in an orthonormal basis for H and hence

$$\left\| K \frac{x}{\|x\|} \right\| \le \|K\|_{HS}.$$

Multiplying this inequality by ||x|| shows  $||Kx|| \leq ||K||_{HS} ||x||$  and hence 
$$\begin{split} \|K\|_{op} &\leq \|K\|_{HS} \,.\\ \textbf{Item 3. For } K_1, K_2 \in L(H,B), \end{split}$$

$$\begin{split} \|K_1 + K_2\|_{HS} &= \sqrt{\sum_{n=1}^{\infty} \|K_1 e_n + K_2 e_n\|^2} \\ &\leq \sqrt{\sum_{n=1}^{\infty} [\|K_1 e_n\| + \|K_2 e_n\|]^2} \\ &= \|\{\|K_1 e_n\| + \|K_2 e_n\|\}_{n=1}^{\infty}\|_{\ell_2} \\ &\leq \|\{\|K_1 e_n\|\}_{n=1}^{\infty}\|_{\ell_2} + \|\{\|K_2 e_n\|\}_{n=1}^{\infty}\|_{\ell_2} \\ &= \|K_1\|_{HS} + \|K_2\|_{HS} \,. \end{split}$$

From this triangle inequality and the homogeneity properties of  $\left\|\cdot\right\|_{HS},$  we now easily see that HS(H,B) is a subspace of L(H,B) and  $\|\cdot\|_{HS}$  is a norm on HS(H, B). Since

$$\sum_{n=1}^{\infty} |\langle K_1 e_n | K_2 e_n \rangle| \le \sum_{n=1}^{\infty} ||K_1 e_n|| \, ||K_2 e_n|| \\ \le \sqrt{\sum_{n=1}^{\infty} ||K_1 e_n||^2} \sqrt{\sum_{n=1}^{\infty} ||K_2 e_n||^2} = ||K_1||_{HS} \, ||K_2||_{HS} \,,$$

the sum in Eq. (9.14) is well defined and is easily checked to define an inner product on HS(H, B) such that  $||K||_{HS}^2 = \langle K|K\rangle_{HS}$ .

The proof that  $\left(HS(H,B), \|\cdot\|_{HS}^2\right)$  is complete is very similar to the proof of Theorem ??. Indeed, suppose  $\{K_m\}_{m=1}^{\infty}$  is a  $\|\cdot\|_{HS}$  – Cauchy sequence in HS(H,B). Because L(H,B) is complete, there exists  $K \in L(H,B)$  such that  $||K - K_m||_{op} \to 0$  as  $m \to \infty$ . Thus, making use of Fatou's Lemma ??,

$$\begin{split} \|K - K_m\|_{HS}^2 &= \sum_{n=1}^{\infty} \|(K - K_m) e_n\|^2 \\ &= \sum_{n=1}^{\infty} \lim \inf_{l \to \infty} \|(K_l - K_m) e_n\|^2 \\ &\leq \lim \inf_{l \to \infty} \sum_{n=1}^{\infty} \|(K_l - K_m) e_n\|^2 \\ &= \lim \inf_{l \to \infty} \|K_l - K_m\|_{HS}^2 \to 0 \text{ as } m \to \infty. \end{split}$$

Hence  $K \in HS(H, B)$  and  $\lim_{m \to \infty} ||K - K_m||_{HS}^2 = 0$ . Item 4. Since  $\operatorname{Nul}(K^*)^{\perp} = \overline{\operatorname{Ran}(K)} = \operatorname{Ran}(K)$ ,

$$||K||_{HS}^2 = ||K^*||_{HS}^2 = \sum_{n=1}^N ||K^*v_n||_H^2 < \infty$$

where  $N := \dim \operatorname{Ran}(K)$  and  $\{v_n\}_{n=1}^N$  is an orthonormal basis for  $\operatorname{Ran}(K) =$ K(H).

Item 5. Simply observe,

$$|K - K_N||_{op}^2 \le ||K - K_N||_{HS}^2 = \sum_{n>N} ||Ke_n||^2 \to 0 \text{ as } N \to \infty.$$

**Item 6.** For  $C \in L(B, Y)$  and  $K \in L(H, B)$  then

$$\|CK\|_{HS}^{2} = \sum_{n=1}^{\infty} \|CKe_{n}\|^{2} \le \|C\|_{op}^{2} \sum_{n=1}^{\infty} \|Ke_{n}\|^{2} = \|C\|_{op}^{2} \|K\|_{HS}^{2}$$

and for  $A \in L(Y, H)$ ,

$$|KA||_{HS} = ||A^*K^*||_{HS} \le ||A^*||_{op} ||K^*||_{HS} = ||A||_{op} ||K||_{HS}.$$

*Remark 9.27.* The separability assumptions made in Proposition 9.26 are unnecessary. In general, we define

$$||K||_{HS}^{2} = \sum_{e \in \beta} ||Ke||^{2}$$

where  $\beta \subset H$  is an orthonormal basis. The same proof of Item 1. of Proposition 9.26 shows  $||K||_{HS}$  is well defined and  $||K||_{HS} = ||K^*||_{HS}$ . If  $||K||_{HS}^2 < \infty$ , then there exists a countable subset  $\beta_0 \subset \beta$  such that Ke = 0 if  $e \in \beta \setminus \beta_0$ . Let  $H_0 := \overline{\operatorname{span}(\beta_0)}$  and  $B_0 := \overline{K(H_0)}$ . Then  $K(H) \subset B_0, K|_{H_0^{\perp}} = 0$  and hence by applying the results of Proposition 9.26 to  $K|_{H_0}: H_0 \to B_0$  one easily sees that the separability of H and B are unnecessary in Proposition 9.26.

Example 9.28. Let  $(X, \mu)$  be a measure space,  $H = L^2(X, \mu)$  and

$$k(x, y) := \sum_{i=1}^{n} f_i(x) g_i(y)$$

where

$$f_i, g_i \in L^2(X, \mu) \text{ for } i = 1, ..., n.$$

Define

$$(Kf)(x) = \int_X k(x, y) f(y) d\mu(y)$$

then  $K: L^2(X,\mu) \to L^2(X,\mu)$  is a finite rank operator and hence Hilbert Schmidt.

**Exercise 9.3.** Suppose that  $(X, \mu)$  is a  $\sigma$ -finite measure space such that  $H = L^2(X, \mu)$  is separable and  $k: X \times X \to \mathbb{R}$  is a measurable function, such that

$$\|k\|_{L^{2}(X \times X, \mu \otimes \mu)}^{2} := \int_{X \times X} |k(x, y)|^{2} d\mu(x) \, d\mu(y) < \infty$$

Define, for  $f \in H$ ,

$$Kf(x) = \int_{X} k(x, y) f(y) d\mu(y),$$

when the integral makes sense. Show:

- 1. Kf(x) is defined for  $\mu$ -a.e. x in X.
- 2. The resulting function Kf is in H and  $K: H \to H$  is linear.
- 3.  $||K||_{HS} = ||k||_{L^2(X \times X, \mu \otimes \mu)} < \infty$ . (This implies  $K \in HS(H, H)$ .)

**Exercise 9.4 (Converse to Exercise 9.3).** Suppose that  $(X, \mu)$  is a  $\sigma$ -finite measure space such that  $H = L^2(X, \mu)$  is separable and  $K : H \to H$  is a Hilbert Schmidt operator. Show there exists  $k \in L^2(X \times X, \mu \otimes \mu)$  such that K is the integral operator associated to k, i.e.

$$Kf(x) = \int_{X} k(x, y) f(y) \, d\mu(y) \,. \tag{9.15}$$

In fact you should show

$$k(x,y) := \sum_{n=1}^{\infty} \left( \left( \overline{K^* \varphi_n} \right)(y) \right) \varphi_n(x) \quad (L^2(\mu \otimes \mu) - \text{convergent sum})$$
(9.16)

where  $\{\varphi_n\}_{n=1}^{\infty}$  is any orthonormal basis for *H*.