TOPICS IN FOURIER ANALYSIS

DANIEL W. STROOCK

0. INTRODUCTION

This is a set of notes that I wrote for a course that I intended to but did not give at MIT during the spring semester of 2024. It covers a number of topics related to the theory and application of Fourier analysis.

I begin in §1 by proving the L^2-convergence of Fourier followed by elementary results about pointwise convergence for sufficiently smooth periodic functions. In §2 I discuss what goes wrong in the absence of periodicity, and in §3 I apply Fourier series to compute the Riemann ζ at odd integers using the Bernoulli polynomials, which I also use to develop the Euler–Maclauren series. After comparing summability methods in §4, I give a brief introduction in §5 to the summability results of Dirichlet, Feijér, and Lebesgue.

In section §6 I introduce the L^1-Fourier transform, followed in §7 by the computation of the Fourier transforms of the Gauss and Poisson kernels and the derivation and application of the Poisson summation formula. The L^1 version of the Fourier inversion formula is proved in §8. In §§9–11 I make preparations for my treatment in §12 of the L^2-Fourier transform via Hermite functions. By the end of §12, I have covered the key results in that theory: Parseval's identity and the Fourier inversion formula.

In §13 I introduce the test function space on which Laurent Schwartz based his theory of tempered distributions. As was the case in my treatment of the L^2-Fourier transform, Hermite functions play a central role here. In §14 I give the definition of and do a few computations with tempered distributions, and in §15 I show how to extend continuous operations on the test function space as continuous operations on tempered distributions.

In §§1–15 I have restricted my results to the one dimensional setting, and it is only in §16 that I describe what has to be done to extend those results to more than one dimension. Once I have done so, in §17 I introduce the weak topology on the space of Borel probability on $\mathbb{R}^N$, and in §18 I show that there is an intimate relationship between that topology and Fourier analysis. The results in §18 are combined with those in §14 to derive in §19 the Lévy–Khinchine formula for infinitely divisible probability measures.

The rest of these notes is devoted to the theory of singular integral operators. After a brief attempt in §20 to provide motivation, in §21 I derive the L^p boundedness of the Hilbert transform when p is an even integer, and in §22 I prove the Riesz–Thorin interpolation theory in order the extend that result to all $p \in (1, \infty)$. Finally, in §23 I use Calderòn and Zygmund's method of rotations to prove L^p boundedness of odd Calderòn–Zygmund kernels.

In so far as possible, I have tried to avoid the use of unfamiliar results, but I am well aware that what is familiar to some may be unfamiliar to others. At a

minimum, the reader is expected to had a rigorous course in Lebesgue integration theory. In addition, I have assumed some comfort with the ideas of elementary functional analysis, especially Hilbert spaces. Other than that, the only prerequisites are an interest in mathematics and a willingness to do computations.

1. Basic Theory of Fourier Series

Set $\mathfrak{e}_m(x) = e^{\imath 2\pi m x}$ for $m \in \mathbb{Z}$ and $x \in \mathbb{R}$, and observe that $\{\mathfrak{e}_m : m \in \mathbb{Z}\}$ is an orthonormal family in $L^2(\lambda_{[0,1)}; \mathbb{C})$.[1] Even though it involves an abuse of notation, we will use $(\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})}$ to denote $\int_{[0,1)} \varphi(y)\mathfrak{e}_{-m}(y)\, dy$ for $\varphi \in L^1(\lambda_{[0,1)}; \mathbb{C})$.

Given a function $\varphi : [0,1) \longrightarrow \mathbb{C}$, define its periodic extension $\tilde{\varphi} : \mathbb{R} \longrightarrow \mathbb{C}$ by $\tilde{\varphi}(x) = \varphi\big(x - \lfloor x \rfloor\big)$, where $\lfloor x \rfloor = \max\{n \in \mathbb{Z} : x \geq n\}$. Notice that if $\varphi \in L^1(\lambda_{[0,1)}; \mathbb{C})$, then

$$\int_{[0,1)} \varphi(x)\, dx = \int_{[a,a+1)} \tilde{\varphi}(x)\, dx \text{ for all } a \in \mathbb{R}.$$

Similarly,

$$\int_{[0,1)} \tilde{\varphi}(-x)\, dx = \int_{[0,1)} \varphi(x)\, dx.$$

For bounded, continuous functions φ and ψ on $[0,1)$, define

$$\varphi * \psi(x) = \int_{[0,1)} \varphi(x-y)\psi(y)\, dy,$$

and use the preceding to check that

$$\varphi * \psi(x) = \int_{[-x,-x+1]} \tilde{\varphi}(y)\tilde{\psi}(x-y)\, dy = \psi * \varphi(x).$$

Finally, by the continuous version of Minkowski's inequality,[2]

$$\|\varphi * \psi\|_{L^p(\lambda_{[0,1)};\mathbb{C})} \leq \|\varphi\|_{L^p(\lambda_{[0,1)};\mathbb{C})}\|\psi\|_{L^1(\lambda_{[0,1)};\mathbb{C})} \wedge \|\psi\|_{L^p(\lambda_{[0,1)};\mathbb{C})}\|\varphi\|_{L^1(\lambda_{[0,1)};\mathbb{C})}$$

for any $p \in [1, \infty)$. Hence, for each $p \in [1, \infty)$, $(\varphi, \psi) \rightsquigarrow \varphi * \psi$ has a unique continuous extension as a map bilinear map from $L^1(\lambda_{[0,1)}; \mathbb{C}) \times L^p(\lambda_{[0,1)}; \mathbb{C})$ into $L^p(\lambda_{[0,1)}; \mathbb{C})$, and

(1.1) $$\|\varphi * \psi\|_{L^p(\lambda_{[0,1)};\mathbb{C})} \leq \|\varphi\|_{L^1(\lambda_{[0,1)};\mathbb{C})}\|\psi\|_{L^p(\lambda_{[0,1)};\mathbb{C})}$$

continues to hold.

Theorem 1.1. *If $\varphi \in L^p\big(\lambda_{[0,1]}; \mathbb{C}\big)$ for some $p \in [1, \infty)$, then*

$$\lim_{r \nearrow 1} \left\| \varphi - \sum_{m \in \mathbb{Z}} r^{|m|} \big(\varphi, \mathfrak{e}_m\big)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m \right\|_{L^p(\lambda_{[0,1]};\mathbb{C})} = 0,$$

and, if $\varphi \in C\big([0,1]; \mathbb{C}\big)$ satisfies $\varphi(0) = \varphi(1)$, then[3]

$$\lim_{r \nearrow 1} \left\| \varphi - \sum_{m \in \mathbb{Z}} r^{|m|} \big(\varphi, \mathfrak{e}_m\big)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m \right\|_{\mathrm{u}} = 0.$$

---

[1] For a measure space $(E, \mathcal{F}, \mu)$ and $p \in [1, \infty]$, $L^p(\mu; \mathbb{C})$ is the associated Lebesgue space. For a Borel measurable subset $S \subseteq \mathbb{R}^N$, λ_S is the Lebesgue's measure resticted to S.

[2] If $\varphi \in L^p(\mu; \mathbb{C})$, then $\|\varphi\|_{L^p(\mu;\mathbb{C})}$ is its L^p-norm.

[3] $\|\cdot\|_{\mathrm{u}}$ is the uniform (i.e., supremum norm).

Proof. Define

$$p_r(x) = \sum_{m \in \mathbb{Z}} r^{|m|} \mathfrak{e}_m(x) \text{ for } r \in [0,1) \text{ and } x \in [0,1).$$

Clearly $\int_0^1 p_r(x)\,dx = 1$, $p_r(-x) = p_r(x)$, and $\widetilde{p}_r$ is continuous. In addition,

$$p_r(x) = \frac{1}{1 - r\mathfrak{e}_1(x)} + \frac{r\mathfrak{e}_{-1}(x)}{1 - r\mathfrak{e}_{-1}(x)} = \frac{1 - r^2}{|1 - r\mathfrak{e}_1(x)|^2} = \frac{1 - r^2}{1 - 2r\cos 2\pi x + r^2} \text{ for } r \in [0,1),$$

and so $p_r \geq 0$.

Obviously,

$$\sum_{m \in \mathbb{Z}} r^{|m|} (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m(x) = p_r * \varphi(x) = \int_{[0,1)} p_r(y)\tilde{\varphi}(x+y)\,dy$$

since p_r is even. Now suppose that $\varphi \in C([0,1] : \mathbb{C})$ with $\varphi(0) = \varphi(1)$. Then, since $\lim_{r \nearrow 1} \int_\delta^1 p_r(y)\,dy = 0$ for each $\delta \in (0,1)$, it is easy to check that

$$\lim_{r \nearrow 1} \sup_{x \in [0,1]} \left| \int_0^1 \left(\varphi(x+y)\,dy - f(x)\right)\right| \leq \omega_\varphi(\delta),$$

where ω_φ is the modulus of continuity of φ. Thus the second part of the theorem has been proved.

To prove the first part, let $\varphi \in L^p(\lambda_{[0,1)};\mathbb{C})$, and choose choose a sequence $\{\varphi_k : k \geq 1\} \subseteq C([0,1];\mathbb{C})$ which satisfy $\varphi_k(0) = \varphi_k(1)$ and $\|\varphi - \varphi_k\|_{L^p(\lambda_{[0,1]};\mathbb{C})} \longrightarrow 0$ as $k \to \infty$. Then, for each k,

$$\|p_r * \varphi - \varphi\|_{L^p(\lambda_{[0,1]};\mathbb{C})}$$
$$\leq \|p_r * (\varphi - \varphi_k)\|_{L^p(\lambda_{[0,1]};\mathbb{C})} + \|p_r * \varphi_k - \varphi_k\|_{L^p(\lambda_{[0,1]};\mathbb{C})} + \|\varphi_k - \varphi\|_{L^p(\lambda_{[0,1]};\mathbb{C})},$$

and so, by (1.1), for all k.

$$\lim_{r \nearrow 1} \|p_r * \varphi - \varphi\|_{L^p(\lambda_{[0,1]};\mathbb{C})} \leq 2\|\varphi_k - \varphi\|_{L^p(\lambda_{[0,1]};\mathbb{C})}.$$

Finally, let $k \to \infty$. □

Theorem 1.2. $\{\mathfrak{e}_m : m \in \mathbb{Z}\}$ *is an orthonormal basis in* $L^2(\lambda_{[0,1)};\mathbb{C})$, *and so, for each* $\varphi \in L^2(\lambda_{[0,1)};\mathbb{C})$,

$$(1.2) \qquad \sum_{m \in \mathbb{Z}} (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m \equiv \lim_{n \to \infty} \sum_{|m| \leq n} (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m = \varphi,$$

where the convergence is in $L^2(\lambda_{[0,1)};\mathbb{C})$. *In addition, for all* $\varphi, \psi \in L^2(\lambda_{[0,1)};\mathbb{C})$,

$$(\varphi, \psi)_{L^2(\lambda_{[0,1)};\mathbb{C})} = \sum_{m \in \mathbb{Z}} (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \overline{(\psi, \mathfrak{e}_m)}_{L^2(\lambda_{[0,1)};\mathbb{C})}.$$

Proof. It suffices to check the first statement, and to do so all we need to know is that $(\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} = 0$ for all $m \in \mathbb{Z}$ implies $\varphi = 0$ for a set of φ's which is dense in $L^2(\lambda_{[0,1)};\mathbb{C})$. But, by Theorem 1.1, we know this for continuous φ's satisfying $\varphi(0) = \varphi(1)$, and these are dense in $L^2(\lambda_{[0,1)};\mathbb{C})$. □

Equation (1.2) is known as *Parseval's identity* for Fourier series.

Define the partial sum $S_n\varphi = \sum_{|m| \leq n} (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m$.

Corollary 1.3. *If $\varphi \in C([0,1];\mathbb{C})$ and*

$$\sum_{m \neq 0} \left| (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \right| < \infty,$$

then the series

$$\sum_{m \in \mathbb{Z}} (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \mathfrak{e}_m(x)$$

is uniformly absolutely convergent to φ. In fact,

$$\left\| S_n(\varphi) - \varphi \right\|_{\mathrm{u}} \leq \sum_{|m|>n} \left| (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \right|.$$

Proof. That the series if uniformly absolutely convergent is obvious. To see that it must be converging to φ, let ψ be uniform limit of $\{S_n\varphi : n \geq 0\}$. Then ψ is continuous and, because φ is the $L^2(\lambda_{[0,1)};\mathbb{C})$ limit of this series, $\psi = \varphi$ $\lambda_{[0,1]}$-almost everywhere, which, since both are continues, means that they are equal everywhere. Given these statements, the final estimate is trivial. $\qquad\square$

Lemma 1.4. *Let $\ell \geq 1$ and assume that $\varphi \in C^\ell([0,1];\mathbb{C})$ satisfies $\varphi^{(k)}(0) = \varphi^{(k)}(1)$ for $0 \leq k \leq \ell - 1$. Then*

$$(\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} = \left(\frac{\imath}{2\pi m} \right)^\ell \left(\varphi^{(\ell)}, \mathfrak{e}_m \right)_{L^2(\lambda_{[0,1)};\mathbb{C})} \text{ for } m \neq 0.$$

Proof. Clearly it suffices that prove the result when $\ell = 1$. To do so, use integration by parts and the condition $\varphi(0) = \varphi(1)$ to check that

$$\int_0^1 \varphi(y) \mathfrak{e}_{-m}(y)\, dy = \frac{1}{-\imath 2\pi m} \int_0^1 \varphi'(y) \mathfrak{e}_{-m}(y)\, dy.$$

$$\square$$

As a consequence of Lemma 1.4, we see that if $\varphi \in C^1([0,1];\mathbb{C})$ satisfies $\varphi(0) = \varphi(1)$, then

$$\sum_{|m|>n} \left| (\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \right| \leq \sum_{|m|>n} \frac{\left| (\varphi', \mathfrak{e}_m)_{L^2(\lambda_{[0,1)};\mathbb{C})} \right|}{2\pi |m|}$$

$$\leq \frac{1}{2\pi} \left(2 \sum_{m>n} m^{-2} \right)^{\frac{1}{2}} \|\varphi'\|_{L^2(\lambda_{[0,1)};\mathbb{C})} \leq \frac{\|\varphi'\|_{L^2(\lambda_{[0,1)};\mathbb{C})}}{\pi (2n)^{\frac{1}{2}}}.$$

Hence, by Corollary 1.3,

$$\|S_n\varphi - \varphi\|_{\mathrm{u}} \leq \frac{\|\varphi'\|_{\mathrm{u}}}{\pi (2n)^{\frac{1}{2}}}.$$

Exercise 1.1. Prove the *Riemann–Lebesgue lemma*, which is the statement that $\lim_{n \to \infty} (\varphi, \mathfrak{e}_n)_{L^2(\lambda_{[0,1)};\mathbb{C})} = 0$ for all $\varphi \in L^1(\lambda_{[0,1)};\mathbb{C})$.

Exercise 1.2. Let φ be a Lipschitz continuous function satisfying $\varphi(0) = \varphi(1)$, and show that

$$\|S_n\varphi - \varphi\|_{\mathrm{u}} \leq \frac{\|\varphi\|_{\mathrm{Lip}}}{\pi (2n)^{\frac{1}{2}}}.$$

Hint: Introduce the functions $\varphi_k = p_{\frac{1}{k}} * \varphi$.

2. Gibbs Phenomenon

Here we will examine what can be said for a $\varphi \in C([0,1]; \mathbb{C})$ that is not periodic. For example, consider the function $\varphi(x) = x$ for $x \in [0,1]$. Clearly

$$(\varphi, \mathfrak{e}_m)_{L^2(\lambda_{[0,1)}; \mathbb{C})} = \frac{\imath}{2\pi m} \text{ for } m \neq 0,$$

and so

$$S_n(x) = \frac{1}{2} - \frac{1}{\pi} \sum_{m=1}^{n} \frac{\sin 2\pi m x}{m},$$

where $S_n \equiv S_n \varphi$. Now set

$$\Phi_m(x) = \sum_{k=1}^{m} \sin 2\pi k x.$$

Then $\Phi_m(x)$ is the imaginary part of

$$\sum_{k=1}^{m} \mathfrak{e}_k(x) = \mathfrak{e}_1(x) \frac{1 - \mathfrak{e}_m(x)}{1 - \mathfrak{e}_1(x)} = \frac{\big(\mathfrak{e}_1(x) - \mathfrak{e}_{m+1}(x)\big)\big(1 - \mathfrak{e}_{-1}(x)\big)}{2(1 - \cos 2\pi x)}$$

$$= \frac{\mathfrak{e}_1(x) - 1 - \mathfrak{e}_{m+1} + \mathfrak{e}_m(x)}{2(1 - \cos 2\pi x)},$$

which is

$$\frac{\sin 2\pi x - \sin 2\pi(m+1)x + \sin 2\pi m x}{2(1 - \cos 2\pi x)}.$$

After using some of trigonometric identities, one sees that

$$(2.1) \qquad \Phi_m(x) = \frac{\cos \pi x \sin^2 \pi m x}{\sin \pi x} + \sin \pi m x \cos \pi m x.$$

In particular, $|\Phi_m(x)| \leq 3\big(\frac{1}{x} \vee \frac{1}{1-x}\big)$.

Summing by parts, one sees that

$$S_n(x) = \frac{1}{2} - \frac{\Phi_n(x)}{\pi n} - \sum_{m=1}^{n-1} \frac{\Phi_m(x)}{\pi m(m+1)},$$

which means that

$$(2.2) \qquad \big|S_n(x) - x\big| \leq \big(\tfrac{1}{x} \vee \tfrac{1}{1-x}\big) \frac{6}{\pi n}.$$

In particular, $S_n(x)$ is converging to x uniformly on compact subsets of $(0,1)$.

To see what happens for x near to 0, consider $x = \frac{k}{2n}$ for $k \geq 1$, and observe that

$$\sum_{m=1}^{n} \frac{\sin \frac{\pi k m}{n}}{m} = \frac{1}{n} \sum_{m=1}^{n} \frac{\sin \frac{\pi k m}{n}}{\frac{m}{n}} \longrightarrow \int_{[0,1]} \frac{\sin \pi k x}{x}\, dx \longrightarrow \int_{[0,\pi k]} \frac{\sin x}{x}\, dx.$$

Hence, since (cf. (7.11) in §7)

$$\lim_{R \to \infty} \int_{[0,R]} \frac{\sin x}{x}\, dx = \frac{\pi}{2},$$

$$S_n\left(\tfrac{k}{2n}\right) = -\frac{1}{\pi}\lim_{R\to\infty}\int_{[\pi k,R]}\frac{\sin x}{x}\,dx$$

$$= \frac{(-1)^{k+1}}{\pi^2 k} - \frac{1}{\pi}\int_{[\pi k,\infty)}\frac{\cos x}{x^2}\,dx = \frac{(-1)^{k+1}}{\pi^2 k} + \frac{2}{\pi}\int_{[\pi k,\infty)}\frac{\sin x}{x^3}\,dx$$

as $n \to \infty$. Therefore

$$S_n\left(\tfrac{k}{2n}\right) = \frac{(-1)^{k+1}}{\pi^2 k}\left(1 - \frac{a_k}{\pi k}\right) + \epsilon_n(k),$$

where

$$a_k = (-1)^k 2(\pi k)^2 \int_{\pi k}^{\infty}\frac{\sin x}{x^3}\,dx \in (-1,1)$$

and $\lim_{n\to\infty}\epsilon_n(k) = 0$. This shows that, for large n, $S_n\left(\tfrac{k}{2n}\right)$ is at least $\frac{1}{2\pi^2 k}$ if k is odd and at most $-\frac{1}{2\pi^2 k}$ if k is even. This sort of oscillatory behavior is known as *Gibbs's phenomenon*, although Gibbs seems not to have been the first to discover it.

Exercise 2.1. By considering $S_n\left(\tfrac{1}{4}\right)$ and using equations (2.1) and (2.2), show that

$$\pi = 8\sum_{\ell=0}^{\infty}\frac{1}{(4\ell+1)(4\ell+3)}.$$

Exercise 2.2. Show that if $\varphi \in C^1\big([0,1];\mathbb{C}\big)$ then,

$$\sup_{x\in[n^{-\frac{1}{2}},1-n^{-\frac{1}{2}}]}\big|S_n\varphi(x) - \varphi(x)\big| \le \frac{8\|\varphi'\|_{L^2(\lambda_{[0,1)};\mathbb{C})}}{\pi n^{\frac{1}{2}}}.$$

3. Bernoulli Polynomials

Theorem 3.1. *Define $\{b_\ell : \ell \ge 0\} \subseteq \mathbb{R}$ inductively by*

$$b_0 = 1 \ \text{and}\ b_{\ell+1} = \sum_{k=0}^{\ell}\frac{(-1)^k b_{\ell-k}}{(k+2)!},$$

and set

(3.1) $$B_\ell(x) = \sum_{k=0}^{\ell}\frac{(-1)^k b_{\ell-k}}{k!}x^k \ \text{for}\ \ell \ge 0.$$

Then $\{B_\ell : \ell \ge 0\}$ are the one and only functions satisfying

(3.2) $B_0 = \mathbf{1}$, $B'_{\ell+1} = -B_\ell$ *for* $\ell \ge 0$, *and* $B_\ell(1) = B_\ell(0)$ *for* $\ell \ge 2$.

Proof. To see that there is at most one set of functions satisfying (3.2), let $\{D_\ell : \ell \ge 0\}$ be the set of differences between two solutions, and set $\ell = \inf\{\ell : D_\ell \ne \mathbf{0}\}$. Then $\ell \ge 1$, and, if $\ell < \infty$, then D_ℓ is a constant a and there is a $b \in \mathbb{R}$ such that $D_{\ell+1}(x) = -ax + b$. But $-a + b = D_{\ell+1}(1) = D_{\ell+1}(0) = b$, and therefore $a = 0$. Since this would mean that $D_\ell = -D'_{\ell+1} = \mathbf{0}$, no such ℓ can exist.

By definition, $B_0 = \mathbf{1}$, and it is easy to check that $B'_{\ell+1} = -B_\ell$. To verify the periodicity property, note that

$$B_{\ell+2}(1) - B_{\ell+2}(0) = \sum_{k=1}^{\ell+2} \frac{(-1)^k b_{\ell+2-k}}{k!}$$

$$= -b_{\ell+1} + \sum_{k=2}^{\ell+2} \frac{(-1)^k b_{\ell+2-k}}{k!} = -b_{\ell+1} + \sum_{k=0}^{\ell} \frac{(-1)^k b_{\ell-k}}{(k+2)!} = 0.$$

$\square$

The functions $\{B_\ell : \ell \geq 0\}$ in (3.1) are known as *Bernoulli polynomials*.

Theorem 3.2. *For $\ell \geq 2$ and $x \in [0,1]$,*

$$(3.3) \qquad B_\ell(x) = \frac{-\imath^\ell}{(2\pi)^\ell} \sum_{n \neq 0} \frac{\mathfrak{e}_n(x)}{n^\ell}.$$

In particular, $b_{2\ell+1} = 0$ and

$$(3.4) \qquad \zeta(2\ell) \equiv \sum_{m=1}^{\infty} \frac{1}{m^{2\ell}} = (-1)^{\ell+1} 2^{2\ell-1} \pi^{2\ell} b_{2\ell}$$

for $\ell \geq 1$.

Proof. First observe that, for $\ell \geq 1$,

$$\left(B_\ell, \mathfrak{e}_0\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} = -\int_0^1 B'_{\ell+1}(x)\,dx = B_{\ell+1}(0) - B_{\ell+1}(1) = 0$$

and, for $\ell \geq 2$ and $n \neq 0$,

$$\left(B_\ell, \mathfrak{e}_n\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} = \frac{\imath}{2\pi n}\left(B_{\ell-1}, \mathfrak{e}_n\right)_{L^2(\lambda_{[0,1]};\mathbb{C})}$$

and therefore

$$\left(\frac{2\pi n}{\imath}\right)^{\ell-1}\left(B_\ell, \mathfrak{e}_n\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} = \left(B_1, \mathfrak{e}_n\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} = \int_0^1 \left(\tfrac{1}{2} - x\right)\mathfrak{e}_{-n}(x)\,dx = \frac{-\imath}{2\pi n}.$$

Hence

$$\left(B_\ell, \mathfrak{e}_n\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} = \frac{-\imath^\ell}{(2\pi n)^\ell}$$

for $\ell \geq 2$ and $n \neq 0$, which completes the proof of (3.3). Finally, because $b_\ell = B_\ell(0)$, it is clear from (3.3) that $b_{2\ell+1} = 0$ and that (3.4) holds. $\square$

Besides (3.4), the Bernoulli polynomials play a critical role in what is known as the *Euler–Maclauren formula*:

$$(3.5) \qquad \begin{aligned} &\int_0^n f(x)\,dx - \sum_{m=1}^n f(m) \\ &= -\sum_{k=1}^{\ell} b_k\left(f^{(k-1)}(n) - f^{k-1}(0)\right) + \int_0^n \tilde{B}_\ell(x) f^{(\ell)}(x)\,dx \end{aligned} \qquad \text{for } \ell \geq 1,$$

where $\tilde{B}_\ell$ is the periodic extension of $B_\ell \restriction [0,1)$ to $\mathbb{R}$. To prove (3.5), first note that

$$\int_0^n f(x)\,dx - \sum_{m=1}^n f(m) = \sum_{m=1}^n \int_{m-1}^m \big(f(x) - f(m)\big)\,dx$$

$$= -\sum_{m=1}^n \int_{m-1}^m \big(x - (m-1)\big)f'(x)\,dx$$

$$= \sum_{m=1}^n \left(-b_1\big(f(m) - f(m-1)\big) + \int_{m-1}^m B_1\big(x - (m-1)\big)f'(x)\,dx\right)$$

$$= -b_1\big(f(n) - f(0)\big) + \int_0^n \tilde{B}_1(x)f'(x)\,dx.$$

Hence, (3.5) holds when $\ell = 1$. Next observe that for any $\ell \geq 1$,

$$\int_0^n \tilde{B}_\ell(x) = n\int_0^1 B_\ell(x)\,dx = n\big(B_{\ell+1}(1) - B_{\ell+1}(0)\big) = 0,$$

and therefore

$$\int_0^n \tilde{B}_\ell(x)f^{(\ell)}(x)\,dx = \sum_{m=1}^n \int_{m-1}^m B_\ell\big(x - (m-1)\big)\big(f^{(\ell)}(x) - f^{(\ell)}(m)\big)\,dx$$

$$= \sum_{m=1}^n \left(-b_{\ell+1}\big(f^{(\ell)}(m) - f^{(\ell)}(m-1)\big) + \int_{m-1}^m B_{\ell+1}\big(x - (m-1)\big)f^{(\ell+1)}(x)\,dx\right)$$

$$= -b_{\ell+1}\big(f^{(\ell)}(n) - f(0)\big) + \int_0^n \tilde{B}_{\ell+1}(x)f^{(\ell+1)}(x)\,dx.$$

Therefore, (3.5) for ℓ implies (3.5) for $\ell + 1$.

Theorem 3.3. *If $\ell \geq 1$ and $\varphi \in C^\ell\big([0,1];\mathbb{C}\big)$, then*

(3.6)
$$\int_0^1 \varphi(x) - \frac{1}{n}\sum_{m=1}^n \varphi\big(\tfrac{m}{n}\big)$$
$$= -\sum_{k=1}^\ell \frac{b_k}{n^k}\big(\varphi^{(k-1)}(1) - \varphi^{(k-1)}(0)\big) + \frac{1}{n^\ell}\int_0^1 \tilde{B}_\ell(nx)\varphi^{(\ell)}(x)\,dx,$$

Proof. Take $f(x) = \varphi\big(\frac{x}{n}\big)$, apply (3.5) to f, and make a simple change of variables. $\qquad\square$

By Schwarz's inequality,

$$\left|\int_0^1 \tilde{B}_\ell(nx)\varphi^{(\ell)}(x)\,dx\right| \leq \left(\int_0^1 \tilde{B}_\ell(nx)^2\,dx\right)^{\frac{1}{2}}\|\varphi^{(\ell)}\|_{L^2(\lambda_{[0,1]};\mathbb{C})},$$

and

$$\int_0^1 \tilde{B}_\ell(nx)^2\,dx = \frac{1}{n}\int_0^n \tilde{B}_\ell(x)^2\,dx = \|B_\ell\|^2_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}.$$

Further, by (1.2) and (3.3),

$$\|B_\ell\|^2_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \frac{1}{(2\pi)^{2\ell}}\sum_{n\neq 0}\frac{1}{n^{2\ell}}.$$

Hence, by (3.6),

$$(3.7) \quad \left| \int_0^1 \varphi(x)\, dx - \frac{1}{n} \sum_{m=1}^n \varphi\left(\tfrac{m}{n}\right) + \sum_{k=1}^\ell \frac{b_k}{n^k} \left(\varphi^{(k-1)}(1) - \varphi^{(k-1)}(0)\right) \right|$$

$$\leq \frac{\sqrt{2\zeta(2\ell)}}{(2\pi n)^\ell} \|\varphi^{(\ell)}\|_{L^2(\lambda_{[0,1]};\mathbb{C})}.$$

From (3.7) one sees that if, for some $n \geq 1$,

$$(3.8) \qquad \lim_{\ell \to \infty} \frac{\|\varphi^{(\ell)}\|_{L^2(\lambda_{[0,1]};\mathbb{C})}}{(2\pi n)^\ell} = 0,$$

then

$$\int_0^1 \varphi(x)\, dx - \frac{1}{n} \sum_{m=1}^n \varphi\left(\tfrac{m}{n}\right) = - \lim_{\ell \to \infty} \sum_{k=1}^\ell \frac{b_k}{n^k} \left(\varphi^{(k-1)}(1) - \varphi^{(k-1)}(0)\right).$$

In particular, if $\varphi \in C^\infty\big([0,1];\mathbb{C}\big)$ and $\varphi^{(k)}$ is periodic for all $k \geq 0$, then (3.8) implies that

$$\int_0^1 \varphi(x)\, dx = \frac{1}{n} \sum_{m=1}^n \varphi\left(\tfrac{m}{n}\right),$$

a result that has a much simpler derivation (cf. Exercise 3.1 below).

   More generally, because $\left|\varphi^{(k-1)}(1) - \varphi^{(k-1)}(0)\right| \leq \|\varphi^{(k)}\|_{L^2(\lambda_{[0,1]};\mathbb{C})}$ and $|b_k| \leq \frac{1}{(2\pi)^k}$,

$$\sum_{k=1}^\infty \frac{\|\varphi^{(k)}\|_{L^2(\lambda_{[0,1]};\mathbb{C})}}{(2\pi n)^k} < \infty$$

implies that

$$(3.9) \qquad \int_0^1 \varphi(x)\, dx - \frac{1}{n} \sum_{m=1}^n \varphi\left(\tfrac{m}{n}\right) = - \sum_{k=1}^\infty \frac{b_k}{n^k} \left(\varphi^{(k-1)}(1) - \varphi^{(k-1)}(0)\right),$$

where the series is absolutely convergent.

Exercise 3.1. Suppose that φ and all its derivatives are periodic on $[0,1]$, and show that

$$\lim_{\ell \to \infty} \frac{\|\varphi^{(\ell)}\|_{L^2(\lambda_{[0,1]};\mathbb{C})}}{(2\pi n)^\ell} = 0 \iff \left(\varphi, \mathfrak{e}_m\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} = 0 \text{ if } |m| \geq n$$

$$\iff \varphi = \sum_{|m| < n} \left(\varphi, \mathfrak{e}_m\right)_{L^2(\lambda_{[0,1]};\mathbb{C})} \mathfrak{e}_m.$$

Next, show that

$$\frac{1}{n} \sum_{j=1}^n \mathfrak{e}_m\left(\tfrac{j}{n}\right) = 0$$

for $1 \leq |m| < n$, and thereby arrive at the conclusion reached above.

4. Comparing Summability Methods

In preparation for the following section, we will review here basic definitions and results for different notions of convergence of a series.

Given a sequence $\{a_m : m \geq 1\} \subseteq \mathbb{C}$, set

$$S_n = \sum_{m=1}^n a_m \text{ and } A_n = \frac{1}{n} \sum_{m=1}^n S_m,$$

and when $\overline{\lim}_{n \to \infty} |a_m|^{\frac{1}{m}} \leq 1$, set

$$A(r) = \sum_{m=1}^\infty a_m r^{m-1} \text{ for } r \in [0, 1).$$

The S_n's are called the *partial sums* of the corresponding series, the A_n's are its *Césaro means*, and $r \rightsquigarrow A(r)$ is its *Abel function*. The series is said to be *summable* to $s \in \mathbb{C}$ if $s = \lim_{n \to \infty} S_n$, it is *Césaro summable* to $s \in \mathbb{C}$ if $\lim_{n \to \infty} A_n = s$, and it is *Abel summable* to $s \in \mathbb{C}$ if $s = \lim_{r \nearrow 1} A(r)$

Here we will show that

summable to $s \implies$ **Césaro summable** to $s \implies$

$$\lim_{m \to \infty} \frac{a_m}{m} = 0 \text{ and } \textbf{Abel summable} \text{ to } s.$$

The Exercise 4.1 below outlines a proof that neither implication can be reversed.

The first implication is trivial. To prove the second, assume Césaro summability, and note that

$$\frac{a_n}{n} = A_n - A_{n-1} + \frac{A_{n-1}}{n} \longrightarrow 0.$$

Next, write

$$a_m = \begin{cases} A_1 & \text{if } m = 1 \\ 2A_2 - A_1 & \text{if } m = 2 \\ mA_m - 2(m-1)A_{m-1} + (m-2)A_{m-2} & \text{if } m \geq 3, \end{cases}$$

and therefore

$$A(r) = \sum_{m=1}^\infty mr^{m-1}A_m - 2\sum_{m=2}^\infty (m-1)r^{m-1}A_{m-1} + \sum_{m=3}^\infty (m-2)r^{m-1}A_{m-2}$$

$$= \sum_{m=1}^\infty (r^{m-1} - 2r^m + r^{m+1})mA_m = (1-r)^2 \sum_{m=1}^\infty mr^{m-1}A_m.$$

Now observe that

$$\sum_{m=1}^n mr^{m-1} = \partial_r \sum_{m=0}^n r^m = \partial_r \frac{1-r^n}{1-r} = \frac{1-r^n - n(1-r)r^{n-1}}{(1-r)^2}.$$

Hence,

$$(1-r)^2 \sum_1^n mr^{m-1} \leq 1 - r^n \text{ and } (1-r)^2 \sum_1^\infty mr^{m-1} = 1.$$

Assume that $A_n \longrightarrow s$, and, given $\epsilon > 0$, choose n so that $|A_m - s| \leq \epsilon$ for $m > n$. Then

$$\left| A(r) - s \right| = (1-r)^2 \left| \sum_{m=1}^{\infty} m r^{m-1} (A_m - s) \right| \leq (1-r)^2 \sum_{m=1}^{n} m r^{m-1} |A_m - s| + \epsilon$$

$$\leq (1 - r^n) \max_{1 \leq m \leq n} |A_m - s| + \epsilon,$$

and therefore $\overline{\lim}_{r \nearrow 1} |A(r) - s| \leq \epsilon$.

Exercise 4.1. Show that
(i) the series for $\{(-1)^{m-1} : m \geq 1\}$ is Césaro summable to $\frac{1}{2}$ but not summable,
(ii) the series for $\{(-1)^{m-1} m : m \geq 1\}$ is Abel summable to $\frac{1}{4}$ but not Césaro summable. In fact, show that $A_{2n} = 0$ and $A_{2n+1} = \frac{n+1}{2n+1} \longrightarrow \frac{1}{2}$.

5. SOME REFINEMENTS

In this section we will apply the notions of summability discussed in the previous section to Fourier series. Observe that we have already considered Abel summability in §1.

To examine further when the series is summable, introduce the function

$$D_n(x) = \sum_{|m| \leq n} \mathfrak{e}_m(x) \text{ for } x \in \mathbb{R}.$$

Then D_n, which is often called the *Dirichlet kernel*, is an even, periodic function with period 1, $\int_0^1 D_n(x)\,dx = 1$, and $S_n \varphi = D_n * \varphi$. In addition

$$D_n(x) = \mathfrak{e}_{-n}(x) \sum_{m=0}^{2n} \mathfrak{e}_m(x) = \mathfrak{e}_{-n}(x) \frac{1 - \mathfrak{e}_{2n+1}(x)}{1 - \mathfrak{e}_1(x)} = \frac{e^{-\imath \pi (2n+1)x} - e^{\imath \pi (2n+1)x}}{e^{-\imath \pi x} - e^{\imath \pi x}}$$

$$= \frac{\sin \pi (2n+1)x}{\sin \pi x}.$$

Hence,

$$S_n \varphi(x) - \varphi(x) = \int_{[0,1]} \frac{\tilde{\varphi}(x+y) - \varphi(x)}{\sin \pi y} \sin \pi (2n+1) y \, dy.$$

Now suppose that φ is an $\mathbb{R}$-valued function for which $\varphi(0) = \varphi(1)$, and assume that $\varphi \in C^\alpha\big([0,1];\mathbb{C}\big)$ [4] is Hölder continuous of order $\alpha \in (0,1)$. Set

$$\psi(y) = e^{\imath \pi y} \frac{\tilde{\varphi}(x+y) - \varphi(x)}{\sin \pi y}.$$

Then $\psi \in L^1(\lambda_{[0,1)}; \mathbb{C})$ and $S_n \varphi(x) - \varphi(x)$ is the imaginary part of

$$\int_{[0,1]} \psi(y) \mathfrak{e}_{-2n+1}(y) \, dt = \big(\psi, \mathfrak{e}_{2n-1} \big)_{L^2(\lambda_{[0,1)};\mathbb{C})},$$

and so, by the Riemann–Lebesgue lemma (cf. Exercise 1.1), $S_n \varphi(x) \longrightarrow \varphi(x)$ as $n \to \infty$. The preceding shows that if $\varphi \in C^\alpha\big([0,1];\mathbb{C}\big)$ satisfies $\varphi(0) = \varphi(1)$, then $S_n \varphi \longrightarrow \varphi$ pointwise, but it does not provide a rate of convergence or even say if the convergence is uniform.

---

[4] $C^\alpha(E;\mathbb{C})$ space of $\mathbb{C}$-valued functions on a metric space E which are uniformly Hölder continuous of order $\alpha \in (0,1)$.

Césaro summability of Fourier series was initiated by Fejér. Obviously,

$$\frac{1}{n} \sum_{m=0}^{n-1} S_m \varphi = F_n * \varphi,$$

where

$$F_n(x) \equiv \frac{1}{n} \sum_{m=0}^{n-1} D_n(x).$$

The function F_n is called the *Fejér kernel*, and it is clear that F_n is a continuous, even function of period 1 for which $\int_{[0,1]} F_n(x)\, dx = 1$. In addition, $nF_n(x)\sin \pi x$ is the imaginary part of

$$e^{\imath \pi x} \sum_{m=0}^{n-1} \mathfrak{e}_{2m}(x) = e^{\imath \pi x} \frac{1 - e^{\imath \pi 2nx}}{1 - e^{\imath 2\pi x}} = \frac{\imath(1 - e^{\imath 2\pi nx})}{2\sin \pi x},$$

and so

(5.1) $$F_n(x) = \frac{1 - \cos 2\pi nx}{2n \sin^2 \pi x} = \frac{1}{n}\left(\frac{\sin \pi nx}{\sin \pi x}\right)^2.$$

Proceeding as in the proof of Theorem 1.1, one sees that

$$F_n * \varphi(x) - \varphi(x) = \int_{[0,1]} F_n(y)\big(\tilde{\varphi}(x+y) - \varphi(x)\big)\, dx \longrightarrow 0$$

uniformly if φ is continuous and satisfies $\varphi(1) = \varphi(0)$. Equivalently,

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{m=0}^{n-1} S_m \varphi - \varphi \right\|_{\mathrm{u}} = 0.$$

It turns out that one can do much better.

Theorem 5.1. *Let* $\varphi : \left[-\frac{1}{2}, \frac{1}{2}\right] \longrightarrow \mathbb{C}$ *be a measurable function, let* $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$, *and assume that there is a* $C \in (0, \infty)$ *and* $\alpha \in (0, 1]$ *such that* $|\tilde{\varphi}(x+y) - \varphi(x)| \leq C|y|^\alpha$ *for* $y \in \left[-\frac{1}{2}, \frac{1}{2}\right]$. *For* $n \geq 5$

(5.2) $$\left|F_n * \varphi(x) - \varphi(x)\right| \leq C \begin{cases} \frac{2}{(1+\alpha)n^\alpha} + \frac{4(n^{1-\alpha}-4^{1-\alpha})}{\pi^2(1-\alpha)n} + \frac{1-2^{-(1+\alpha)}}{2^\alpha(1+\alpha)n} & \text{if } \alpha \in (0,1) \\ \frac{19}{16n} + \frac{4\log\frac{n}{4}}{\pi^2 n(1-\alpha)} & \text{if } \alpha = 1. \end{cases}$$

Hence

$$\overline{\lim_{n\to\infty}}\, n^\alpha |F_n * \varphi(x) - \varphi(x)| \leq \frac{2}{1+\alpha} + \frac{4}{\pi^2(1-\alpha)} \quad \text{if } \alpha \in (0,1)$$

and

$$\overline{\lim_{n\to\infty}}\, \frac{n}{\log n} |F_n * \varphi(x) - \varphi(x)| \leq \frac{4}{\pi^2} \quad \text{if } \alpha = 1.$$

Proof. Without loss in generality, I will assume that $C = 1$.

The proof turns on the estimates

(5.3) $$F_n(y) \leq \begin{cases} n & \text{for all } y \in \left[-\frac{1}{2}, \frac{1}{2}\right] \\ \frac{2}{\pi^2 ny^2} & \text{when } |y| \in \left(0, \frac{1}{4}\right] \\ \frac{2}{n} & \text{when } |y| \in \left[\frac{1}{4}, \frac{1}{2}\right] \end{cases}$$

That $F_n(y) \leq n$ is clear from the fact that $\|D_m\|_u \leq 1$ and therefore that $nF_n(y) \leq 2\sum_{m=1}^{n-1} m + n = n^2$. To see second inequality, note that $\cos \pi t \geq 2^{-\frac{1}{2}}$ when $|y| \in \left(0, \frac{1}{4}\right]$ and therefore that

$$|\sin \pi y| = \int_0^{\pi|y|} \cos t \, dt \geq 2^{-\frac{1}{2}} \pi |y|.$$

As for $F_n(y) \leq \frac{2}{n}$ when $|y| \in \left[\frac{1}{4}, \frac{1}{2}\right]$, simply remember that $|\sin \pi y| \geq 2^{-\frac{1}{2}}$ for such y's.

Assume that $\alpha \in (0, 1)$. Because $\int_{-\frac{1}{2}}^{\frac{1}{2}} F_n(y) = 1$

$$\left| F_n * \varphi(x) - \varphi(x) \right| \leq \int_{-\frac{1}{2}}^{\frac{1}{2}} F_n(y) \left| \tilde{\varphi}(x+y) - \varphi(x) \right| dy$$

$$\leq n \int_0^{\frac{1}{n}} |y|^\alpha \, dy + \frac{2}{\pi^2 n} \int_{\frac{1}{n}}^{\frac{1}{4}} |y|^{\alpha-2} \, dy + \frac{2}{n} \int_{\frac{1}{4} \leq |y| \leq \frac{1}{2}} |y|^\alpha \, dy$$

$$\leq \frac{2}{(1+\alpha)n^\alpha} + \frac{4(n^{1-\alpha} - 4^{1-\alpha})}{\pi^2(1-\alpha)n} + \frac{1 - 2^{-(1+\alpha)}}{2^\alpha(1+\alpha)n}.$$

If $\alpha = 1$, the top line in (5.2) holds for all $\alpha \in (0, 1)$, and therefore one need only examine what happens as $\alpha \nearrow 1$. Clearly $\frac{2}{(1+\alpha)n^\alpha} \searrow \frac{1}{n}$ and $\frac{1-2^{-(1+\alpha)}}{2^\alpha(1+\alpha)n} \searrow \frac{3}{16n}$ as $\alpha \nearrow 1$. To handle the remaining term, note that it can be written as

$$\frac{4^{2-\alpha}}{\pi^2 n} \frac{\left(\frac{n}{4}\right)^{1-\alpha} - 1}{1 - \alpha}$$

which decreases to $\frac{4 \log \frac{n}{4}}{\pi^2 n}$ as $\alpha \nearrow 1$. $\qquad \square$

One could of course have derived the estimate when $\alpha = 1$ directly by the same argument as was used when $\alpha < 1$. However, the derivation given has the advantage that it shows the estimates get stronger for all $n \geq 5$, not just asymptotically, as α increases.

Obviously, results like those in Theorem 5.1 turn on the continuity properties of φ, properties that a generic element of $L^1(\lambda_{[0,1)}; \mathbb{C})$ will not possess. Nonetheless, Lebesgue showed that every locally $\lambda_\mathbb{R}$-integrable φ does have a continuity property at almost everywhere point. Namely, he showed that

$$\lim_{r \searrow 0} \frac{1}{r} \int_0^r |\tilde{\varphi}(x \pm t) - \varphi(x)| \, dt = 0 \quad \text{for } \lambda_\mathbb{R}\text{-almost every } x \in \mathbb{R},$$

and he used this fact to prove the following theorem.

Theorem 5.2. *If $\varphi \in L^1\left(\lambda_{[-\frac{1}{2}, \frac{1}{2}]}; \mathbb{C}\right)$, then*

$$\lim_{n \to \infty} F_n * \varphi(x) = \varphi(x) \text{ for } \lambda_{[-\frac{1}{2}, \frac{1}{2}]}\text{-almost every } x \in [0, 1].$$

Proof. Set $\varphi_x(y) = |\tilde{\varphi}(x+y) - \varphi(x)|$ and

$$\Phi_x(y) = \frac{1}{|y|} \int_0^{|y|} \varphi_x(\text{sgn}(y)t) \, dt.$$

By Lebesgue's theorem, $\lim_{|y| \searrow 0} \Phi_x(y) = 0$ for $\lambda_{[-\frac{1}{2}, \frac{1}{2}]}$-almost every $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$.

Let x be such a point. Then

$$\left|F_n * \varphi(x) - \varphi(x)\right| \le \int_{-\frac{1}{2}}^{0} F_n(y)\varphi_x(y)\,dy + \int_{0}^{\frac{1}{2}} F_n(y)\varphi_x(y)\,dy.$$

We will show only that $\lim_{n\to\infty}\int_0^{\frac{1}{2}} F_n(y)\varphi_x(y)\,dy = 0$ because the proof that $\lim_{n\to\infty}\int_{-\frac{1}{2}}^{0} F_n(y)\varphi_x(y)\,dy = 0$ is essentially the same.

Using our estimates for F_n in (5.3), one has

$$\int_0^{\frac{1}{2}} F_n(y)\varphi_x(y)\,dy = \int_0^{\frac{1}{n}} F_n(y)\varphi_x(y)\,dy + \int_{\frac{1}{n}}^{\frac{1}{2}} F_n(y)\varphi_x(y)\,dy$$

$$\le n\int_0^{\frac{1}{n}} \varphi_x(y)\,dy + \frac{2}{n}\int_{\frac{1}{n}}^{\frac{1}{2}} \frac{\varphi_x(y)}{y^2}\,dy.$$

Since

$$n\int_0^{\frac{1}{n}} \varphi_x(y)\,dy = \Phi_x\!\left(\tfrac{1}{n}\right),$$

the first term tends to 0. As for the second, use integration by parts to see that it is dominated by

$$\frac{4\Phi_x\!\left(\frac{1}{2}\right)}{n} + \frac{4}{n}\int_{\frac{1}{n}}^{\frac{1}{2}} \frac{\Phi_x(y)}{y^2}\,dy.$$

Finally, given $\epsilon > 0$, choose $\delta \in \left(0, \frac{1}{2}\right)$ so that $\Phi_x(y) \le \epsilon$ for $0 \le y \le \delta$. Then, for $n > \frac{1}{\delta}$,

$$\frac{1}{n}\int_{\frac{1}{n}}^{\frac{1}{2}} \frac{\Phi_x(y)}{y^2}\,dy \le \frac{\epsilon}{n}\int_{\frac{1}{n}}^{\delta} \frac{1}{y^2}\,dy + \frac{1}{n}\int_{\delta}^{\frac{1}{2}} \frac{\Phi_x(y)}{y^2}\,dy \le 2\epsilon + \frac{\|\Phi_x\|_u}{\delta n},$$

and so

$$\varlimsup_{n\to\infty}\int_0^{\frac{1}{2}} F_n(y)\varphi_x(y)\,dy \le 4\epsilon.$$

$\square$

Theorem 5.2 is a stark contrast to a famous example produced in 1926 by Kolmogorov[5] of a function in $L^1\!\left(\lambda_{[-\frac{1}{2},\frac{1}{2}]}; \mathbb{C}\right)$ for which $\{S_n\varphi(x) : n \ge 0\}$ diverges at *every* x. It is also interesting to compare it to more recent results by L. Carleson and R. Hunt. Namely, Carleson showed that $S_n\varphi \longrightarrow \varphi$ $\left(\text{a.e.}, \lambda_{[-\frac{1}{2},\frac{1}{2}]}\right)$ if $\varphi \in L^2\!\left(\lambda_{[-\frac{1}{2},\frac{1}{2}]}; \mathbb{C}\right)$, and Hunt showed that the same is true for $\varphi \in L^p\!\left(\lambda_{[-\frac{1}{2},\frac{1}{2}]}; \mathbb{C}\right)$ for $p \in (1, \infty)$.

Exercise 5.1. Show that

$$\varliminf_{n\to\infty} n^\alpha \int_{-\frac{1}{2}}^{\frac{1}{2}} F_n(y)|y|^\alpha\,dy > 0 \text{ for } \alpha \in (0,1)$$

and that

$$\varliminf_{n\to\infty} \frac{n}{\log n}\int_{-\frac{1}{2}}^{\frac{1}{2}} F_n(y)|y|\,dy > 0.$$

[5] A.N. Kolmogorov, *Une série de Fourier-Lebesgue divergente partout*, C.R. 183 (1926), pp. 1327-1328.

Hence the rates given Theorem 5.1 are optimal.

Hint: If $0 \leq m \leq n - 1$, show that

$$F_n(y) \geq \frac{1}{2\pi^2 ny^2} \text{ if } \frac{4m+1}{n4} \leq y \leq \frac{2m+1}{2n}.$$

6. The L^1 Fourier Transform

By an easy rescaling argument, one knows that, for any $L \in \mathbb{Z}^+$ and $f \in C^1([-L, L]; \mathbb{C})$ satisfying $f(-L) = f(L)$,

$$f(x) = \frac{1}{2L} \sum_{m \in \mathbb{Z}} \int_{-L}^{L} e^{i\frac{2\pi m(y-x)}{2L}} f(y) \, dy = \lim_{R \to \infty} \int_{-L}^{L} \left(\frac{1}{2L} \sum_{|m| \leq R} e^{i\frac{2\pi m(y-x)}{2L}} \right) f(y) \, dx.$$

Now suppose that $f \in C_c^1(\mathbb{R}; \mathbb{C})$. Then

$$f(x) = \lim_{L \to \infty} \lim_{R \to \infty} \int_{-L}^{L} \left(\frac{1}{2L} \sum_{|m| \leq R} e^{i\frac{2\pi m(y-x)}{2L}} \right) f(y) \, dy.$$

Thus, if one can justify reversing the order in which the limits are taken, one would have that

$$f(x) = \lim_{R \to \infty} \int \left(\int_{-R}^{R} e^{i\xi 2\pi(x-y)} \, d\xi \right) f(y) \, dy$$

$$= \lim_{R \to \infty} \frac{1}{2\pi} \int_{-2\pi R}^{2\pi R} e^{-i\xi x} \left(\int e^{i\xi y} f(y) \, dy \right) d\xi.$$

In other words, there is reason to hope that, under suitable conditions on f,

(6.1) $$f(x) = \frac{1}{2\pi} \int e^{-i\xi x} \hat{f}(\xi) \, d\xi \text{ where } \hat{f}(\xi) \equiv \int e^{i\xi y} f(y) \, dy.$$

The function $\hat{f}$ is called the *Fourier transform* of f, and our primary goal here will be to find out in what sense (6.1) is true, first when $f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ and then when $f \in L^2(\lambda_{\mathbb{R}}; \mathbb{C})$. However, we will begin with some computations involving $\hat{f}$ that don't require our knowing when (6.1) holds.

7. Computations and Applications of L^1 Fourier Transforms

If $f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$, then it is clear that $\hat{f}$ is continuous and that

(7.1) $$\|\hat{f}\|_u \leq \|f\|_{L^1(\lambda_{\mathbb{R}}; \mathbb{C})}.$$

Lemma 7.1. *If $f \in C^1(\mathbb{R}, \mathbb{C}) \cap L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ and $f' \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$, then*

(7.2) $$\widehat{f'}(\xi) = -i\xi \hat{f}(\xi).$$

Proof. If f has compact support, then (7.2) is an easy application of integration by parts. To prove it under the given conditions, choose a function $\eta \in C^\infty(\mathbb{R}; [0, 1])$ for which $\eta(y) = 1$ when $y \in [-1, 1]$ and $\eta(y) = 0$ when $y \notin [-2, 2]$, and set $f_n(y) = \eta\left(\frac{y}{n}\right) f(y)$. Then $f_n \longrightarrow f$ and $f_n' \longrightarrow f'$ in $L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ and so

$$\widehat{f'}(\xi) = \lim_{n \to \infty} \widehat{f_n'}(\xi) = -i\xi \lim_{n \to \infty} \widehat{f_n}(\xi) = -i\xi \hat{f}(\xi).$$

$\square$

As a consequence of Lemma 7.1, it is easy to prove the *Riemann-Lebesgue lemma* in this context. Namely, (7.2) makes it clear for compactly support $f \in C^1(\mathbb{R}; \mathbb{C})$, and (7.1) makes it clear that the set of f's for which it is holds is closed in $L^1(\lambda_{\mathbb{R}}; \mathbb{C})$.

We next turn to the computation of $\hat{f}$ in two important cases.

Set $g_t(x) = (2\pi t)^{-\frac{1}{2}} e^{-\frac{x^2}{2}}$ for $(t, x) \in (0, \infty) \times \mathbb{R}$, and check that $\partial_t g_t(x) = \frac{1}{2} \partial_x^2 g_t(x)$. Hence, for any $\zeta \in \mathbb{C}$, integration by parts leads to

$$\partial_t \int e^{\zeta x} g_t(x) \, dx = \frac{1}{2} \int e^{\zeta x} \partial_x^2 g_t(x) \, dx = \frac{\zeta^2}{2} \int e^{\zeta x} g_t(x) \, dx.$$

Since

$$\int e^{\zeta x} g_t(x) \, dx = \int e^{t^{\frac{1}{2}} \zeta x} g_1(x) \, dx \longrightarrow 1$$

as $t \searrow 0$,

$$\int e^{\zeta x} g_t(x) \, dx = e^{\frac{t\zeta^2}{2}}.$$

In particular

(7.3) $$\widehat{g_t}(\xi) = e^{-\frac{\xi^2}{2}}$$

Equivalently, $\widehat{g_t} = \left(\frac{2\pi}{t}\right)^{\frac{1}{2}} g_{\frac{1}{t}}$ and so

(7.4) $$\left(\widehat{g_t}\right)^{\wedge} = 2\pi g_t.$$

Set $p_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2}$ for $(y, x) \in (0, \infty) \times \mathbb{R}$, and note that

$$\int p_y(x) \, dx = \int p_1(x) \, dx = 1 \text{ for all } y > 0.$$

In addition, because $p_y(x)$ is the real part of $\frac{\imath}{\pi z}$ with $z = x + \imath y$, $(x, y) \rightsquigarrow p_y(x)$ is harmonic. Thus, $\partial_x^2 p_y = -\partial_y^2 p_y$, and so, by (7.2),

$$\partial_y^2 \widehat{p_y}(\xi) = \xi^2 \widehat{p_y}(\xi).$$

Thus, for each ξ,

$$\widehat{p_y}(\xi) = a(\xi) e^{y\xi} + b(\xi) e^{-y\xi},$$

where, since $\widehat{p_y}(0) = 1$, $a(\xi) + b(\xi) = 1$. Because $|\widehat{p_y}(\xi)| \leq 1$, $\xi \geq 0 \implies a(\xi) = 0$ & $b(\xi) = 1$ and $\xi < 0 \implies a(\xi) = 1$ & $b(\xi) = 0$. Hence

(7.5) $$\widehat{p_y}(\xi) = e^{-y|\xi|}.$$

Here is an interesting application of equations (7.3) and (7.5). Since

$$\frac{1}{\xi^2 + y^2} = \int_0^\infty e^{-t(\xi^2 + y^2)} \, dx = \int_0^\infty e^{-ty^2} \widehat{g_{2t}}(\xi) \, dt$$

and $\left(\widehat{g_{2t}}\right)^{\wedge} = 2\pi g_{2t}$,

$$\frac{\pi}{y} e^{-y|x|} = 2\pi \int_0^\infty e^{-ty^2} g_{2t}(x) \, dt = \pi^{\frac{1}{2}} \int_0^\infty t^{-\frac{1}{2}} e^{-ty^2} e^{-\frac{x^2}{4t}} \, dt.$$

Thus, for $x, y \in (0, \infty)$,

(7.6) $$\int_0^\infty t^{-\frac{1}{2}} e^{-ty^2} e^{-\frac{x^2}{t}} \, dt = \frac{\pi^{\frac{1}{2}} e^{-2yx}}{y},$$

a computation which can also be done using a somewhat tricky change of variables.

Theorem 7.2. (Poisson Sum) *Let $f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C}) \cap C(\mathbb{R}; \mathbb{C})$, and assume that*

$$\sum_{n \in \mathbb{Z}} \left(\sup_{x \in [0,1]} |f(x+n)| + |\hat{f}(2\pi n)| \right) < \infty.$$

Then

$$\tag{7.7} \sum_{n \in \mathbb{Z}} f(n) = \sum_{n \in \mathbb{Z}} \hat{f}(2\pi n).$$

Proof. Define $\tilde{f}(x) = \sum_{n \in \mathbb{Z}} f(x+n)$. Then $\tilde{f}$ is a continuous periodic function with period 1, and

$$\left(\tilde{f}, \mathfrak{e}_m \right)_{L^2(\lambda_{[0,1]}; \mathbb{C})} = \sum_{n \in \mathbb{Z}} \int_0^1 e^{-\imath 2\pi m x} f(x+n) \, dx = \int e^{-\imath 2\pi m x} f(x) \, dx = \hat{f}(-2\pi m).$$

Thus, $\sum_{m \in \mathbb{Z}} \left| (\tilde{f}, \mathfrak{e}_m)_{L^2(\lambda_{[0,1]}; \mathbb{C})} \right| < \infty$, and therefore

$$\tilde{f}(x) = \sum_{m \in \mathbb{Z}} \hat{f}(-2\pi m) \mathfrak{e}_m(x) = \sum_{m \in \mathbb{Z}} \hat{f}(2\pi m) \mathfrak{e}_{-m}(x),$$

where the convergence of the series is absolute and uniform. By taking $x = 0$, (7.7) follows. $\square$

Equation (7.7) is known as the *Poisson summation formula*. Among its many applications is the following.

When $f = p_y$, (7.7) says that

$$\frac{y}{\pi} \sum_{n \in \mathbb{Z}} \frac{1}{y^2 + n^2} = \sum_{n \in \mathbb{Z}} e^{-2\pi y |n|} = \frac{1 + e^{-2\pi y}}{1 - e^{-2\pi y}} = \coth \pi y,$$

and so

$$\tag{7.8} \sum_{n \in \mathbb{Z}} \frac{1}{y^2 + n^2} = \frac{\pi \coth \pi y}{y}$$

for $y > 0$.

A famous application of (7.8) is *Euler's product formula*:

$$\tag{7.9} \sin \pi x = \pi x \prod_{m=1}^{\infty} \left(1 - \frac{x^2}{m^2} \right).$$

To prove it, first observe that

$$\frac{1}{x^2 + m^2} = \frac{1}{2x} \partial_x \log(x^2 + m^2) = \frac{1}{2x} \partial_x \log \left(1 + \frac{x^2}{m^2} \right) \text{ for } m \neq 0$$

and that $\pi \coth \pi y = \partial_y \log(\sinh \pi y)$. Hence, by (7.8)

$$\frac{1}{x} \partial_x \log \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2} \right) + \frac{1}{x^2} = \frac{1}{x} \partial_x \log(\sinh \pi x),$$

which means that

$$\partial_x \log \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2} \right) = \partial_x \log(x^{-1} \sinh \pi x).$$

Integrating both sides from 0 to x, one gets

$$\log x \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2}\right) = \log(\sinh \pi x) - \log \pi = \log \frac{\sinh \pi x}{\pi x},$$

which means that

$$(7.10) \qquad\qquad \sinh \pi x = \pi x \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{n^2}\right)$$

from which (7.9) follows by analytic continuation.

Another application of (7.5) is a proof[6] that

$$(7.11) \qquad\qquad \lim_{R \to \infty} \int_{-R}^{R} \frac{\sin \xi x}{x} \, dx = \operatorname{sgn}(\xi)\pi \quad \text{for } \xi \neq 0.$$

We begin with the more or less trivial observation that

$$\int_{-R}^{R} \frac{\sin \xi x}{x} \, dx = \operatorname{sgn}(\xi) \int_{-R}^{R} \frac{\sin |\xi| x}{x} \, dx = \operatorname{sgn}(\xi) \int_{-|\xi|R}^{|\xi|R} \frac{\sin x}{x} \, dx.$$

Thus, what we have to show is that

$$\lim_{R \to \infty} \int_{-R}^{R} \frac{\sin x}{x} \, dx = \pi. \tag{$*$}$$

The first step in the proof $(*)$, is to show that if

$$g_R(\xi, y) \equiv \int_{-R}^{R} \frac{x \sin \xi x}{x^2 + y^2} \, dx \longrightarrow \pi e^{-y\xi} \quad \text{for } \xi > 0, \tag{$**$}$$

then $(*)$ holds. Indeed,

$$\left| \int_{-R}^{R} \frac{\sin \xi x}{x} \, dx - g_R(\xi, y) \right| \leq 2y^2 \left| \int_{0}^{\infty} \frac{|\sin \xi x|}{x(x^2 + y^2)} \, dx \right| \leq \xi \pi y,$$

and so $(**)$ implies $(*)$.

The next step is to show that for each $y > 0$ there exists a continuous $\xi \in (0, \infty) \longmapsto g(\xi, y) \in \mathbb{C}$ such that $g_R(\xi, y) \longrightarrow g(\xi, y)$ uniformly for ξ compact subsets of $(0, \infty)$. To this end, note that

$$g_R(\xi, y) = 2 \int_{0}^{R} \frac{x \sin \xi x}{x^2 + y^2} \, dx = \frac{2}{\xi} \left(-\frac{R \cos \xi R}{R^2 + y^2} + 2 \int_{0}^{R} \frac{(y^2 - x^2) \cos \xi x}{(x^2 + y^2)^2} \, dx \right)$$

$$\longrightarrow \frac{4}{\xi} \int_{0}^{\infty} \frac{(y^2 - x^2) \cos \xi x}{(x^2 + y^2)^2} \, dx$$

uniformly for ξ in compacts subsets of $(0, \infty)$.

The final step is the identify $g(\xi, y)$ as $\pi e^{-y\xi}$. For this purpose, observe that

$$g_R(\xi, y) = -\imath \int_{-R}^{R} \frac{x e^{\imath \xi x}}{x^2 + y^2} \, dx = \partial_\xi f_R(\xi, y)$$

where

$$f_R(\xi, y) = -\frac{\pi}{y} \int_{-R}^{R} p_y(x) e^{\imath \xi x} \, dx \longrightarrow -\frac{\pi}{y} e^{-y\xi}.$$

---

[6]The most commonly given proof is based on contour integration and Cauchy's theorem.

Hence

$$f_R(\eta) - f_R(\xi) = \int_\xi^\eta g_R(t, y)\, dt,$$

and therefore

$$\frac{\pi}{y}\left(e^{-y\xi} - e^{-y\eta}\right) = \int_\xi^\eta g(t, y)\, dt,$$

from which $g(\xi, y) = \pi e^{-y\xi}$ follows easily.

Exercise 7.1. Show that if $f \in L^1(\lambda_\mathbb{R}; \mathbb{C})$ and $f_t(x) = t^{-1} f(t^{-1}x)$, then $\hat{f}_t(\xi) = \hat{f}(t\xi)$.

Exercise 7.2. Show that if $f \in C^2(\mathbb{R}; \mathbb{C}) \cap L^1(\lambda_\mathbb{R}; \mathbb{C})$ and both f' and f'' are in $L^1(\lambda_\mathbb{R}; \mathbb{C})$, then $\hat{f} \in L^1(\lambda_\mathbb{R}; \mathbb{C})$.

Exercise 7.3. Using $\cosh t = 1 + \frac{t^2}{2} + O(t^4)$ and $\sinh t = t + \frac{t^3}{6} + O(t^5)$, prove from (7.8) that $\sum_{n=1}^\infty \frac{1}{n^2} = \frac{\pi^2}{6}$.

Exercise 7.4. Show that

$$\sum_{n \in \mathbb{Z}} e^{-\frac{\pi n^2}{t}} = t^{\frac{1}{2}} \sum_{n \in \mathbb{Z}} e^{-\pi t n^2},$$

a formula that plays an important role in the theory of Theta functions.

8. The L^1 Fourier Inversion Formula

Here we will see to what extent (6.1) can be justified, and the idea is to use the fact that we already know (cf. (7.4)) that it holds for g_t. With this in mind, we have, by Fubini's theorem,

$$2\pi g_t * f(x) = 2\pi \int g_t(y) f(x - y)\, dy = 2\pi \int g_t(y) f(x + y)\, dy$$

$$= \int (\hat{g}_t)^\wedge(y) f(x + y)\, dy = \int \widehat{g_t}(\xi) \left(\int e^{\imath \xi y} f(x + y)\, dy \right) d\xi = \int e^{-\frac{t\xi^2}{2}} e^{-\imath \xi x} \hat{f}(\xi)\, d\xi,$$

and so

$$g_t * f(x) = \frac{1}{2\pi} \int e^{-\frac{t\xi^2}{2}} e^{-\imath \xi x} \hat{f}(\xi)\, d\xi.$$

Let $f \in L^1(\lambda_\mathbb{R}; \mathbb{C})$. If f is continuous at x, then $\lim_{t \searrow 0} g_t * f(x) = f(x)$, and so

$$(8.1) \qquad f(x) = \frac{1}{2\pi} \lim_{t \searrow 0} \int e^{-\frac{t\xi^2}{2}} e^{-\imath \xi x} \hat{f}(\xi)\, d\xi \quad \text{if } f \text{ is continuous at } x.$$

In particular

$$(8.2) \qquad f(x) = \frac{1}{2\pi} \int e^{-\imath \xi x} \hat{f}(\xi)\, d\xi \quad \text{if } \hat{f} \in L^1(\lambda_\mathbb{R}; \mathbb{C}).$$

More generally, for any $f \in L^1(\lambda_\mathbb{R}; \mathbb{C})$, $g_t * f \longrightarrow f$ in $L^1(\lambda_\mathbb{R}; \mathbb{C})$, and so

$$(8.3) \qquad \frac{1}{2\pi} \int e^{-\frac{t\xi^2}{2}} e^{-\imath \xi x} \hat{f}(\xi)\, d\xi \longrightarrow f(x) \text{ in } L^1(\lambda_\mathbb{R}; \mathbb{C}),$$

which can be thought of the Abel version of (6.1). As an immediate consequence, we know that $f = \mathbf{0} \iff \hat{f} = \mathbf{0}$.

Exercise 8.1. Show that if $f \in C^2(\mathbb{R}; \mathbb{C}) \cap L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ and both f' and f'' are in $L^1(\lambda_{\mathbb{R}}; \mathbb{C})$, then $\hat{f} \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ and therefore $f = (2\pi)^{-1} \int e^{-\imath \xi x} \hat{f}(\xi) \, d\xi$.

Exercise 8.2. Using Exercise 8.1, give another proof that $\widehat{p_t}(\xi) = e^{-t|\xi|}$.

Exercise 8.3. There is nothing sacrosanct about g_t in producing formulas like (8.1) and (8.3). Indeed, give a $\rho \in C(\mathbb{R}, [0, \infty))$ for which $\int \rho(x) \, dx = 1$, set $\rho_t(x) = t^{-1} \rho(t^{-1} x)$. Then it is well known that, as $t \searrow 0$, $\rho_t * f(x) \longrightarrow f(x)$ if $f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ is continuous at x and that $\rho_t * f \longrightarrow f$ in $L^1(\lambda_{\mathbb{R}}; \mathbb{C})$ for any $f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C})$. Now suppose that $\rho \in C^2(\mathbb{R}, [0, \infty))$ and that ρ' and ρ'' are in $L^1(\lambda_{\mathbb{R}}; \mathbb{C})$, and show that

$$\frac{1}{2\pi} \int e^{-\imath \xi x} \hat{\rho}(t\xi) \hat{f}(\xi) \, d\xi \longrightarrow f(x) \quad \text{if } f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C}) \text{ is continuous at } x$$

and

$$\frac{1}{2\pi} \int \hat{e}^{-\imath \xi x} \hat{\rho}(t\xi) \hat{f}(\xi) \, d\xi \longrightarrow f(x) \text{ in } L^1(\lambda_{\mathbb{R}}; \mathbb{C}) \text{ for any } f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C}).$$

9. The Ornstein–Uhlenbeck Semigroup

Set $g_t(x) = (2\pi t)^{-\frac{1}{2}} e^{-\frac{x^2}{2t}}$, and note that

$$(9.1) \qquad \int g_s(x - \xi) g_t(\xi - y) \, d\xi = g_{s+t}(y - x) \text{ and } \partial_t g_t(x) = \tfrac{1}{2} \partial_x^2 g_t(x).$$

For $(t, x, y) \in (0, \infty) \times \mathbb{R} \times \mathbb{R}$, define

$$p(t, x, y) = g_{1 - e^{-t}}\left(y - e^{-\frac{t}{2}} x\right)$$

$$(9.2) \qquad = \left(2\pi(1 - e^{-t})\right)^{-\frac{1}{2}} \exp\left(-\frac{(y - e^{-\frac{t}{2}} x)^2}{2(1 - e^{-t})}\right) = e^{\frac{t}{2}} g_{e^t - 1}\left(x - e^{\frac{t}{2}} y\right).$$

From the first part of (9.1) and the third equality in (9.2), we see that

$$\int p(s, x, \xi) p(t, \xi, y) \, d\xi = e^{\frac{t}{2}} \int g_{1 - e^{-s}}\left(\xi - e^{-\frac{s}{2}} y\right) g_{e^t - 1}\left(\xi - e^{\frac{t}{2}} x\right) d\xi$$

$$= e^{\frac{t}{2}} g_{e^t - e^{-s}}\left(e^{\frac{t}{2}} y - e^{-\frac{s}{2}} x\right) = p(s + t, x, y).$$

Hence $p(t, x, y)$ satisfies the *Chapman–Kolmogorov equation*

$$(9.3) \qquad p(s + t, x, y) = \int p(s, x, \xi) p(t, \xi, y) \, d\xi.$$

In addition, using the second part of (9.1), one sees that

$$(9.4) \qquad \partial_t p(t, x, y) = \mathcal{L}_x p(t, x, y) \text{ where } \mathcal{L}_x = \tfrac{1}{2}\left(\partial_x^2 - x \partial_x\right).$$

Next define

$$(9.5) \qquad P_t \varphi(x) = \int \varphi(y) p(t, x, y) \, dy$$

for $\varphi \in C(\mathbb{R}; \mathbb{C})$ with at most exponential growth at ∞, and use (9.3) to see that $\{P_t : t > 0\}$ is a semigroup (i.e., $P_{s+t} = P_s \circ P_t$). In addition, use (9.4) to show that

$$(9.6) \qquad \partial_t P_t \varphi = \mathcal{L} P_t \varphi.$$

After making the change of variable $y \to e^{\frac{t}{2}}y$, one sees that another expression for $P_t\varphi$ is

$$(9.7) \qquad P_t\varphi(x) = \int \varphi\big(e^{-\frac{t}{2}}y\big)g_{e^t-1}(y-x)\,dy = \int g_1(y)\varphi\big((1-e^{-t})^{\frac{1}{2}}y+x\big)\,dy,$$

from which it is easy to see that $P_t\varphi \longrightarrow \varphi$ uniformly on compact subsets as $t \searrow 0$. Further, if $p \in [1,\infty)$, then, by Minkowski's inequality,

$$\|P_tf\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \le \int g_1(y)\left(\int \big|f\big((1-e^{-t})^{\frac{1}{2}}y+x\big)\big|^p\,dx\right)^{\frac{1}{p}}\,dy = \|f\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})},$$

and, as $t \searrow 0$,

$$\|P_tf-f\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \le \int g_1(y)\left(\int \big|f\big((1-e^{-t})^{\frac{1}{2}}y+x\big)-f(x)\big|^p\,dy\right)^{\frac{1}{p}}\,dx \longrightarrow 0$$

since

$$2\|f\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \ge \left(\int \big|f\big((1-e^{-t})^{\frac{1}{2}}y+x\big)-f(x)\big|^p\,dx\right)^{\frac{1}{p}} \longrightarrow 0.$$

Therefore we know that

$$(9.8) \qquad \|P_tf\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \le \|f\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \text{ and } \lim_{t\searrow 0}\|P_tf-f\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} = 0.$$

In particular, we have now shown that $\{P_t : t > 0\}$ is a continuous contraction semigroup, known as the *Ornstein–Uhlenbeck* semigroup, on $L^p(\lambda_{\mathbb{R}};\mathbb{C})$ for each $p \in [1,\infty)$.

Although $\{P_t : t > 0\}$ is a continuous semigroup on the Lebesgue spaces $L^p(\lambda_{\mathbb{R}};\mathbb{C})$, these are not the Lebesgue spaces on which it acts most naturally. Instead, one should consider its action on the spaces $L^p(\gamma;\mathbb{C})$, where γ is the standard Gauss measure $\gamma(dx) = (2\pi)^{-\frac{1}{2}}e^{-\frac{x^2}{2}}\lambda_{\mathbb{R}}(dx)$. The reason why is that

$$e^{-\frac{x^2}{2}}p(t,x,y) = p(t,y,x)e^{-\frac{y^2}{2}},$$

which means that

$$(9.9) \qquad \big(\varphi, P_t\psi\big)_{L^2(\gamma;\mathbb{C})} = \big(P_t\varphi, \psi\big)_{L^2(\gamma;\mathbb{C})}.$$

Hence, since $P_t\mathbf{1} = \mathbf{1}$,

$$\int P_t\varphi\,d\gamma = \big(\varphi, P_t\mathbf{1}\big)_{L^2(\gamma;\mathbb{C})} = \int \varphi\,d\gamma.$$

At the same time, by Jensen's inequality, $|P_t\varphi|^p \le P_t|\varphi|^p$, and so,

$$\int |P_t\varphi|^p\,d\gamma \le \int P_t|\varphi|^p\,d\gamma = \int |\varphi|^p\,d\gamma.$$

Thus,

$$(9.10) \qquad \|P_t\varphi\|_{L^p(\gamma;\mathbb{C})} \le \|\varphi\|_{L^p(\gamma;\mathbb{C})} \text{ for all } p \in [1,\infty).$$

In addition, if $\varphi \in C_{\mathrm{b}}(\mathbb{R};\mathbb{C})$, then $\|P_t\varphi\|_{\mathrm{u}} \le \|\varphi\|_{\mathrm{u}}$ and $P_t\varphi \longrightarrow \varphi$ pointwise as $t \searrow 0$, and therefore, for each $p \in [1,\infty)$, $\|P_t\varphi - \varphi\|_{L^p(\gamma;\mathbb{C})} \longrightarrow 0$ as $t \searrow 0$. Finally, if $\varphi \in L^p(\mathbb{R};\mathbb{C})$, then there exists a sequence $\{\varphi_n : n \ge 1\} \subseteq C_{\mathrm{b}}(\mathbb{R};\mathbb{C})$ such that $\lim_{n\to\infty}\|f_n - f\|_{L^p(\gamma;\mathbb{C})} = 0$, and

$$\|P_t\varphi - \varphi\|_{L^p(\gamma;\mathbb{C})} \le \|P_t(\varphi-\varphi_n)\|_{L^p(\gamma;\mathbb{C})} + \|P_t\varphi_n - \varphi_n\|_{L^p(\gamma;\mathbb{C})} + \|\varphi_n - \varphi\|_{L^p(\gamma;\mathbb{C})}$$

$$\le 2\|\varphi_n - \varphi\|_{L^p(\gamma;\mathbb{C})} + \|P_t\varphi_n - \varphi_n\|_{L^p(\gamma;\mathbb{C})}.$$

Thus, after first letting $t \searrow 0$ and then $n \to \infty$, we see that

$$(9.11) \qquad \lim_{t \searrow 0} \|P_t \varphi - \varphi\|_{L^p(\gamma;\mathbb{C})} = 0 \text{ for all } p \in [1, \infty).$$

Summarizing, $\{P_t : t > 0\}$ is a continuous contraction semigroup on $C_b(\mathbb{R};\mathbb{C})$ and on $L^p(\gamma;\mathbb{C})$ for each $p \in [1, \infty)$, and P_t is self-adjoint on $L^2(\gamma;\mathbb{C})$.

Exercise 9.1. Show that

$$(9.12) \qquad \left\|\varphi - (\varphi, \mathbf{1})_{L^2(\gamma;\mathbb{C})}\right\|^2_{L^2(\gamma;\mathbb{C})} \leq \|\varphi'\|^2_{L^2(\gamma;\mathbb{C})} \text{ for } \varphi \in C_b^1(\mathbb{R};\mathbb{C})$$

and that

$$(9.13) \qquad \left\|P_t\varphi - (\varphi, \mathbf{1})_{L^2(\gamma;\mathbb{C})}\right\|^2_{L^2(\gamma;\mathbb{C})} \leq e^{-t}\left\|\varphi - (\varphi, \mathbf{1})_{L^2(\gamma;\mathbb{C})}\right\|^2_{L^2(\gamma;\mathbb{C})}$$

for $\varphi \in L^2(\gamma;\mathbb{C})$. The inequality in (9.12) is the *Poincaré inequality* for γ.

Hint: Note that if suffices to handle $\varphi \in C_b^2(\mathbb{R};\mathbb{C})$ for which $(\varphi, \mathbf{1})_{L^2(\gamma;\mathbb{C})} = 0$. Next, given such a φ, show that

$$(P_t\varphi)' = e^{-\frac{t}{2}}P_t\varphi' \text{ and } -\partial_t\|P_t\varphi\|^2_{L^2(\gamma;\mathbb{C})} = \|(P_t\varphi)'\|^2_{L^2(\gamma;\mathbb{C})}.$$

Use these to show that

$$-\partial_t\|P_t\varphi\|^2_{L^2(\gamma;\mathbb{C})} = \left\|(P_t\varphi)'\right\|^2_{L^2(\gamma;\mathbb{C})} \leq e^{-t}\|\varphi'\|^2_{L^2(\gamma;\mathbb{C})}.$$

10. Hermite Polynomials

Define $H_n(x) = (-1)^n e^{\frac{x^2}{2}} \partial_x^n e^{-\frac{x^2}{2}}$. Then H_n is an nth order monic polynomial known as the nth *Hermite polynomial*. Define the operator $A_+ = x\mathbf{1} - \partial_x$, and note that $A_+H_n = H_{n+1}$, for which reason it is called the *raising operator*. Using this, check that $H_n(-x) = (-1)^n H_n(x)$.

Next note that if $\varphi, \psi \in C^1(\mathbb{R};\mathbb{C})$ which together with their derivatives have at most exponential growth, then

$$(10.1) \qquad (A_+\varphi, \psi)_{L^2(\gamma;\mathbb{C})} = (\varphi, \partial\psi)_{L^2(\gamma;\mathbb{C})}.$$

Hence, if $0 \leq m \leq n$, then

$$(H_n, H_m)_{L^2(\gamma;\mathbb{C})} = (H_0, \partial^n H_m)_{L^2(\gamma;\mathbb{C})} = \begin{cases} m! & \text{if } n = m \\ 0 & \text{if } n > m. \end{cases}$$

Next, observe that if $n \geq 1$, then $\partial H_n \in \text{span}\{H_m : 0 \leq m < n\}$, and so

$$\partial H_n = \sum_{m=0}^{n-1} \frac{(\partial H_n, H_m)_{L^2(\gamma;\mathbb{C})} H_m}{m!}$$
$$= \sum_{m=0}^{n-1} \frac{(H_n, H_{m+1})_{L^2(\gamma;\mathbb{C})} H_m}{m!} = \frac{(H_n, H_n)_{L^2(\gamma;\mathbb{C})} H_{n-1}}{(n-1)!}.$$

Hence $\partial H_n = nH_{n-1}$, and for this reason $A_- \equiv \partial$ is called the *lowering operator*.

Theorem 10.1. $\|H_m\|_{L^2(\gamma;\mathbb{C})} = (m!)^{\frac{1}{2}}$ and $\{H_m : m \geq 0\}$ is an orthogonal basis in $L^2(\gamma;\mathbb{C})$. Equivalently, if $\tilde{H}_m = \frac{H_m}{\sqrt{m!}}$, then $\{\tilde{H}_m : m \geq 0\}$ is an orthonormal basis in $L^2(\gamma;\mathbb{C})$.

Proof. All that we need to show is that if $\varphi \in L^2(\gamma; \mathbb{C})$ and $(\varphi, H_m)_{L^2(\gamma;\mathbb{C})} = 0$ for all $m \geq 0$, then $\varphi = \mathbf{0}$. To this end, use Taylor's theorem for $e^{-\frac{x^2}{2}}$ to see that, for all $\zeta \in \mathbb{C}$,

$$(10.2) \qquad e^{\zeta x - \frac{\zeta^2}{2}} = \sum_{m=0}^{\infty} \frac{\zeta^m}{m!} H_m(x),$$

where the series converges uniformly on compact subsets of $\mathbb{C} \times \mathbb{R}$, and, by calulation above, in $L^2(\gamma; \mathbb{C})$ uniformly for ζ in compact subsets of $\mathbb{C}$. Now suppose that $\varphi \in L^2(\gamma; \mathbb{C})$, and set $\psi(x) = e^{-\frac{s^2}{2}} \varphi(x)$. Then

$$\|\psi\|_{L^1(\lambda_{\mathbb{R}},\mathbb{C})} = \int_{\mathbb{R}} e^{-\frac{x^2}{4}} \left(e^{-\frac{x^2}{4}} |\varphi(x)| \right) ds \leq (2\pi)^{\frac{1}{2}} \|\varphi\|_{L^2(\gamma;\mathbb{C})},$$

and

$$e^{\frac{\xi^2}{2}} \hat{\psi}(\xi) = (2\pi)^{\frac{1}{2}} \int_{\mathbb{R}} e^{\iota \xi x - \frac{(\iota\xi)^2}{2}} \varphi(x) \gamma(dx) = (2\pi)^{\frac{1}{2}} \sum_{m=0}^{\infty} \frac{(\iota\xi)^m (\varphi, H_m)_{L^2(\gamma;\mathbb{C})}}{m!}.$$

Hence $\hat{\psi}$ and therefore φ vanish if $(\varphi, H_m)_{L^2(\gamma;\mathbb{C})} = 0$ for all $m \geq 0$. $\qquad\square$

Observe that $\mathcal{L} = -\frac{A_+ A_-}{2}$, and therefore, by (10.1)

$$\left(\mathcal{L}\varphi, \psi \right)_{L^2(\gamma;\mathbb{C})} = -\tfrac{1}{2} \left(\varphi', \psi' \right)_{L^2(\gamma;\mathbb{C})} = \left(\varphi, \mathcal{L}\psi \right)_{L^2(\gamma;\mathbb{C})}$$

for $\varphi, \psi \in C^2(\mathbb{R}; \mathbb{C})$ which together with their derivatives have at most exponential growth. Thus, by (9.6) and (9.9),

$$\left(\mathcal{L}P_t\varphi, \psi \right)_{L^2(\gamma;\mathbb{C})} = \partial_t \left(P_t\varphi, \psi \right)_{L^2(\gamma;\mathbb{C})} = \partial_t \left(\varphi, P_t\psi \right)_{L^2(\gamma;\mathbb{C})}$$

$$= \left(\varphi, \mathcal{L}P_t\psi \right)_{L^2(\gamma;\mathbb{C})} = \left(P_t\mathcal{L}\varphi, \psi \right)_{L^2(\gamma;\mathbb{C})},$$

and therefore $\mathcal{L}P_t = P_t\mathcal{L}$. In particular, because $-2\mathcal{L}H_n = nA_+H_{n-1} = nH_n$,

$$\partial_t P_t H_n = \mathcal{L}P_t H_n = P_t \mathcal{L} H_n = -\frac{n}{2} P_t H_n,$$

and so, because $\lim_{t \searrow 0} P_t H_n = H_n$,

$$(10.3) \qquad P_t H_n = e^{-\frac{nt}{2}} H_n.$$

Exercise 10.1. Using (10.3), give another proof of (9.13), and, using $A_+ H_m = H_{m+1}$, give another proof of (9.12).

11. HERMITE FUNCTIONS

Define $T : L^2(\gamma; \mathbb{C}) \longrightarrow L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ so that $T\varphi(x) = \pi^{-\frac{1}{4}} e^{-\frac{x^2}{2}} \varphi(2^{\frac{1}{2}}x)$, and check that

$$\|T\varphi\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \|\varphi\|_{L^2(\gamma;\mathbb{C})} \text{ and } T^{-1}f(x) = \pi^{\frac{1}{4}} e^{\frac{x^2}{4}} f(2^{-\frac{1}{2}}x).$$

Thus T is an isometric isomorphism from $L^2(\gamma; \mathbb{C})$ onto $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$.

Set $h_m = TH_m$ and $\tilde{h}_m = h_m = T\tilde{H}_m$. Then, because $\{\tilde{H}_m : m \geq 0\}$ is an orthonormal basis in $L^2(\gamma; \mathbb{C})$, $\{\tilde{h}_m : m \geq 0\}$ is an orthonormal bases in $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$.

Assuming that $\varphi \in C^1(\mathbb{R}; \mathbb{C})$, it easy to show that

$$TA_{\pm}\varphi = a_{\pm}T\varphi \text{ where } a_{\pm} = 2^{-\frac{1}{2}}(x\mathbf{1} \mp \partial_x)$$

and therefore that

$$(11.1) \qquad a_+ h_m = h_{m+1} \text{ and } a_- h_m = m h_{m-1}.$$

Theorem 11.1. *For all $m \geq 0$, $\widehat{h_m} = (2\pi)^{\frac{1}{2}} \imath^m h_m$.*

Proof. Certainly $\widehat{h_0} = (2\pi)^{\frac{1}{2}} h_0$. Assuming that $\widehat{h_m} = (2\pi)^{\frac{1}{2}} \imath^m h_m$, use integration by parts to see that

$$\widehat{h_{m+1}}(\xi) = 2^{-\frac{1}{2}} \int e^{\imath \xi x} a_+ h_m(x) \, dx = 2^{-\frac{1}{2}} \int x e^{\imath \xi x} h_m(x) \, dx + 2^{-\frac{1}{2}} \imath \xi \widehat{h_m}(\xi)$$

$$= 2^{-\frac{1}{2}} \left(-\imath (\widehat{h_m})'(\xi) + \imath \xi \widehat{h_m}(\xi) \right) = (2\pi)^{\frac{1}{2}} \imath^{m+1} a_+ h_m(\xi) = (2\pi)^{\frac{1}{2}} \imath^{m+1} h_{m+1}(\xi).$$

$\square$

Corollary 11.2. *For all $m \geq 0$,*

(11.2)
$$\|\tilde{h}_m\|_{L^1(\lambda_{\mathbb{R}};\mathbb{C})} \leq (2\pi)^{\frac{1}{2}} (m+1)^{\frac{1}{2}}, \|\tilde{h}_m\|_{\mathrm{u}} \leq (m+1)^{\frac{1}{2}} \text{ and}$$

$$\|x\tilde{h}_m\|_{\mathrm{u}} \vee \|\partial \tilde{h}_m\|_{\mathrm{u}} \leq 2m + \frac{3}{2}.$$

Proof. Since $\|\tilde{h}_0\|_{L^1(\lambda_{\mathbb{R}};\mathbb{C})} = 2^{\frac{1}{2}} \pi^{\frac{1}{4}}$, $\|\tilde{h}_0\|_{\mathrm{u}} \leq \pi^{-\frac{1}{4}}$, and

$$\pi^{\frac{1}{4}} \left\| (\tilde{h}_0)' \right\|_{\mathrm{u}} = \sup_{x \geq 0} x e^{-\frac{x^2}{2}} = e^{-\frac{1}{2}},$$

there is nothing to do when $m = 0$.

Now assume that $m \geq 1$. Using the facts that $x h_m(x) - h_m' = 2^{\frac{1}{2}} h_{m+1}$ and $x h_m + h_m' = 2^{\frac{1}{2}} m h_{m-1}$, one sees that

(11.3)
$$x\tilde{h}_m(x) = \frac{m^{\frac{1}{2}} \tilde{h}_{m-1}(x) + (m+1)^{\frac{1}{2}} \tilde{h}_{m+1}(x)}{2^{\frac{1}{2}}}$$

$$(\tilde{h}_m)'(x) = \frac{m^{\frac{1}{2}} \tilde{h}_{m-1}(x) - (m+1)^{\frac{1}{2}} \tilde{h}_{m+1}(x)}{2^{\frac{1}{2}}}.$$

Hence,

$$\int x^2 \tilde{h}_m(x)^2 \, dx = m + \frac{1}{2},$$

and so

$$\int (1 + x^2) \tilde{h}_m(x)^2 \, dx = m + \frac{3}{2},$$

which, by Schwarz's inequality, means that

$$\|\tilde{h}_m\|_{L^1(\lambda_{\mathbb{R}};\mathbb{C})} = \int (1+x^2)^{-\frac{1}{2}} (1+x^2)^{\frac{1}{2}} \tilde{h}_m(x)^2 \, dx \leq \pi^{\frac{1}{2}} \left(m + \frac{3}{2}\right)^{\frac{1}{2}} \leq (2\pi)^{\frac{1}{2}} (m+1)^{\frac{1}{2}}.$$

Because $(\tilde{h}_m)^\wedge = (2\pi)^{\frac{1}{2}} \imath^m \tilde{h}_m$,

$$\|\tilde{h}_m\|_{\mathrm{u}} = (2\pi)^{-\frac{1}{2}} \left\| (\tilde{h}_m)^\wedge \right\|_{\mathrm{u}} \leq (2\pi)^{-\frac{1}{2}} \|\tilde{h}_m\|_{L^1(\lambda_{\mathbb{R}};\mathbb{C})} \leq (m+1)^{\frac{1}{2}}.$$

To complete the proof, use the second part of (11.3) plus the preceding to see that

$$\left\| \partial \tilde{h}_m \right\|_{\mathrm{u}} \leq \left(m^{\frac{1}{2}} \|\tilde{h}_{m-1}\|_{\mathrm{u}} + (m+1)^{\frac{1}{2}} \|\tilde{h}_{m+1}\|_{\mathrm{u}} \right)$$

$$\leq \left(m + (m+1)^{\frac{1}{2}} (m+2)^{\frac{1}{2}} \right) \leq 2m + \frac{3}{2}.$$

The same argument, only this time using the first part of (11.3), proves the same estimate for $\|x\tilde{h}_m\|_{\mathrm{u}}$.

$\square$

The kernel which plays the role for the Hermite functions that the Ornstein–Uhlenbeck kernel (cf. (9.2)) $p(t, x, y)$ plays for the Hermite polynomial is

(11.4)
$$q(t, x, y) = 2^{\frac{1}{2}} e^{-\frac{t}{2}} e^{-\frac{x^2}{2}} p\big(2t, 2^{\frac{1}{2}}x, 2^{\frac{1}{2}}y\big) e^{\frac{y^2}{2}}$$
$$= (2\pi \sinh t)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\tanh t} + \frac{xy}{\sinh t} - \frac{y^2}{2\tanh t}\right).$$

Observe that $q(t, x, \cdot) \in L^p(\lambda_{\mathbb{R}}; \mathbb{C})$ for all $p \in [1, \infty]$ and that

$$e^{\frac{t}{2}} \int q(t, x, y) f(y)\, dy = e^{-\frac{x^2}{2}} \int p\big(2t, 2^{\frac{1}{2}}x, y\big) e^{\frac{y^2}{4}} f\big(2^{-\frac{1}{2}}y\big)\, dy = \big(T P_{2t} T^{-1} f\big)(x).$$

Hence, the operator Q_t given by

$$Q_t f(x) = \int q(t, x, y) f(y)\, dy$$

is well defined on $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ and is equal to $e^{-\frac{t}{2}} T P_{2t} T^{-1}$. In particular, by (9.10),

$$e^{\frac{t}{2}} \|Q_t f\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} = \|P_{2t} T^{-1} f\|_{L^2(\gamma; \mathbb{C})} \le \|T^{-1} f\|_{L^2(\gamma; \mathbb{C})} = \|f\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})},$$

and, by (9.11)

$$\left\| e^{\frac{t}{2}} Q_t f - f \right\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} = \left\| T(P_{2t} T^{-1} f - T^{-1} f) \right\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})}$$
$$= \left\| P_{2t} T^{-1} f - T^{-1} f \right\|_{L^2(\gamma; \mathbb{C})} \longrightarrow 0 \text{ as } t \searrow 0.$$

Hence

$$\|Q_t f\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \le e^{-\frac{t}{2}} \|f\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \text{ and } \lim_{t \searrow 0} \|Q_t f - f\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} = 0.$$

In addition, by (10.3), $Q_t h_m = e^{-\frac{t}{2}} T P_{2t} H_m = e^{-(m+\frac{1}{2})t} h_m$.

Theorem 11.3. *If $f \in L^1(\lambda_{\mathbb{R}}; \mathbb{C}) \cup L^2(\lambda_{\mathbb{R}}; \mathbb{C})$, then*

$$\int q(t, x, y) f(y)\, dy = e^{-\frac{1}{2}} \sum_{m=0}^{\infty} e^{-mt} (f, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m \text{ for } t > 0,$$

where the convergence of the series is absolute uniformly for $x \in \mathbb{R}$.

Proof. First observe that, by the estimates in Corollary 11.2, the series is absolutely convergent uniformly in $x \in \mathbb{R}$ and that both sides are continuous as functions of $f \in L^1(\mathbb{R}; \mathbb{C})$ or of $f \in L^2(\lambda_{\mathbb{R}}; \mathbb{C})$. In particular, it suffices to prove the equality when $f \in C_c(\mathbb{R}; \mathbb{C})$.

Given $f \in C_c(\mathbb{R}; \mathbb{C})$, set $f_n = \sum_{m=0}^{n} (f, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m$. Then

$$\int q(t, x, y) f_n(y)\, dy = e^{-\frac{t}{2}} \sum_{m=0}^{n} e^{-mt} (f, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m(x).$$

Because $q(t, x, \cdot) \in L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ and $f_n \longrightarrow f$ in $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$, the left hand side converges to $\int q(t, x, y) f(y)\, dy$. $\qquad\square$

Exercise 11.1. Define the *Mehler kernel* $M(\theta, x, y)$ for $(\theta, x, y) \in (0, 1) \times \mathbb{R} \times \mathbb{R}$ by

$$M(\theta, x, y) = \big(2\pi(1 - \theta^2)\big)^{-\frac{1}{2}} \exp\left(-\frac{\theta^2 x^2 - 2\theta xy + \theta^2 y^2}{2(1 - \theta^2)}\right),$$

and show that

$$M(\theta, x, y) = \left(1 - \theta^2\right)^{-\frac{1}{2}} \sum_{m=0}^{\infty} \theta^m \frac{H_m(x) H_m(y)}{m!},$$

where the series converges uniformly for (x, y) in compact subsets. This famous equation is known as *Mehler's formula*.

12. The Fourier Transform for $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$

The basic goal here is to extend the Fourier transform on $L^1(\mathbb{R}; \mathbb{C}) \cap L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ as a bounded operation on $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ into $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$. We will then examine some of the fundamental properties of this extension.

Lemma 12.1. *If $f \in L^1(\mathbb{R}; \mathbb{C})$, then*

(12.1)
$$\frac{e^{-\frac{\xi^2 \tanh t}{2}}}{(2\pi \cosh t)^{\frac{1}{2}}} \int e^{\frac{\imath \xi x}{\cosh t}} e^{-\frac{x^2 \tanh t}{2}} f(x) \, dx$$
$$= e^{-\frac{t}{2}} \sum_{m=0}^{\infty} (\imath e^{-t})^m (f, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m(\xi)$$

for $(t, \xi) \in (0, \infty) \times \mathbb{R}$.

Proof. Since both sides of (12.1) are continuous functions of $f \in L^1(\mathbb{R}; \mathbb{C})$, we may and will assume that $f \in C_c(\mathbb{R}; \mathbb{C})$.

Set $\mathbb{D} = \{\zeta \in \mathbb{C} : |\zeta| < 1 \,\&\, \zeta \notin (-1, 0]\}$, and define $\alpha_{\pm}(\zeta) = \frac{1}{2}\left(\frac{1}{\zeta} \mp \zeta\right)$ for $\zeta \in \mathbb{D}$. Next, for fixed $\xi \in \mathbb{R}$ and all $\zeta \in \mathbb{D}$, define

$$\Phi(\zeta) = \left(2\pi\alpha_+(\zeta)\right)^{-\frac{1}{2}} e^{-\frac{\alpha_-(\zeta)}{2\alpha_+(\zeta)}\xi^2} \int e^{\frac{\xi x}{\alpha_+(\zeta)}} e^{-\frac{\alpha_-(\zeta)}{2\alpha_+(\zeta)}x^2} f(x) \, dx$$

and

$$\Psi(\zeta) = \zeta^{\frac{1}{2}} \sum_{m=0}^{\infty} \zeta^m (f, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m(\xi),$$

and observe that both Φ and Ψ are analytic functions on $\mathbb{D}$. Furthermore, since $\alpha_+(e^{-t}) = \sinh t$ and $\alpha_-(e^{-t}) = \cosh t$, Lemma 11.3 says that $\Phi = \Psi$ on $(0, 1)$, and therefore, by analytic continuation, $\Phi = \Psi$ on $\mathbb{D}$. In particular, $\Phi(\imath e^{-t}) = \Psi(\imath e^{-t})$. Finally, because $\alpha_+(\imath e^{-t}) = \frac{\cosh t}{\imath}$ and $\alpha_-(\imath e^{-t}) = \frac{\sinh t}{\imath}$, one sees that the left hand and right sides of (12.1) equal, respectively $\imath^{\frac{1}{2}}\Phi(\imath e^{-t})$ and $\imath^{\frac{1}{2}}\Psi(\imath e^{-t})$. □

Theorem 12.2. *If $f \in L^1(\mathbb{R}; \mathbb{C}) \cap L^2(\mathbb{R}; \mathbb{C})$, then*

(12.2)
$$\hat{f} = (2\pi)^{\frac{1}{2}} \sum_{m=0}^{\infty} \imath^m (f, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m$$

almost everywhere.

Proof. Because $f \in L^1(\mathbb{R}; \mathbb{C})$, the left hand side of (12.1) tends pointwise to $(2\pi)^{-\frac{1}{2}}\hat{f}$ as $t \searrow 0$, and because $f \in L^2(\lambda_{\mathbb{R}}; \mathbb{C})$, the right hand side tends in $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ to the series on the right hand side of (12.2). □

As a consequence of Theorem 12.2, we know that $\|\hat{f}\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (2\pi)^{\frac{1}{2}}\|f\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$ for $f \in L^1(\lambda_{\mathbb{R}};\mathbb{C}) \cap L^2(\lambda_{\mathbb{R}};\mathbb{C})$. Hence the map $f \in L^1(\lambda_{\mathbb{R}};\mathbb{C}) \cap L^2(\lambda_{\mathbb{R}};\mathbb{C}) \rightsquigarrow \hat{f}$ admits a unique continuous extension as a linear map with norm $(2\pi)^{\frac{1}{2}}$ from $L^2(\lambda_{\mathbb{R}};\mathbb{C})$ into $L^2(\lambda_{\mathbb{R}};\mathbb{C})$, and (12.2) continuous to hold for this extension.

Define $\check{f}(x) = f(-x)$, and observe that $\check{h}_m = (-1)^m h_m$, $(\check{f}, g)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (f, \check{g})_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$, and $\hat{\check{f}} = \widehat{\check{f}}$. In addition, by Fubini's theorem,

$$\left(\hat{\varphi}, \tilde{h}_m\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \int\int e^{\imath\xi x}\varphi(x)\tilde{h}_m(\xi)\,dxd\xi = \int \varphi(x)\widehat{\tilde{h}}_m(x)\,dx = (2\pi)^{\frac{1}{2}}\imath^m(\varphi, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},$$

and so, for $f, g \in L^2(\lambda_{\mathbb{R}};\mathbb{C})$,

$$\begin{aligned}
\left(\hat{f}, \hat{g}\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} &= \sum_{m=0}^{\infty}\left(\hat{f}, \tilde{h}_m\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}\overline{\left(\hat{g}, \tilde{h}_m\right)}_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \\
&= 2\pi\sum_{m=0}^{\infty}\left(f, \tilde{h}_m\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}\overline{\left(g, \tilde{h}_m\right)}_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = 2\pi(f, g)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},
\end{aligned}$$

which means that *Parseval's identity*

$$(12.3) \qquad\qquad \left(\hat{f}, \hat{g}\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = 2\pi\left(f, g\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$$

holds. Finally, set $\check{f} = \hat{\check{f}}$, $(\hat{f}, g)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (f, \check{g})_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$ and therefore, by (12.3), that

$$\left((\hat{f})^{\vee}, g\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \left(\hat{f}, \hat{g}\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = 2\pi(f, g)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}.$$

Similarly, $\left((\check{f})^{\wedge}, g\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = 2\pi\left(f, g\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$, and so we have proved the *Fourier inversion formula*

$$(12.4) \qquad\qquad \left(\hat{f}\right)^{\vee} = 2\pi f = (\check{f})^{\wedge}.$$

It is important to keep in mind that $\hat{f}$ is not given by a Lebesgue integral for $f \in L^2(\lambda_{\mathbb{R}};\mathbb{C})$ unless $f \in L^1(\lambda_{\mathbb{R}};\mathbb{C})$ as well. On the other hand, because $f_R \equiv \mathbf{1}_{[-R,R]}f \in L^1(\lambda_{\mathbb{R}};\mathbb{C}) \cap L^2(\lambda_{\mathbb{R}};\mathbb{C})$ and $f_R \longrightarrow f$ in $L^2(\lambda_{\mathbb{R}};\mathbb{C})$,

$$\hat{f}(\xi) = \lim_{R\to\infty}\int_{-R}^{R} e^{\imath\xi x}f(x)\,dx,$$

where the convergence is in $L^2(\lambda_{\mathbb{R}};\mathbb{C})$.

Exercise 12.1. Define $\mathcal{F}f(\xi) = \hat{f}(2\pi\xi)$, and show that $\mathcal{F}$ is an orthogonal operator on $L^2(\lambda_{\mathbb{R}};\mathbb{C})$. Further, show that if $\mathcal{F}^*$ is the adjoint of $\mathcal{F}$, then equals $\mathcal{F}^{-1}f = \mathcal{F}^*f = (\mathcal{F}f)^{\cup} = \mathcal{F}\check{f}$.

13. Schwartz Test Functions

In this section we will study a space of functions introduced by Laurent Schwartz[7] and used by him to construct the class of distributions discussed in the next section.

The function space alluded to above is denoted by $\mathscr{S}(\mathbb{R};\mathbb{C})$ and consists of functions $\varphi \in C^{\infty}(\mathbb{R};\mathbb{C})$ with the property that $x \rightsquigarrow x^k\partial^{\ell}\varphi(x)$ is bounded for all $k, \ell \in \mathbb{N}$. Obviously, $\mathscr{S}(\mathbb{R};\mathbb{C})$ is a vector space. In addition, it is closed under

---

[7]There are many books in which Schwartz's theory is presented, but his own original treatment in *Théorie des distributions, I* published in 1950 by Hermann, Paris remains one of the best accounts.

differentiation as well as products with smooth functions which, together with all their derivatives, have at most polynomial growth (i.e., grow no faster than some power of $(1 + x^2)$). Thus the Hermite functions are all in $\mathscr{S}(\mathbb{R}; \mathbb{C})$. Finally, since, for $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{C})$,

$$\int |\varphi(x)|^p \, dx \leq \|(1 + x^2)\varphi\|_{\mathrm{u}}^p \int (1 + x^2)^{-p} \, dx,$$

$\mathscr{S}(\mathbb{R}; \mathbb{C}) \subseteq \bigcap_{p \in [1, \infty]} L^p(\lambda_{\mathbb{R}}; \mathbb{C})$.

There is an obvious notion of convergence for sequences in $\mathscr{S}(\mathbb{R}; \mathbb{C})$. Namely, define the norms

$$\|\varphi\|_{\mathrm{u}}^{(k,\ell)} = \|x^k \partial^\ell \varphi\|_{\mathrm{u}}$$

for k, $\ell \in \mathbb{N}$, and say that $\varphi_j \longrightarrow \varphi$ in $\mathscr{S}(\mathbb{R}; \mathbb{C})$ if $\lim_{n \to \infty} \|\varphi_j - \varphi\|_{\mathrm{u}}^{(k,\ell)} = 0$ for all k, $\ell \in \mathbb{N}$. The corresponding topology is the one for which G is open if and only if for each $\varphi \in G$ there an $m \in \mathbb{N}$ and $r > 0$ such that

$$\left\{ \psi : \|\psi - \varphi\|_{\mathrm{u}}^{(m)} < r \right\} \subseteq G,$$

where

$$\| \cdot \|_{\mathrm{u}}^{(m)} \equiv \max_{\substack{k, \ell \in \mathbb{N} \\ k + \ell \leq m}} \| \cdot \|_{\mathrm{u}}^{(j,\ell)}.$$

We will now develop a more convenient description of the topology on $\mathscr{S}(\mathbb{R}; \mathbb{C})$, one that shows that $\mathscr{S}(\mathbb{R}; \mathbb{C})$ shares many properties with Hilbert spaces. Define the operator $\mathcal{H}$ on $\mathscr{S}(\mathbb{R}; \mathbb{C})$ into itself by

$$\mathcal{H}\varphi = x^2 \varphi - \partial^2 \varphi.$$

Since (cf. (11.1)) $\mathcal{H} = (2a_+ a_- + \mathbf{1})$,

(13.1) $$\mathcal{H}\tilde{h}_k = \mu_k \tilde{h}_k \quad \text{where } \mu_k = 2k + 1,$$

and so we can define operators $\mathcal{H}^s$ for any $s \in \mathbb{R}$ by

$$\mathcal{H}^s \varphi = \sum_{m=0}^{\infty} \mu_m^s (\varphi, \tilde{h}_m)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \tilde{h}_m.$$

For each $m \geq 0$, set

$$\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C}) = \left\{ \varphi \in L^2(\lambda_{\mathbb{R}}; \mathbb{C}) : \sum_{k=1}^{\infty} \mu_k^m \big|(\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})}\big|^2 < \infty \right\},$$

and define

$$(\varphi, \psi)_{\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})} = \sum_{k=0}^{\infty} \mu_k^m (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} (\tilde{h}_k, \psi)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} = \left(\varphi, \mathcal{H}^m \psi\right)_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})}$$

$$\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})} = (\varphi, \varphi)_{\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})}^{\frac{1}{2}} = \left(\varphi, \mathcal{H}^m \varphi\right)^{\frac{1}{2}}.$$

Clearly $\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})$ is a vector space for which $(\varphi, \psi)_{\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})}$ is an inner product. Below we will show below that it is a separable Hilbert space.

Lemma 13.1. *For each $m \geq 0$,*

$$\|x\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})} \vee \|\partial\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})} \leq 3^m \|\varphi\|_{\mathscr{S}^{(m+1)}(\mathbb{R}; \mathbb{C})}.$$

Proof. By the first part of (11.3),

$$\|x\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}^2 = \sum_{k=0}^{\infty} \mu_k^m |(x\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2$$

$$= \sum_{k=1}^{\infty} k\mu_k^m |(\varphi, \tilde{h}_{k-1})_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2 + \sum_{k=0}^{\infty} (k+1)\mu_k^m |(\varphi, \tilde{h}_{k+1})_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2$$

$$= \mu_1^m |(\varphi, \tilde{h}_0)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2 + \sum_{k=1}^{\infty} ((k+1)\mu_{k+1}^m + k\mu_{k-1}^m)|(\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2$$

$$\leq 3^m \mu_0^{m+1} m |(\varphi, \tilde{h}_0)_{L^2(\lambda_{[0,1)};\mathbb{C})}|^2 + \sum_{k=1}^{\infty} (2^m(k+1)+k)\mu_k^m |(\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2$$

$$\leq 3^m \|\varphi\|_{\mathscr{S}^{(m+1)}(\mathbb{R};\mathbb{C})}.$$

Using the second part of (11.3) and the same argument, one can show that $\|\partial\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}$ $\leq 3^m \|\varphi\|_{\mathscr{S}^{(m+1)}(\mathbb{R};\mathbb{C})}.$ $\square$

Theorem 13.2. *For each $m \in \mathbb{N}$, $\mathscr{S}(\mathbb{R};\mathbb{C})$ is a dense subset of $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$. In addition, for each $m \geq 0$, there exists a $K_m \in (0,\infty)$ such that*

$$(13.2) \qquad \|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \leq K_m \|\varphi\|_{\mathrm{u}}^{(m+1)}$$

and

$$(13.3) \qquad \|\varphi\|_{\mathrm{u}}^{(m)} \leq K_m \|\varphi\|_{\mathscr{S}^{(m+3)}(\mathbb{R};\mathbb{C})}.$$

for all $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$. Thus $\varphi_n \longrightarrow \varphi$ in $\mathscr{S}(\mathbb{R};\mathbb{C})$ if and only if

$$\lim_{n\to\infty} \|\varphi_n - \varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = 0$$

for all $m \in \mathbb{N}$. In particular, for each $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$,

$$\sum_{k=0}^{n} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \tilde{h}_k \longrightarrow \varphi \text{ in } \mathscr{S}(\mathbb{R};\mathbb{C}) \text{ as } n \to \infty.$$

Proof. Since $\mathcal{H} \upharpoonright \mathscr{S}(\mathbb{R};\mathbb{C})$ is a symmetric operator, (13.1) implies that

$$\mu_k^m (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (\varphi, \mathcal{H}^m \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (\mathcal{H}^m \varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},$$

for all $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and $m \geq 0$, from which it is clear that $\mathscr{S}(\mathbb{R};\mathbb{C}) \subseteq \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ for all $m \geq 0$. Moreover, since, for each $\varphi \in \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$,

$$\mathscr{S}(\mathbb{R};\mathbb{C}) \ni \sum_{k=0}^{n} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \tilde{h}_k \longrightarrow \varphi \text{ in } \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C}) \text{ as } n \to \infty,$$

$\mathscr{S}(\mathbb{R};\mathbb{C})$ is dense in $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$.

Next observe that there exist constants $c_{j,\ell}^{(m)} \in \mathbb{R}$ such that

$$(x^2 - \partial^2)^m \varphi = \sum_{\substack{k,\ell \in \mathbb{N} \\ k+\ell \leq 2m}} c_{j.\ell}^{(m)} x^k \partial^\ell \varphi,$$

and use integration by parts to see that

$$(\varphi, x^k \partial^\ell \varphi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (-1)^{\ell'} (\partial^{\ell'}(x^{k'}\varphi), x^{k-k'}\partial^{\ell-\ell'}\varphi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},$$

where
$$k' = \begin{cases} \frac{k}{2} & \text{if } k \text{ is even} \\ \frac{k-1}{2} & \text{if } k \text{ is odd} \end{cases} \quad \text{and} \quad \ell' = \begin{cases} \frac{\ell}{2} & \text{if } \ell \text{ is even} \\ \frac{\ell+1}{2} & \text{if } \ell \text{ is odd.} \end{cases}$$

Hence there exist constants $b^{(m)}_{(k_1,\ell_1),(k_2,\ell_2)} \in \mathbb{R}$ such that

$$\big(\varphi, \mathcal{H}^m \varphi\big)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \leq \sum_{\substack{(k_1,\ell_1),(j_2,\ell_2)\in\mathbb{N}^2 \\ (k_1+\ell_1)\vee(j_2+\ell_2)\leq m}} \big|b^{(m)}_{(k_1,\ell_1),(k_2,\ell_2)}\big(x^{k_1}\partial^{\ell_1}\varphi, x^{k_2}\partial^{\ell_2}\varphi\big)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}\big|$$

$$\leq \sum_{\substack{(k_1,\ell_1),(k_2,\ell_2)\in\mathbb{N}^2 \\ (k_1+\ell_1)\vee(k_2+\ell_2)\leq m}} \big|b^{(m)}_{(k_1,\ell_1),(k_2,\ell_2)}\big|\big\|x^{k_1}\partial^{\ell_1}\varphi\big\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}\big\|x^{k_2}\partial^{\ell_2}\varphi\big\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}.$$

Finally, observe that

$$\|x^k\partial^\ell\varphi\|^2_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \int (1+x^2)^{-1}\big|(1+x^2)^{\frac{1}{2}}x^k\partial^\ell\varphi(x)\big|^2\,dx$$
$$\leq \pi\big(\|x^k\partial^\ell\varphi\|^2_{\mathrm{u}} + \|x^{k+1}\partial^\ell\varphi\|^2_{\mathrm{u}}\big),$$

and combine this with the preceding to arrive at (13.2).

To prove (13.3), begin by making repeated application of Lemma 13.1 to show that
$$\|x^k\partial^\ell\varphi\|_{\mathscr{S}^{(3)}(\mathbb{R};\mathbb{C})} \leq 3^{\frac{m(m+1)}{2}}\|\varphi\|_{\mathscr{S}^{(k+\ell+3)}(\mathbb{R};\mathbb{C})} \text{ if } k+\ell\leq m.$$

Thus, if we show that there is a $K \in (0,\infty)$ such that
$$\|\varphi\|_{\mathrm{u}} \leq K\|\varphi\|_{\mathscr{S}^{(3)}(\mathbb{R};\mathbb{C})}, \tag{$*$}$$

then
$$\|x^k\partial^\ell\varphi\|_{\mathrm{u}} \leq K\|x^k\partial^\ell\varphi\|_{\mathscr{S}^{(3)}(\mathbb{R};\mathbb{C})} \leq K\|\varphi\|_{\mathscr{S}^{(k+\ell+3)}(\mathbb{R};\mathbb{C})},$$

in which case we would know that $\|\varphi\|_{\mathrm{u}}^{(m)} \leq 3^{\frac{m(m+1)}{2}}K\|\varphi\|_{\mathscr{S}^{(m+3)}(\mathbb{R};\mathbb{C})}$. To prove $(*)$, use the estimate in (11.2) to see that

$$\|\varphi\|_{\mathrm{u}} \leq \sum_{k=0}^{\infty}|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|\|\tilde{h}_k\|_{\mathrm{u}}$$
$$\leq \sum_{k=0}^{\infty}(k+1)^{\frac{1}{2}}|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}| = \sum_{k=0}^{\infty}\left(\frac{k+1}{\mu_k^3}\right)^{\frac{1}{2}}\mu_k^{\frac{3}{2}}|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|$$

$$\leq \left(\sum_{k=0}^{\infty}\frac{k+1}{\mu_k^3}\right)^{\frac{1}{2}}\left(\sum_{k=0}^{\infty}\mu_k^3|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2\right)^{\frac{1}{2}} = K\|\varphi\|_{\mathscr{S}^{(3)}(\mathbb{R};\mathbb{C})}.$$

where $K = \left(\sum_{k=0}^{\infty}\frac{k+1}{\mu_k^3}\right)^{\frac{1}{2}}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

As a consequence of Theorem 13.2, we know that

$$\rho_{\mathscr{S}}(\varphi,\psi) \equiv \sum_{m=0}^{\infty}\frac{1}{2^{m+1}}\frac{\|\varphi-\psi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}}{1+\|\varphi-\psi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}}$$

is a metric for the topology on $\mathscr{S}(\mathbb{R};\mathbb{C})$. In addition, $\mathscr{S}(\mathbb{R};\mathbb{C}) = \bigcap_{m=0}^{\infty}\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$, and so we can learn about properties of $\mathscr{S}(\mathbb{R};\mathbb{C})$ by understanding those of the $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$'s.

For each $m \geq 0$, let $\mathfrak{s}^{(m)}(\mathbb{N}; \mathbb{C})$ be the space of functions $s : \mathbb{N} \longrightarrow \mathbb{C}$ such that

$$\|s\|_{\mathfrak{s}^{(m)}(\mathbb{N};\mathbb{C})} \equiv \left(\sum_{k=0}^{\infty} \mu_k^m |s(k)|^2 \right)^{\frac{1}{2}} < \infty,$$

and define

$$(s, t)_{\mathfrak{s}^{(m)}(\mathbb{N};\mathbb{C})} = \sum_{k=0}^{\infty} \mu_k^m s(k) \overline{t(k)} \text{ for } s, \, t \in \mathfrak{s}^{(m)}(\mathbb{N}, \mathbb{C}),$$

Clearly each $\mathfrak{s}^{(m)}(\mathbb{N}; \mathbb{C})$ is a vector space with inner product $(s, t)_{\mathfrak{s}^{(m)}(\mathbb{N};\mathbb{C})}$. Finally, set $\mathfrak{s}(\mathbb{N}; \mathbb{C}) = \bigcap_{m=0}^{\infty} \mathfrak{s}^{(m)}(\mathbb{N}; \mathbb{C})$, and turn $\mathfrak{s}(\mathbb{N}; \mathbb{C})$ into a metric space with metric

$$\rho_{\mathfrak{s}}(s, t) \equiv \sum_{m=0}^{\infty} \frac{1}{2^{m+1}} \frac{\|t - s\|_{\mathfrak{s}^{(m)}(\mathbb{N};\mathbb{C})}}{1 + \|t - s\|_{\mathfrak{s}^{(m)}(\mathbb{N};\mathbb{C})}}.$$

The following corollary is essentially a reformulation of the results in Theorem 13.2. It is the analogue for $\mathscr{S}(\mathbb{R}; \mathbb{C})$ of the fact that every separable Hilbert space is isomorphic to $\ell^2(\mathbb{N}; \mathbb{C})$.

Corollary 13.3. *Define the map* $S : L^2(\lambda_{\mathbb{R}}; \mathbb{C}) \longrightarrow \ell^2(\mathbb{N}; \mathbb{C})$ *by*

$$[S(\varphi)](k) = (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}.$$

Then, for each $m \geq 0$, $S \restriction \mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})$ *is an isometric isomorphism from* $\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})$ *onto* $\mathfrak{s}^{(m)}(\mathbb{N}; \mathbb{C})$, *and so* $S \restriction \mathscr{S}(\mathbb{R}; \mathbb{C})$ *is isometric homeomorphism from* $\mathscr{S}(\mathbb{R}; \mathbb{C})$ *onto* $\mathfrak{s}(\mathbb{N}; \mathbb{C})$.

Corollary 13.3 means that any topological property of $\mathfrak{s}^{(m)}(\mathbb{N}; \mathbb{C})$ or $\mathfrak{s}(\mathbb{N}; \mathbb{C})$ is a property of $\mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})$ or $\mathscr{S}(\mathbb{R}; \mathbb{C})$, and the following lemma facilitates the study of such properties.

Lemma 13.4. *Let* $\{\alpha_k : k \geq 0\} \subseteq (0, \infty)$, *and define the measure* ν *on* $\mathbb{N}$ *by* $\nu(\{k\}) = \alpha_k$. *Then* $L^2(\nu; \mathbb{C})$ *is a separable Hilbert space. In addition, a set* $B \subseteq L^2(\nu; \mathbb{C})$ *is relatively compact if and only if* B *is bounded and tight in the sense that*

$$\lim_{K \to \infty} \sup_{s \in B} \sum_{k > K} \alpha_k |s(k)|^2 = 0.$$

Proof. Since the L^2-space for any measure on a countably generated σ-algebra is a separable Hilbert space, $L^2(\nu; \mathbb{C})$ is a separable Hilbert space.

Since $L^2(\nu; \mathbb{C})$ is complete, to prove that a bounded, tight subset B is relatively compact it suffices to show that B is totally bounded (i.e., for every $r > 0$ there is a finite cover of B by balls of radius r with centers in B). To that end, let $r > 0$ be given, and choose K so that

$$\sup_{s \in B} \sum_{k > K} \alpha_k |s(k)|^2 < \frac{r^2}{4}.$$

Next, note that $\{(s(0), \ldots, s(K)) : s \in B\}$ is a bounded subset of $\mathbb{C}^{K+1}$ and therefore totally bounded there. Hence there exists a finite set $\{s_j : 1 \leq j \leq J\} \subseteq B$ such that, for each $s \in B$,

$$\min_{1 \leq j \leq J} \sum_{k=0}^{K} \alpha_k |s(k) - s_j(k)|^2 < \frac{r^2}{2},$$

which means that, for each $s \in B$ there exists a $1 \le j \le J$ such that

$$\|s - s_j\|_{L^2(\nu;\mathbb{C})}^2 = \sum_{k=0}^{K} \alpha_k |s(k) - s_j(k)|^2 + \sum_{k>K} \alpha_k |s(k) - s_j(k)|^2 \le r^2.$$

Finally, suppose that B is relatively compact. Certainly it is bounded. To see that it must be tight, suppose it were not. Then there would exist an $\epsilon > 0$ such that, for each $K \in \mathbb{N}$,

$$\sup_{s \in B} \sum_{k>K} \alpha_k |s(k)|^2 > \epsilon.$$

Thus we could find a sequence $\{s_K : K \ge 0\} \subseteq B$ with the property that $\sum_{k>K} \alpha_k |s_K(k)|^2 \ge \epsilon$, and, because B is relatively compact, we could choose it to be a sequence which converges to some $t \in L^2(\nu;\mathbb{C})$. But this would mean that

$$\sum_{k>K} \alpha_k |t(k)|^2 \ge \sum_{k>K} \alpha_k |s_K(k)|^2 - \|t - s_K\|_{L^2(\nu;\mathbb{C})}^2 \ge \frac{\epsilon}{2}$$

for sufficient large K, and that would mean the t can't be in $L^2(\nu;\mathbb{C})$. $\square$

Say that $B \subseteq \mathscr{S}(\mathbb{R};\mathbb{C})$ is *bounded* in $\mathscr{S}(\mathbb{R};\mathbb{C})$ if

$$\sup_{\varphi \in B} \|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} < \infty \text{ for each } m \ge 0.$$

Theorem 13.5. $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ *is a separable Hilbert space for each $m \ge 0$, and $\mathscr{S}(\mathbb{R};\mathbb{C})$ is a complete separable metric space. Moreover, a subset $B \subseteq \mathscr{S}(\mathbb{R};\mathbb{C})$ is relatively compact if and only if it is bounded in $\mathscr{S}(\mathbb{R};\mathbb{C})$.*

Proof. By Lemma 13.4 applied with $\alpha_k = \mu_k^m$, we know that each of the spaces $\mathfrak{s}^{(m)}(\mathbb{N};\mathbb{C})$ is a separable Hilbert space, and therefore, by Corollary 13.3, so is each $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$. Thus, since $\mathscr{S}(\mathbb{R};\mathbb{C})$ is dense in every $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$, we can use a diagonalization argument to find a sequence $\{\varphi_n : n \ge 1\} \subseteq \mathscr{S}(\mathbb{R};\mathbb{C})$ which is dense in $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ for all $m \ge 0$. Since this means that

$$\inf_{n \ge 1} \|\varphi - \varphi_n\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = 0 \text{ for all } \varphi \in \mathscr{S}(\mathbb{R};\mathbb{C}) \text{ and } m \ge 0,$$

it follows that

$$\inf_{n \ge 1} \rho_{\mathscr{S}(\mathbb{R};\mathbb{C})}(\varphi, \varphi_n) = 0 \text{ for all } \varphi \in \mathscr{S}(\mathbb{R};\mathbb{C}).$$

That is, $\{\varphi_n : n \ge 1\}$ is dense in $\mathscr{S}(\mathbb{R};\mathbb{C})$, and so $\mathscr{S}(\mathbb{R};\mathbb{C})$ is separable.

To see that $\mathscr{S}(\mathbb{R};\mathbb{C})$ is complete, first use Lemma 13.4 and Corollary 13.3 to see that each $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ is complete. Now suppose that $\{\varphi_n : n \ge 1\} \subseteq \mathscr{S}(\mathbb{R};\mathbb{C})$ is $\rho_{\mathscr{S}(\mathbb{R};\mathbb{C})}$-Cauchy convergent. Then it is $\|\cdot\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}$-Cauchy convergent for each $m \ge 0$, and therefore it is convergent in each $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ to some element of $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$. But if $\varphi_n \longrightarrow \varphi$ in $\mathscr{S}^{(m+1)}(\mathbb{R};\mathbb{C})$, then $\varphi_n \longrightarrow \varphi$ in $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$, and so there is a unique $\varphi \in \bigcap_{m=0}^{\infty} \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ to which $\{\varphi_n : n \ge 1\}$ converges in $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ for all $m \ge 0$. Therefore $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and $\lim_{n \to \infty} \rho_{\mathscr{S}}(\varphi, \varphi_n) = 0$.

Finally, suppose that $B \subseteq \mathscr{S}(\mathbb{R};\mathbb{C})$ is relatively compact in $\mathscr{S}(\mathbb{R};\mathbb{C})$. Because B is then relatively compact in each $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ and therefore bounded there, it is a bounded subset of $\mathscr{S}(\mathbb{R};\mathbb{C})$. Conversely, if B is bounded in $\mathscr{S}(\mathbb{R};\mathbb{C})$, in order to show that it is relatively compact in $\mathscr{S}(\mathbb{R};\mathbb{C})$ we need only show that it is totally bounded there. To that end, first observe that it is bounded in each $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$.

Thus, by Lemma 13.4 and Corollary 13.3, we will know that it is relatively compact in $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ if

$$\lim_{K\to\infty}\sup_{\varphi\in B}\sum_{k>K}\mu_k^m|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2 = 0. \qquad (*)$$

But

$$\sum_{k>K}\mu_k^m|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2 \le \frac{1}{\mu_{K+1}}\|\varphi\|^2_{\mathscr{S}^{(m+1)}(\mathbb{R};\mathbb{C})},$$

and so, since B is bounded in $\mathscr{S}^{(m+1)}(\mathbb{R};\mathbb{C})$, $(*)$ holds. To complete the proof that B $\rho_{\mathscr{S}}$-totally bounded, let $r > 0$ be given, and choose m so that $2^{-m} < \frac{r}{2}$. Next, using the fact that B is relatively compact in $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$, choose $\{\varphi_j : 1 \le j \le J\} \subseteq B$ so that

$$\sup_{\varphi\in B}\min_{1\le j\le J}\|\varphi - \varphi_j\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} < \frac{r}{2},$$

and conclude that

$$B \subseteq \bigcup_{j=1}^{J}\{\varphi : \rho_{\mathscr{S}(\mathbb{R};\mathbb{C})}(\varphi,\varphi_j) < r\}.$$

$\square$

The assertion in the following is one of the many virtues possessed by $\mathscr{S}(\mathbb{R};\mathbb{C})$.

Theorem 13.6. *The map $\varphi \rightsquigarrow \hat{\varphi}$ is an isomorphism from $\mathscr{S}(\mathbb{R};\mathbb{C})$ onto itself, and, for each $m \ge 0$, $\|\hat{\varphi}\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = (2\pi)^{\frac{1}{2}}\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}$.*

Proof. We already know that the Fourier transform is an isomorphism of $L^2(\lambda_{\mathbb{R}};\mathbb{C})$ onto $L^2(\lambda_{\mathbb{R}};\mathbb{C})$. In addition, by Theorem 12.2,

$$(\hat{\varphi},\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (2\pi)^{\frac{1}{2}}(-i)^k(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},$$

and so

$$\|\hat{\varphi}\|^2_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = 2\pi\sum_{k=0}^{\infty}\mu_k^m|(\varphi,\tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|^2 = 2\pi\|\varphi\|^2_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}.$$

$\square$

Exercise 13.1. Show that for each $(m,n) \in \mathbb{N}^2$ there is a $C_{n,m} \in (0,\infty)$ such that

$$\frac{1}{C_{n,m}}\max_{\substack{k,\ell\in\mathbb{N}\\k+\ell\le m}}\|x^k\partial^\ell\varphi\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})} \le \|\varphi\|_{\mathscr{S}^{(n+m)}(\mathbb{R};\mathbb{C})} \le C_{n,m}\max_{\substack{k,\ell\in\mathbb{N}\\k+\ell\le m}}\|x^k\partial^\ell\varphi\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})}.$$

Hint: In proving the upper bound, consider using the equation

$$\left(a_+^n\varphi,\tilde{h}_{k+n}\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \left(\frac{(k+n)!}{k!}\right)^{\frac{1}{2}}\left(\varphi,\tilde{h}_k\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}.$$

.

Exercise 13.2. Let $\{\varphi_n : n \geq 1\}$ be a bounded sequence in $\mathscr{S}(\mathbb{R};\mathbb{C})$ such that $\lim_{n\to\infty} \varphi_n(x)$ exists for each $x \in \mathbb{R}$. Show that there is a $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ such that $\varphi_n \longrightarrow \varphi$ in $\mathscr{S}(\mathbb{R};\mathbb{C})$.

Hint: Use Theorem 13.5.

Exercise 13.3. This exercise deals with the relationship between various function spaces.

(**i**) Show that $C_c^\infty(\mathbb{R};\mathbb{C})$ is a dense subset of $\mathscr{S}(\mathbb{R};\mathbb{C})$

(**ii**) Set
$$C_0(\mathbb{R};\mathbb{C}) = \left\{ f \in C(\mathbb{R};\mathbb{C}) : \lim_{|x|\to\infty} f(x) = 0 \right\}.$$
Show that $C_0(\mathbb{R};\mathbb{C})$ with the uniform norm is a Banach space in which both C_c^∞ and $\mathscr{S}(\mathbb{R};\mathbb{C})$ are dense subsets.

Exercise 13.4. For $x \in \mathbb{R}$ and $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$, define $\tau_x\varphi(y) = \varphi(x+y)$. Show that $\tau_x\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and that $\|\tau_x\varphi\|_u^{(m)} \leq 2^m(|x| \vee 1)^m\|\varphi\|_u^{(m)}$ for all $m \geq 0$. In addition, show that
$$\|\tau_{x_2}\varphi - \tau_{x_1}\varphi\|_u^{(m)} \leq 2^m\big(|x_1| \vee |x_2| \vee 1\big)^m\|\varphi\|_u^{(m+1)}|x_2 - x_1|.$$

Hint: To prove the first estimate, check that
$$|y^k\partial^\ell\tau_x\varphi(y)| \leq \begin{cases} (2|x|)^m\big|(\partial^\ell\varphi)(x+y)\big| & \text{if } |y| \leq 2|x| \\ 2^m\big|(x+y)^k(\partial^\ell\varphi)(x+y)\big| & \text{if } |y| \geq 2|x|. \end{cases}$$

To prove the second estimate, assume that $x_1 \leq x_2$, note that
$$\tau_{x_2}\varphi - \tau_{x_1}\varphi = \int_{x_1}^{x_2} \tau_t\varphi' \, dt,$$
and therefore that
$$\|\tau_{x_2}\varphi - \tau_{x_1}\varphi\|_u^{(m)} \leq \int_{x_1}^{x_2} \|\tau_t\varphi'\|_u^{(m)} \, dt.$$

Finally, apply the first estimate.

14. Tempered Distributions

   Schwartz developed the theory of distribution in order to provide a mathematically rigorous way to describe the sort of generalized functions that appear in the work by Boole and Heaviside in connection with applications of the Laplace transform to ordinary differential equations, and those that were somewhat later introduced by Sobolev for applications to partial differential equations. What Schwartz realized is that generalized functions should be thought of in terms of their *action* (i.e., their $L^2(\lambda_\mathbb{R};\mathbb{C})$ inner product with) on smooth functions rather than their value (which won't exist in general) at points.

   To make that idea mathematically precise, he said a generalized function, which he called a *distribution*, should be a continuous linear functional on a topological vector space of smooth functions. One of the spaces Schwartz considered is $C_c^\infty(\mathbb{R};\mathbb{C})$, but the appropriate topology on that space is rather cumbersome (for instance, elements don't have countable neighborhood bases). A second, and much more tractable, choice is $\mathscr{S}(\mathbb{R};\mathbb{C})$. Because elements of $\mathscr{S}(\mathbb{R};\mathbb{C})$ need not have compact support, an element of its dual space must satisfy restricted growth conditions and is therefore called a *tempered distribution*.

Recall that the dual space X^* of a topological vector space X over $\mathbb{C}$ is the space of continuous, $\mathbb{C}$-valued linear functions on X. When, like $\mathscr{S}(\mathbb{R};\mathbb{C})$, all the elements of X have a countable neighborhood basis, a linear function Λ on X is an element of X^* if $\Lambda x_n \longrightarrow \Lambda x$ whenever $x_n \longrightarrow x$ in X. Because we want to think of elements of $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ as generalized functions which act via their L^2-inner product with elements of $\mathscr{S}(\mathbb{R};\mathbb{C})$, we will use letters like u to denote elements of $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ and write their action on $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ as $\langle \varphi, u \rangle$.

Lemma 14.1. *For each $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$ there is an $m \geq 0$ and a $C \in (0,\infty)$ such that*
$$|\langle \varphi, u \rangle| \leq C\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \text{ for all } \varphi \in \mathscr{S}(\mathbb{R};\mathbb{C}).$$

Proof. Because sets of the form $\left\{\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \leq r\right\}$ form a neighborhood basis for $\mathbf{0}$ in $\mathscr{S}(\mathbb{R};\mathbb{C})$, there is an $m \geq 0$ and $r > 0$ such that $|\langle \varphi, u \rangle| \leq 1$ when $\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \leq r$. Hence $|\langle \varphi, u \rangle| \leq r^{-1}\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}$. $\square$

Simple as it is, Lemma 14.1 has many consequences. For example, it allows us to say that

$$(14.1) \qquad \langle \varphi, u \rangle = \sum_{k=0}^{\infty} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle,$$

where the series is absolutely convergent. Indeed, if $|\langle \varphi, u \rangle| \leq C\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}$, then $|\langle \tilde{h}_k, u \rangle| \leq C\mu_k^m$, and so, since $|(\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}| = \mu_k^{-n}|(\mathcal{H}^n\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}|$ for all $n \geq 0$, the series $\sum_{k=0}^{\infty} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle$ is absolutely convergent. Hence, if $\varphi_n = \sum_{k=0}^{n} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$, then $\varphi_n \longrightarrow \varphi$ in $\mathscr{S}(\mathbb{R};\mathbb{C})$ and therefore

$$\langle \varphi, u \rangle = \lim_{n \to \infty} \langle \varphi_n, u \rangle = \lim_{n \to \infty} \sum_{k=0}^{n} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle$$

$$= \sum_{k=0}^{\infty} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle.$$

Obviously, given a measurable function $f : \mathbb{R} \longrightarrow \mathbb{C}$ with at most polynomial growth, one can think of it as the element $f\lambda_{\mathbb{R}}$ of $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ given by $\langle \varphi, f\lambda_{\mathbb{R}} \rangle = \int \varphi \bar{f} \, d\lambda_{\mathbb{R}}$, and in this way $\mathscr{S}(\mathbb{R};\mathbb{C})$ can be thought of as a subset of $\mathscr{S}(\mathbb{R};\mathbb{C})^*$. Although the distribution corresponding to f is $f\lambda_{\mathbb{R}}$, it is conventional to denote it by f instead, and we will adopt this convention.

We will need to know that $\mathscr{S}(\mathbb{R};\mathbb{C})$ is dense in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$. To see that it is, let $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$, and set

$$\psi_n = \sum_{k=0}^{n} \overline{\langle \tilde{h}_k, u \rangle} \tilde{h}_k.$$

Clearly $\psi_n \in \mathscr{S}(\mathbb{R};\mathbb{C})$, and, for each $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$,

$$(\varphi, \psi_n)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \sum_{k=0}^{n} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \overline{(\psi_n, \tilde{h}_k)}_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$$

$$= \sum_{k=0}^{n} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle \longrightarrow \sum_{k=0}^{\infty} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle = \langle \varphi, u \rangle.$$

The importance of this density result is that it tells us how to extend continuous operators like $\mathcal{H}^s$ as continuous operators on $\mathscr{S}(\mathbb{R};\mathbb{C})^*$. Namely, because

$(\varphi, \mathcal{H}^s \psi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (\mathcal{H}^s \varphi, \psi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$ for $\varphi, \psi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and $\mathscr{S}(\mathbb{R};\mathbb{C})$ is dense in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$, the one and only continuous extension of $\mathcal{H}^s$ to $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ is given by

(14.2) $$\langle \varphi, \mathcal{H}^s u \rangle \equiv \langle \mathcal{H}^s \varphi, u \rangle.$$

Since $\mathscr{S}(\mathbb{R};\mathbb{C})$ can be written as the intersection of the spaces $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$, $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ must be able to be written as the union of the spaces $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*$. Of course, because $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ is a Hilbert space, Riesz's theorem provides an isomorphism between $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*$ and $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$. However, in order to be consistent with the idea that $\langle \varphi, u \rangle$ is a generalization of the L^2 inner product, this is not the way we will think about $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*$. Instead, we want to identify $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*$ as the Hilbert space

$$\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C}) = \left\{ u \in \mathscr{S}(\mathbb{R};\mathbb{C})^* : \sum_{k=0}^{\infty} \mu_k^{-m} |\langle \tilde{h}_k, u \rangle|^2 < \infty \right\}$$

with inner product

$$(u, v)_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})} = \sum_{k=0}^{\infty} \mu_k^{-m} \overline{\langle \tilde{h}_k, u \rangle} \langle \tilde{h}_k, v \rangle.$$

Recall that if X is a Banach space and $\Lambda \in X^*$, then $\|\Lambda\|_{X^*} = \sup\{|\Lambda(x)| : \|x\|_X = 1\}$. Thus

$$\|u\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*} = \sup\{|\langle \varphi, u \rangle| : \|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = 1\}.$$

Theorem 14.2. *For each $m \geq 0$, $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ is a separable Hilbert space in which $\mathscr{S}(\mathbb{R};\mathbb{C})$ is a dense subset, and*

$$u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C}) \iff \mathcal{H}^{-\frac{m}{2}} u \in L^2(\lambda_{\mathbb{R}};\mathbb{C}) \,\&\, |\mathcal{H}^{-\frac{m}{2}} u\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$$
$$\iff u \in \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*.$$

Moreover, if $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$, then $\|u\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*} = \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$ and therefore

(14.3) $$|\langle \varphi, u \rangle| \leq \|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}.$$

Proof. That $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ is a seperable Hilbert space follows from Lemma 13.4. Next, let $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ and set $u_n = \sum_{k=0}^{n} \overline{\langle \tilde{h}_k, u \rangle} \tilde{h}_k$. Then $u_n \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and

$$\|u - u_n\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}^2 = \sum_{k>n} \mu_k^{-m} |\langle \tilde{h}_k, u \rangle|^2 \longrightarrow 0.$$

Hence $\mathscr{S}(\mathbb{R};\mathbb{C})$ is dense in $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$.

If $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$, then

$$|\langle \varphi, \mathcal{H}^{-\frac{m}{2}} u \rangle| = \left| \sum_{k=0}^{\infty} \mu_k^{-\frac{m}{2}} (\varphi, \tilde{h}_k)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \langle \tilde{h}_k, u \rangle \right| \leq \|\varphi\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})},$$

and so $\mathcal{H}^{-\frac{m}{2}} u \in L^2(\lambda_{\mathbb{R}};\mathbb{C})$ and $\|\mathcal{H}^{-\frac{m}{2}} u\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \leq \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$. Conversely, if $\mathcal{H}^{-\frac{m}{2}} u \in L^2(\lambda_{\mathbb{R}};\mathbb{C})$, then

$$\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}^2 = \sum_{k=0}^{\infty} \mu_k^{-m} |\langle \tilde{h}_k, u \rangle|^2 = \sum_{k=0}^{\infty} |\langle \tilde{h}_k, \mathcal{H}^{-\frac{m}{2}} u \rangle|^2 = \|\mathcal{H}^{-\frac{m}{2}} u\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}.$$

To prove the second equivalence, first suppose that $u \in \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*$. Then, since $\|\mathcal{H}^{-\frac{m}{2}}\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = \|\varphi\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$,

$$
\begin{aligned}
|\langle \varphi, \mathcal{H}^{-\frac{m}{2}}u \rangle| &= |\langle \mathcal{H}^{-\frac{m}{2}}\varphi, u \rangle| \\
&\leq \|\mathcal{H}^{-\frac{m}{2}}\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}\|u\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*} = \|u\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*}\|\varphi\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},
\end{aligned}
$$

and so $\mathcal{H}^{-\frac{m}{2}}u \in L^2(\lambda_{\mathbb{R}};\mathbb{C})$ and $\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})} \leq \|u\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*}$. Conversely, if $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$, set $f = \mathcal{H}^{-\frac{m}{2}}u$, then

$$
\begin{aligned}
|\langle \varphi, u \rangle| &= |(\mathcal{H}^{\frac{m}{2}}\varphi, f)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}| \\
&\leq \|\mathcal{H}^{\frac{m}{2}}\varphi\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}\|f\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})},
\end{aligned}
$$

and so $u \in \mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*$ and $\|u\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})^*} \leq \|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$. $\qquad\square$

By combining Lemma 14.1 and Theorem 14.2, we know that

$$
\mathscr{S}(\mathbb{R};\mathbb{C})^* = \bigcup_{m=0}^{\infty} \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C}).
$$

Theorem 14.3. *If $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ is non-negative in the sense that $\langle \varphi, u \rangle \geq 0$ whenever $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ is non-negative, then there exists a Borel measure μ on $\mathbb{R}$ such that*

$$
\int (1+x^2)^{-\frac{m+2}{2}}\,\mu(dx) < \infty \text{ and } \langle \varphi, \mu \rangle = \int \varphi\,d\mu.
$$

Conversely, if μ is a Borel measure on $\mathbb{R}$ satisfying

$$
\int (1+x^2)^{-\frac{m}{2}}\,\mu(dx) < \infty
$$

and $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$ is defined by $\langle \varphi, u \rangle = \int \varphi\,d\mu$, then $u \in \mathscr{S}^{(-m-3)}(\mathbb{R};\mathbb{C})$.

Proof. Assume that $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ is non-negative. Choose $\eta \in C^{\infty}(\mathbb{R};[0,1])$ so that $\eta = 1$ on $[-1,1]$ and $\eta = 0$ off $[-2,2]$, set $\eta_R(x) = \eta\left(\frac{x}{R}\right)$ for $R \geq 1$, and define $u_R \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$ by $\langle \varphi, u_R \rangle = \langle \eta_R\varphi, u \rangle$. Given an $\mathbb{R}$-valued $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$, $\|\varphi\|_u\eta_R \pm \varphi\eta_R \geq 0$, and therefore $|\langle \varphi, u_R \rangle| \leq \|\varphi\|_u\langle \eta_R, u \rangle$. Thus there is a unique extension of $\varphi \rightsquigarrow \langle \varphi, u_R \rangle$ as a continuous, non-negative linear functional on $C\left([-2R;2R],\mathbb{R}\right)$, which, by the Riesz representation theorem, means that there is a finite Borel measure μ_R on $\mathbb{R}$ such that $\langle \varphi, u_R \rangle = \int \varphi\,d\mu_R$. In particular, $\mu_R(\mathbb{R}) = \langle \eta_R, u \rangle \leq \|\eta_R\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$. Since $\|\eta_R\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}^2 = \left(\eta_R, \mathcal{H}^m\eta_R\right)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$ and $\mathcal{H}^m\eta_R$ is a linear combinations of terms of the form $\frac{x^k}{R^\ell}\eta^{(\ell)}\left(\frac{x}{R}\right)$, where $0 \leq k+\ell \leq 2m$, there exists a $C < \infty$ such that

$$
\left(\int \eta_R(x)\mathcal{H}^m\eta_R(x)\,dx\right)^{\frac{1}{2}} \leq CR^{m+\frac{1}{2}},
$$

and so $\mu_R(\mathbb{R}) \leq C\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}R^{m+\frac{1}{2}}$.

Note that $R \leq R' \implies \mu_{R'}\restriction [-R,R] = \mu_R \restriction [-R,R]$, and therefore there is a Borel measure μ on $\mathbb{R}$ such that $\mu \restriction [-R,R] = \mu_R \restriction [-R,R]$ for all $R \geq 1$.

Furthermore

$$\int (1+x^2)^{-\frac{m+2}{2}} \mu(dx) = \sum_{n=-\infty}^{\infty} \int_{[n,n+1]} (1+|x|^2)^{-\frac{m+2}{2}} \mu_n(dx)$$

$$\leq 2C\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})} \sum_{n=0}^{\infty} \frac{(n+1)^{m+\frac{1}{2}}}{(1+n^2)^{\frac{m+2}{2}}} = 2^{\frac{m+2}{2}} C\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})} \sum_{n=0}^{\infty} (1+n)^{-\frac{3}{2}} < \infty.$$

Finally,

$$\langle \varphi, u \rangle = \lim_{R\to\infty} \langle \eta_R \varphi, u \rangle = \lim_{R\to\infty} \int \eta_R \varphi\, d\mu = \int \varphi\, d\mu.$$

Conversely, suppose that μ is a Borel measure on $\mathbb{R}$ and that

$$C \equiv \int (1+x^2)^{-\frac{m}{2}}\, d\mu(dx) < \infty.$$

Clearly $\varphi \rightsquigarrow \int \varphi\, d\mu$ determines a distribution u. In fact, by (13.3),

$$|\langle \varphi, u \rangle| \leq C\|(1+x^2)^{\frac{m}{2}}\varphi\|_{\mathrm{u}} \leq C\|(1+|x|)^m\varphi\|_{\mathrm{u}} \leq CK_m\|\varphi\|_{\mathscr{S}^{(m+3)}(\mathbb{R};\mathbb{C})},$$

and therefore $u \in \mathscr{S}^{(-m-3)}(\mathbb{R};\mathbb{C})$.   □

As a consequence of Theorem 14.3, we know that for any measurable $f : \mathbb{R} \longrightarrow \mathbb{C}$ for which there exists an $m \in \mathbb{Z}$ such that

$$\int (1+x^2)^{-\frac{m}{2}}|f(x)|\, dx < \infty,$$

there is a distribution $f \in \mathscr{S}^{(-m-3)}(\mathbb{R};\mathbb{C})$ such that

$$\langle \varphi, f \rangle = \int \varphi(x)\bar{f}(x)\, dx.$$

The following generalizes the preceding observation.

Theorem 14.4. *Let μ be a Borel measure on $\mathbb{R}$, and assume that*

$$M_\mu \equiv \int (1+x^2)^{-\frac{m}{2}} \mu(dx) < \infty.$$

If $f \in L^p(\mu;\mathbb{C})$, then there is a distribution $f\mu$ given by

$$\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C}) \longmapsto \int \varphi \bar{f}\, d\mu \in \mathbb{C}.$$

Moreover, if $m_p = \min\{n : m \leq 2p'n\}$, where p' is the Hölder conjugate of p, then $f\mu \in \mathscr{S}^{(-m_p-3)}(\mathbb{R};\mathbb{C})$ and

$$\|f\mu\|_{\mathscr{S}^{(-m_p-3)}(\mathbb{R};\mathbb{C})} \leq K_{m_p} M_\mu^{\frac{1}{p'}} \|f\|_{L^p(\mu;\mathbb{C})}.$$

Proof. By Hölder's inequality,

$$\left| \int \varphi \bar{f}\, d\mu \right| \leq \|f\|_{L^p(\mu;\mathbb{C})} \|\varphi\|_{L^{p'}(\mu;\mathbb{C})}.$$

At the same time,

$$\|\varphi\|_{L^{p'}(\mu;\mathbb{C})} \leq \left(\int (1+x^2)^{-\frac{m}{2}} (1+x^2)^{\frac{m}{2}} |\varphi(x)|^{p'} \mu(dx) \right)^{\frac{1}{p'}}$$

$$\leq M_\mu^{\frac{1}{p'}} \left\| (1+x^2)^{\frac{m}{2p'}}\varphi \right\|_{\mathrm{u}} \leq K_{m_p} M_\mu^{\frac{1}{p'}} \|\varphi\|_{\mathscr{S}^{(m_p+3)}(\mathbb{R};\mathbb{C})}.$$

Hence,

$$|\langle \varphi, f\mu \rangle| \le K_{m_p} M_\mu^{\frac{1}{p'}} \|f\|_{L^p(\mu;\mathbb{C})} \|\varphi\|_{\mathscr{S}^{(m_p+3)}(\mathbb{R};\mathbb{C})}.$$

$\square$

Loosely related to the preceding is the following theorem of Schwartz. Given a $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$, its *support* is the smallest closed set F such that $\langle \varphi, u \rangle = 0$ for all φ that are 0 on $F^{\complement}$. Equivalently, $\langle \varphi_1, u \rangle = \langle \varphi_2, u \rangle$ if $\varphi_1 = \varphi_2$ on an open set containing F.

Theorem 14.5. *If $u \in \mathscr{S}^{(-n+1)}(\mathbb{R};\mathbb{C})$, then u is supported on $\{0\}$ if and only if there exist $\{a_0, \ldots, a_n\} \subseteq \mathbb{C}$ for which*

$$\langle \varphi, u \rangle = \sum_{m=0}^{n} a_m \partial^m \varphi(0)$$

for all $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$. .

Proof. The sufficiency statement is trivial. To prove the necessity assertion, first note that, by Theorem 13.2, there is a $C \in [0, \infty)$ such that $|\langle \varphi, u \rangle| \le C\|\varphi\|_{\mathrm{u}}^{(n)}$. Next, choose $\eta \in C^\infty(\mathbb{R};[0,1])$ so that $\eta = 1$ on $[-1, 1]$ and $\eta = 0$ off of $[-2, 2]$, and define $\eta_r(x) = \eta\left(\frac{x}{r}\right)$ for $r \in (0, 1]$. Because 0 is the support of u, $\langle \varphi, u \rangle = \langle \eta_r \varphi, u \rangle$ for all $r \in (0, 1]$. In particular, this means that

$$|\langle \varphi, u \rangle| \le C \sum_{\ell=0}^{n} \|\eta \varphi^{(\ell)}\|_{\mathrm{u}}$$

for some other $C < \infty$.

We will now show that $\langle \varphi, u \rangle = 0$ if $\varphi(x) = x^{n+1}\eta(x)\psi(x)$ for some $\psi \in C^\infty(\mathbb{R};\mathbb{C})$. To this end, set $\varphi_r(x) = x^{n+1}\eta_r(x)\psi(x)$, and note $\langle \varphi, u \rangle = \langle \varphi_r, u \rangle$ for all $r \in (0, 1]$. Next observe that $\partial^\ell \varphi_r$ is a linear combination of terms of the form

$$x^{n+1-i}r^{-j}\eta^{(j)}\left(\tfrac{x}{r}\right)\psi^{(k)}(x) = x^{n+1-i-j}\left(\frac{x}{r}\right)^j \eta^{(j)}\left(\tfrac{x}{r}\right)\psi^{(k)}(x)$$

where $i + j + k = \ell$. Since

$$\left| x^{n+1-i-j}\left(\frac{x}{r}\right)^j \eta^{(j)}\left(\tfrac{x}{r}\right)\psi^{(k)}(x) \right| \le (2r)^{n+1-i-j}\|x^j\eta^{(j)}\|_{\mathrm{u}}\|\psi^{(k)}\|_{\mathrm{u}},$$

$\lim_{r \searrow 0} \|\varphi_r^{(\ell)}\|_{\mathrm{u}} = 0$ for $\ell \le n$, and so

$$\langle \varphi, u \rangle = \lim_{r \searrow 0} \langle \varphi_r, u \rangle = 0.$$

Now let $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and use Taylor's theorem to write

$$\varphi(x) = \sum_{m=0}^{n} \frac{\varphi^{(m)}(0)}{m!}x^m + \frac{x^{n+1}}{n!}\int_0^1 (1-t)^n \varphi^{(n+1)}(tx)\,dt.$$

Set $\psi(x) = \frac{1}{n!}\int_0^1 (1-t)^n \varphi^{(n+1)}(tx)\,dt$, and apply the preceding to see that $\langle x^{n+1}\eta\psi, u \rangle = 0$ and therefore that

$$\langle \varphi, u \rangle = \langle \eta\varphi, u \rangle = \sum_{m=0}^{n} \frac{\varphi^{(m)}(0)}{m!}\langle x^m\eta, u \rangle.$$

Hence we can take $a_m = \frac{\langle x^m\eta, u \rangle}{m!}$. $\square$

The next result characterizes distributions $u \in \mathscr{S}(\mathbb{R}; \mathbb{C})^*$ which satisfy the *minimum principle*

(14.4) $\langle \varphi, u \rangle \geq 0$ if $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{R})$ and $\varphi(0) = \min\{\varphi(x) : x \in \mathbb{R}\}$

and are *quasi-local* in the sense that

(14.5) $\lim_{R \to \infty} \langle \varphi_R, u \rangle = 0$ for all $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{C})$,

where $\varphi_R(x) = \varphi\left(\frac{x}{R}\right)$.

In preparation for the proof of the characterization, I have to introduce the following partition of unity for $\mathbb{R} \setminus \{\mathbf{0}\}$. Choose $\psi \in C^\infty\big(\mathbb{R}; [0,1]\big)$ so that ψ has compact support in $(0, 2) \setminus \overline{\left(0, \frac{1}{4}\right)}$ and $\psi(y) = 1$ when $\frac{1}{2} \leq |y| \leq 1$, and set $\psi_m(y) = \psi(2^m y)$ for $m \in \mathbb{Z}$. Then, if $y \in \mathbb{R}$ and $2^{-m-1} \leq |y| \leq 2^{-m}$, $\psi_m(y) = 1$ and $\psi_n(y) = 0$ unless $-m - 2 \leq n \leq -m + 1$. Hence, if $\Psi(y) = \sum_{m \in \mathbb{Z}} \psi_m(y)$ for $y \in \mathbb{R} \setminus \{0\}$, then Ψ is a smooth function with values in $[1, 4]$; and therefore, for each $m \in \mathbb{Z}$, the function χ_m given by $\chi_m(0) = 0$ and $\chi_m(y) = \frac{\psi_m(y)}{\Psi(y)}$ for $y \in \mathbb{R} \setminus \{0\}$ is a smooth, $[0,1]$-valued function that vanishes off of $(0, 2^{-m+1}) \setminus \overline{(0, 2^{-m-2})}$. In addition, for each $y \in \mathbb{R} \setminus \{0\}$, $\sum_{m \in \mathbb{Z}} \chi_m(y) = 1$ and $\chi_m(y) = 0$ unless $2^{-m-2} \leq |y| \leq 2^{-m+1}$.

Lemma 14.6. *If $u \in \mathscr{S}(\mathbb{R}; \mathbb{R})$ satisfies (14.4) and (14.5), then there exists a unique Borel measure M on $\mathbb{R}$ such that $M(\{0\}) = 0$, $\int \frac{y^2}{1+y^2} M(dy) < \infty$, and*

$$\langle \varphi, u \rangle = \int \varphi(y) \, M(dy)$$

if φ, φ', and φ'' vanish at 0.

Proof. Referring to the partition of unity described above, define $\Lambda_m \varphi = \langle \chi_m \varphi, u \rangle$ for $\varphi \in C^\infty\big(\overline{(0, 2^{-m+1})} \setminus (0, 2^{-m-2}); \mathbb{R}\big)$, where

$$\chi_m \varphi(y) = \begin{cases} \chi_m(y)\varphi(y) & \text{if } 2^{-m-2} \leq |y| \leq 2^{-m+1} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly Λ_m is linear. In addition, if $\varphi \geq 0$, then $\chi_m \varphi \geq 0 = \chi_m \varphi(0)$, and so, by (14.4), $\Lambda_m \varphi \geq 0$. Similarly, for any $\varphi \in C^\infty\big(\overline{(0, 2^{-m+1})} \setminus (0, 2^{-m-2}); \mathbb{R}\big)$, $\|\varphi\|_{\mathrm{u}} \chi_m \pm \chi_m \varphi \geq 0 = \big(\|\varphi\|_{\mathrm{u}} \chi_m \pm \chi_m \varphi\big)(0)$, and therefore $|\Lambda_m \varphi| \leq K_m \|\varphi\|_{\mathrm{u}}$, where $K_m = \langle \chi_m, u \rangle$. Hence, Λ_m admits a unique extension as a continuous linear functional on $C\big(\overline{(0, 2^{-m+1})} \setminus (0, 2^{-m-2}); \mathbb{R}\big)$ that is non-negativity preserving and has norm K_m; and so, by the Riesz representation theorem, we know that there is a unique non-negative Borel measure M_m on $\mathbb{R}$ such that M_m is supported on $\overline{(0, 2^{-m+1})} \setminus (0, 2^{-m-2})$, $K_m = M_m(\mathbb{R})$, and $\langle \chi_m \varphi, u \rangle = \int_{\mathbb{R}} \varphi(y) \, M_m(dy)$ for all $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{R})$.

Now define the Borel measure M on $\mathbb{R}$ by $M = \sum_{m \in \mathbb{Z}} M_m$. Clearly, $M(\{0\}) = 0$. In addition, if $\varphi \in C_{\mathrm{c}}^\infty\big(\mathbb{R} \setminus \{0\}; \mathbb{R}\big)$, then there is an $n \in \mathbb{Z}$ such that $\chi_m \varphi \equiv 0$ unless $|m| \leq n$. Thus,

$$\langle \varphi, u \rangle = \sum_{m=-n}^{n} A(\chi_m \varphi) = \sum_{m=-n}^{n} \int_{\mathbb{R}} \varphi(y) \, M_m(dy)$$

$$= \int_{\mathbb{R}^N} \left(\sum_{m=-n}^{n} \chi_m(y)\varphi(y) \right) M(dy) = \int_{\mathbb{R}^N} \varphi(y) \, M(dy),$$

and therefore

(14.6)
$$\langle \varphi, u \rangle = \int_{\mathbb{R}} \varphi(y) \, M(dy)$$

for $\varphi \in C_c^\infty \big(\mathbb{R} \setminus \{\mathbf{0}\}; \mathbb{R} \big)$.

Before taking the next step, observe that, as an application of (14.4), if $\varphi_1, \varphi_2 \in \mathscr{S}(\mathbb{R}; \mathbb{R})$, then

$$\varphi_1 \leq \varphi_2 \text{ and } \varphi_1(0) = \varphi_2(0) \implies \langle \varphi_1, u \rangle \leq \langle \varphi_2, u \rangle. \tag{$*$}$$

Indeed, this reduces to the observation that $\varphi_2 - \varphi_1 \geq 0 = (\varphi_2 - \varphi_1)(0)$.

With these preparations, we can show that, for any $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{C})$,

$$\varphi \geq 0 = \varphi(\mathbf{0}) \implies \int_{\mathbb{R}} \varphi(y) \, M(d\mathbf{y}) \leq \langle \varphi, u \rangle. \tag{$**$}$$

To check this, apply $(*)$ to $\varphi_n = \sum_{m=-n}^n \chi_m \varphi$ and φ, and use (14.6) together with the monotone convergence theorem to conclude that

$$\int_{\mathbb{R}} \varphi(y) \, M(dy) = \lim_{n \to \infty} \int_{\mathbb{R}} \varphi_n(y) \, M(dy) = \lim_{n \to \infty} \langle \varphi_n, u \rangle \leq \langle \varphi, u \rangle.$$

Now let $\eta \in C^\infty \big(\mathbb{R}; [0,1] \big)$ satisfy $\eta = 0$ on $[-1,1]$ and $\eta = 0$ off $(-2,2)$, and set $\eta_R(y) = \eta(R^{-1}y)$ for $R > 0$. By $(**)$ with $\varphi(y) = |y|^2 \eta(y)$ we know that

$$\int_{\mathbb{R}} |y|^2 \eta(y) \, M(dy) \leq \langle \varphi, u \rangle < \infty.$$

At the same time, by (14.6), for $R \geq 2$,

$$\int_{\mathbb{R}^N} \big(\eta_R(y) - \eta(y) \big) \, M(dy) = \langle (\eta_R - \eta), u \rangle = \langle \eta_R, u \rangle - \langle \eta, u \rangle$$

and therefore, by (14.5) and Fatou's Lemma,

$$\int_{\mathbb{R}} \big(1 - \eta(y) \big) \, M(dy) \leq -\langle \eta, u \rangle < \infty.$$

Hence, we have proved that

(14.7)
$$\int_{\mathbb{R}} \frac{y^2}{1+y^2} \, M(dy) < \infty.$$

We are now in a position to show that (14.6) continues to hold for any $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{R})$ that vanishes along with its first and second order derivatives at 0. To this end, first suppose that φ vanishes in a neighborhood of 0. Then, for each $R > 0$, (14.6) applies to $\eta_R \varphi$, and so

$$\int_{\mathbb{R}} \eta_R(y)\varphi(y) \, M(dy) = \langle \eta_R \varphi, u \rangle = \langle \varphi, u \rangle + \langle (\mathbf{1} - \eta_R)\varphi, u \rangle.$$

Since φ is M-integrable and $(1 - \eta_R)\varphi \longrightarrow 0$ in $\mathscr{S}(\mathbb{R}; \mathbb{R})$ as $R \to \infty$, Lebesgue's dominated convergence theorem implies that,

$$\langle \varphi, u \rangle = \lim_{R \to \infty} \int_{\mathbb{R}} \eta_R(y)\varphi(y) \, M(dy) = \int_{\mathbb{R}} \varphi(y) \, M(dy).$$

We still have to replace the assumption that φ vanishes in a neighborhood of 0 by the assumption that it vanishes to second order there. For this purpose, first note that, by (14.7), φ is certainly M-integrable, and therefore

$$\int_{\mathbb{R}^N} \varphi(y)\, M(dy) = \lim_{r \searrow 0} \langle (\mathbf{1} - \eta_r)\varphi, u \rangle = \langle \varphi, u \rangle - \lim_{r \searrow 0} \langle \eta_r \varphi, u \rangle.$$

By our assumptions about φ at 0, we can find a $C < \infty$ such that $|\eta_r \varphi(y)| \leq Cr y^2 \eta(y)$ for all $r \in (0, 1]$. Hence, by $(*)$ and the M-integrability of $y^2 \eta(y)$, there is a $C' < \infty$ such that $\langle \eta_r \varphi, u \rangle \leq C'r$ for small $r > 0$, and therefore $\langle \eta_r \varphi, u \rangle \longrightarrow 0$ as $r \searrow 0$. $\square$

Theorem 14.7. *If $u \in \mathscr{S}(\mathbb{R}; \mathbb{R})$ satisfies (14.4) and (14.5), then there exist an $a \geq 0$, a $b \in \mathbb{R}$, and Borel measure M on $\mathbb{R}$ such that $M(\{0\}) = 0$, (14.7) holds, and*

$$\langle \varphi, u \rangle = \tfrac{a}{2}\varphi''(0) + b\varphi'(0) + \int \big(\varphi(y) - \varphi(0) - \mathbf{1}_{[0,1]}(y)\varphi'(0)y\big) M(dy).$$

In fact, M is determined by

$$\langle \varphi, u \rangle = \int \varphi(y)\, M(dy) \ \text{if } \varphi \in C_{\mathrm{c}}^\infty\big(\mathbb{R} \setminus \{0\}\big),$$

and, for any $\eta \in C^\infty\big(\mathbb{R}; [0,1]\big)$ which is 1 on $[-1, 1]$ and 0 off $(-2, 2)$

$$a = \langle y^2\eta^2, u \rangle - \int y^2\eta(y)^2\, M(dy)$$

and

$$b = \langle y\eta, y \rangle - \int y\big(\eta(y) - \mathbf{1}_{[0,1]}(y)\big) M(dy).$$

Proof. Let η be as in the statement, set $\eta_R(x) = \eta\big(\tfrac{x}{R}\big)$ for $R > 0$, and define

$$\psi_R(y) = \varphi(y) - \varphi(0)\eta_R(y) - \varphi'(0)y\eta(y) - \tfrac{1}{2}\varphi''(0)y^2\eta(y)^2.$$

Then $\psi \in \mathscr{S}(\mathbb{R}; \mathbb{C})$ and vanishes to second order at 0, and so, by Lemma 14.6, $\langle \psi_R, u \rangle = \int \psi(y)\, M(dy)$. Hence,

$$\langle \varphi, u \rangle = \varphi(0)\langle \eta_R, u \rangle + \varphi'(0)\langle y\eta, u \rangle + \tfrac{1}{2}\varphi''(0)\langle y^2\eta^2, u \rangle$$
$$+ \int \big(\varphi(y) - \varphi(0)\eta_R(y) - \varphi'(0)y\eta(y) - \tfrac{1}{2}\varphi''(0)y^2\eta(y)^2\big) M(dy),$$

and so

$$\langle \varphi, u \rangle = \varphi(0)\left(\langle \eta_R, u \rangle + \int \big(\mathbf{1} - \eta_R(y)\big) M(dy)\right)$$

$$+ \varphi'(0)\langle y\eta, u \rangle - \tfrac{1}{2}\varphi''(0)\left(\langle y^2\eta^2, u \rangle - \int y^2\eta(y)^2\, M(dy)\right)$$

$$+ \int \big(\varphi(y) - \varphi(0) - \varphi'(0)y\eta(y)\big) M(dy).$$

By (14.5) and the Lebesgue dominated convergence theorem, as $R \to \infty$ both $\langle \eta_R, u \rangle$ and $\int (\mathbf{1} - \eta_R)\, dM$ tend to 0. Finally, because $y\big(\eta(y) - \mathbf{1}_{[-1,1]}(y)\big)$ vanishes on $[-1, 1]$ and is therefore M-integrable, we can replace $\varphi'(0)\langle y\eta, u \rangle$ by

$$\varphi'(0)\left(\langle y\eta, u \rangle - \int y\big(\eta(y) - \mathbf{1}_{[-1,1]}(y)\big)\right) M(dy)$$

and $\int \big(\varphi(y) - \varphi(0) - \varphi'(0)y\eta(y)\big)M(dy)$ by

$$\int \big(\varphi(y) - \varphi(0) - \varphi'(0)y\mathbf{1}_{[-1,1]}(y)\big)M(dy).$$

$\square$

Exercise 14.1. Let $f \in C_{\mathrm{b}}^1(\mathbb{R};\mathbb{C})$, set $u = f(|x|)$, and show that $u' = \mathrm{sgn}(x)f'(|x|)$. Next assume that $f \in C_{\mathrm{b}}^2(\mathbb{R};\mathbb{C})$, and show that $u'' = f'(0)\delta_0 + f''(|x|)$.

15. Extending Continuous Operators on $\mathscr{S}(\mathbb{R};\mathbb{C})$ to $\mathscr{S}(\mathbb{R};\mathbb{C})^*$

The extension that we made of the operators $\mathcal{H}^s$ to $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ is a special case of the fact that many continuous linear maps of $\mathscr{S}(\mathbb{R};\mathbb{C})$ into $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ determine a unique continuous extension as a continuous map from $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ itself. The key to making such an extension is contained in the following theorem.

Theorem 15.1. *Let A be a continuous map of $\mathscr{S}(\mathbb{R};\mathbb{C})$ into $\mathscr{S}(\mathbb{R};\mathbb{C})^*$, and assume that there is a continuous operator A^* on $\mathscr{S}(\mathbb{R};\mathbb{C})$ such that*

$$\big(A^*\varphi,\psi\big)_{L^2(\lambda_\mathbb{R};\mathbb{C})} = \big(\varphi, A\psi\big)_{L^2(\lambda_\mathbb{R};\mathbb{C})} \text{ for all } \varphi,\psi \in \mathscr{S}(\mathbb{R};\mathbb{C}).$$

If Au is defined for $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$ by

(15.1) $$\langle\varphi, Au\rangle = \langle A^*\varphi, u\rangle \text{ for } \varphi \in \mathscr{S}(\mathbb{R};\mathbb{C}),$$

then $u \rightsquigarrow Au$ is the unique extension of A as a continuous operator on $\mathscr{S}(\mathbb{R};\mathbb{C})^*$.

Proof. Because A^* maps $\mathscr{S}(\mathbb{R};\mathbb{C})$ continuously into itself, for each $m \geq 0$ there exists an $n \geq 0$ and $C < \infty$ such that $\|A^*\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \leq C\|\varphi\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})}$, and therefore, if $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$, then

$$|\langle\varphi, Au\rangle| = |\langle A^*\varphi, u\rangle| \leq \|A^*\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})} \leq C\|\varphi\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})}\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}.$$

Hence $\|Au\|_{\mathscr{S}^{(-n)}(\mathbb{R};\mathbb{C})} \leq C\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$, and so A maps $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ continuously into $\mathscr{S}^{(-n)}(\mathbb{R};\mathbb{C})$. Furthermore, since $\mathscr{S}(\mathbb{R};\mathbb{C})$ is dense in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ and $\langle\varphi, A\psi\rangle = \big(A^*\varphi,\psi\big)_{L^2(\lambda_\mathbb{R};\mathbb{C})}$ for $\psi \in \mathscr{S}(\mathbb{R};\mathbb{C})$, A is the one and only continuous extension to $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ of $A \restriction \mathscr{S}(\mathbb{R};\mathbb{C})$. $\square$

If $A : \mathscr{S}(\mathbb{R};\mathbb{C}) \longrightarrow \mathscr{S}(\mathbb{R};\mathbb{C})^*$ is a continuous map, we will say that a continuous operator A^* on $\mathscr{S}(\mathbb{R};\mathbb{C})$ is its *adjoint* if $(A^*\varphi,\psi)_{L^2(\lambda_\mathbb{R};\mathbb{C})} = \langle\varphi, A\psi\rangle$ for all $\varphi,\psi \in \mathscr{S}(\mathbb{R};\mathbb{C})$.

Given a continuous operator A on $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ and $m, n \in \mathbb{Z}$

$$\|A\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})\to\mathscr{S}(\mathbb{R};\mathbb{C})(m)} = \sup\big\{|Au\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} : \|u\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})} = 1\big\}.$$

The argument given in the proof of Theorem 15.1 shows that, for $m, n \in \mathbb{N}$,

(15.2) $$\|A\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})\to\mathscr{S}^{(-n)}(\mathbb{R};\mathbb{C})} = \|A^*\|_{\mathscr{S}^{(n)}(\mathbb{R};\mathbb{C})\to\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}.$$

Among the simplest maps to which Theorem 15.1 applies are $\varphi \rightsquigarrow x^k\varphi$ and $\varphi \rightsquigarrow \partial^\ell\varphi$. Indeed, the first of these is its own adjoint, and the adjoint of ∂^ℓ is $(-\partial)^\ell$. By Lemma 13.1, the extensions of these maps take, respectively, $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ into $\mathscr{S}^{(-m-k)}(\mathbb{R};\mathbb{C})$ and $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ into $\mathscr{S}^{(-m-\ell)}(\mathbb{R};\mathbb{C})$.

The Fourier transform is a particularly important operator on $\mathscr{S}(\mathbb{R};\mathbb{C})^*$, and its adjoint is given by $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C}) \longmapsto \check{\varphi} \in \mathscr{S}(\mathbb{R};\mathbb{C})$. Hence

$$\langle\varphi, u\rangle = \langle\check{\varphi}, u\rangle,$$

and, since $\|\check{\varphi}\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} = (2\pi)^{\frac{1}{2}}\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}$ for all $m \geq 0$, (15.2) says that $\|\hat{u}\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})} = (2\pi)^{\frac{1}{2}}\|u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$ for all $m \geq 0$. In addition, since both sides of the equation $\langle \hat{\varphi}, \hat{u} \rangle = 2\pi \langle \varphi, u \rangle$ are continuous functions of $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$ and, by (12.3), the equation holds when $u \in \mathscr{S}(\mathbb{R};\mathbb{C})$,

$$\langle \hat{\varphi}, \hat{u} \rangle = (2\pi)\langle \varphi, u \rangle.$$

The same continuity argument shows that

$$\widehat{\partial u} = -\imath \xi \hat{u}, \ \widehat{xu} = \imath \partial \hat{u}$$

and that the Fourier inversion formula

$$(\hat{u})^\vee = 2\pi u = (\check{u})^\wedge$$

holds.

Computing most Fourier transforms of functions is hard, and computing them of distributions can be even harder. Among those that are easy are those of $x^k \delta_a$, $\partial^\ell \delta_a$ and $f \in L^1(\lambda_\mathbb{R};\mathbb{C}) \cup L^2(\lambda_\mathbb{R};\mathbb{C})$ when thought of as a tempered distribution. Indeed,

$$\langle \varphi, \hat{\delta}_a \rangle = \check{\varphi}(a) = \int e^{-\imath ax} \varphi(x)\, dx = \langle \varphi, e_a \rangle, \quad \text{where } e_a(x) = e^{\imath ax},$$

Hence, $\widehat{\partial^\ell \delta_a} = (-\imath \xi)^\ell e_a$. To compute $\hat{f}$ when f is thought of as a distribution, note that

$$\langle \check{\varphi}, f \rangle = \int \bar{f}(\xi) \left(\int e^{-\imath \xi x} \varphi(x)\, dx \right) d\xi = \int \varphi(x) \overline{\hat{f}(x)}\, dx = \langle \varphi, \hat{f} \rangle,$$

and therefore, when $f \in L^1(\lambda_\mathbb{R};\mathbb{C})$ is thought of as a distribution, its Fourier transform is the distribution determined by the function $\hat{f} \in C_\mathrm{b}(\mathbb{R};\mathbb{C})$. When $f \in L^2(\lambda_\mathbb{R};\mathbb{C})$, one uses the fact that, as $R \to \infty$, $\mathbf{1}_{[-R,R]}f \longrightarrow f$ in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ and therefore $\hat{f} = \hat{f}$ where $\hat{f} = \lim_{R\to\infty} \widehat{f_R}$ is the L^2-Fourier transform of f. Similarly, when μ is a finite Borel measure on $\mathbb{R}$, $\hat{\mu}$ as a distribution is equal to the function $\hat{\mu}$ given by

$$(15.3) \qquad\qquad \hat{\mu}(\xi) = \int e^{\imath \xi x} \mu(dx).$$

Trickier is the computation of the Fourier transform of distributions like $\log|x|$. One way to do so is to observe that $\partial \log|x| = \frac{1}{x}$ and first compute $\widehat{x^{-1}}$. For that purpose, set $f_y(x) = \frac{x}{x^2+y^2}$ for $y > 0$, and observe that, as $y \searrow 0$, $f_y \longrightarrow x^{-1}$ and therefore $\widehat{f_y} \longrightarrow \widehat{x^{-1}}$ in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$. Next observe that observe that, by (7.11),

$$\widehat{f_y}(\xi) = \lim_{R\to\infty} \int_{-R}^R \frac{xe^{\imath \xi x}}{x^2+y^2}\, dx = \imath \lim_{R\to\infty} \int_{-R}^R \frac{x \sin x}{x^2+y^2}\, dx = \imath \pi \mathrm{sgn}(\xi)e^{-y|\xi|}.$$

Hence

$$(15.4) \qquad\qquad \widehat{x^{-1}} = \imath \pi \mathrm{sgn}.$$

Knowing (15.4) one might expect that one can use $\widehat{\partial u} = -\imath \xi \hat{u}$ to compute $\widehat{\log|x|}$. However to do so it is necessary to confront a technical difficulty. Namely, $\frac{\imath \pi \mathrm{sgn}(\xi)}{-\imath \xi} = -\frac{\pi}{|\xi|}$, and $|\xi|^{-1}$ is *not* a distribution. On the other hand,

$$\varphi \rightsquigarrow \int \frac{\varphi(\xi) - \varphi(0)e^{-\frac{\xi^2}{2}}}{|\xi|}\, d\xi$$

is a distribution. Thus, to overcome the problem, set $u = \log|x|$ and write

$$\langle \varphi, \hat{u} \rangle = \langle \varphi - \varphi(0)\widehat{g_1}, \hat{u} \rangle + \varphi(0)\langle \widehat{g_1}, \hat{u} \rangle.$$

and note that $\langle \widehat{g_1}, \hat{u} \rangle = 2\pi \int g_1(x) \log|x| \, dx$. At the same time,

$$\langle \varphi - \varphi(0)\widehat{g_1}, \hat{u} \rangle = \left\langle \frac{\varphi - \varphi(0)e^{-\frac{\xi^2}{2}}}{\imath\xi}, -\imath\xi\hat{u} \right\rangle$$

$$= \left\langle \frac{\varphi - \varphi(0)e^{-\frac{\xi^2}{2}}}{\imath\xi}, \widehat{\partial u} \right\rangle = -\pi \left\langle \frac{\varphi - \varphi(0)e^{-\frac{\xi^2}{2}}}{|\xi|}, \lambda_{\mathbb{R}} \right\rangle.$$

Hence

$$\langle \varphi, \widehat{\log|x|} \rangle = -\pi \int \frac{\varphi(\xi) - \varphi(0)e^{-\frac{\xi^2}{2}}}{|\xi|} \, d\xi + 2\pi\varphi(0) \int g_1(x) \log|x| \, dx.$$

Next, consider a differential operator $L = \sum_{j=0}^{J} a_j \partial^j$ where $\{a_0, \ldots, a_J\} \subseteq C^\infty(\mathbb{R}; \mathbb{C})$ and all the a_j's and their derivatives have at most polynomial growth. If

$$L^*\varphi \equiv \sum_{j=0}^{J} (-1)^j \partial^j (a_j\varphi),$$

then it is clear that $(L^*\varphi, \psi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = (\varphi, L\varphi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$. To see that L^* is a continuous operator on $\mathscr{S}(\mathbb{R}; \mathbb{C})$, we need the following lemma.

Lemma 15.2. *Let $f \in C^\infty(\mathbb{R}; \mathbb{R})$, and assume that for each $m \geq 0$ there exists an $k_m \geq 0$ such that*

$$F_m \equiv \max_{1 \leq j \leq m} \sup_{x \in \mathbb{R}} \frac{|\partial^j f(x)|}{|x|^{k_m} \vee 1} < \infty.$$

Then, for each $m \geq 0$, there is a $C_m < \infty$ such that

$$\|\varphi f\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})} \leq C_m F_m \|\varphi\|_{\mathscr{S}^{(m+k_m)}(\mathbb{R};\mathbb{C})}.$$

Proof. By Exercise 13.1 with $n = 0$, it is sufficient for us to show that for each $k, \ell \in \mathbb{N}$ with $k + \ell \leq m$, there is a $c_{k,\ell}$ such that

$$\|x^k \partial^\ell (\varphi f)\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \leq c_{k,\ell} F_m \|\varphi\|_{\mathscr{S}^{(m+k_m)}(\mathbb{R};\mathbb{C})}.$$

To this end, remember that

$$\partial^\ell (\varphi f) = \sum_{j=0}^{\ell} \binom{\ell}{j} \partial^j \varphi \partial^{\ell-j} f,$$

and

$$\|x^k \partial^j \varphi \partial^{\ell-j} f\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} \leq F_m \big\|(1 + |x|^{k_m})x^k \partial^j \varphi\big\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$$

$$\leq 2 \cdot 3^{m+k_m} F_m \|\varphi\|_{\mathscr{S}^{(m+k_m)}(\mathbb{R};\mathbb{C})}\big\|.$$

$\square$

Using Lemma 15.2, it easy to check that L^* is a continuous operator on $\mathscr{S}(\mathbb{R}; \mathbb{C})$, and therefore L extends as a continuous operator on $\mathscr{S}(\mathbb{R}; \mathbb{C})^*$.

Another important operation is convolution. That is, given $\psi \in \mathscr{S}(\mathbb{R}; \mathbb{C})$, consider the operator $\mathcal{C}_\psi$ on $\mathscr{S}(\mathbb{R}; \mathbb{C})$ given by $\mathcal{C}_\psi \eta = \eta * \psi$. Because $\widehat{\eta * \psi} = \hat{\eta}\hat{\psi}$,

Lemma 15.2 guarantees that $\mathcal{C}_\psi$ maps $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ continuously into itself for all $m \geq 0$. In addition,

$$\langle \varphi, \psi * \eta \rangle = \iint \varphi(x)\bar\psi(x-y)\bar\eta(y)\,dxdy = \iint \varphi(x+y)\bar\psi(x)\bar\eta(y)\,dxdy = \langle \mathcal{C}_\psi^* \varphi, \eta \rangle$$

where

$$\mathcal{C}_\psi^* \varphi(y) = \int \varphi(x+y)\bar\psi(x)\,dx.$$

Since $\widehat{\mathcal{C}_\psi^* \varphi}(\xi) = \hat\varphi(\bar\psi)^\vee$, Lemma 15.2 again guarantees that, for all $m \geq 0$, $\mathcal{C}_\psi^*$ maps $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ continuously into itself, and so $\mathcal{C}_\psi$ has a unique continuous extension to $\mathscr{S}(\mathbb{R};\mathbb{C})^*$, and this extention is a continuous map of $\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})$ into itself for all $m \in \mathbb{Z}$.

In order to gain a better understanding of $\mathcal{C}_\psi$, we need to use the translation maps $\tau_x : \mathscr{S}(\mathbb{R};\mathbb{C}) \longrightarrow \mathscr{S}(\mathbb{R};\mathbb{C})$ defined in Exercise 13.4, and define $\psi * u(x) = \langle \tau_{-x}\psi, u \rangle$ for $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$ and $x \in \mathbb{R}$.

Theorem 15.3. *For $\psi \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$, $\psi * u$ is a continuous function with at most polynomial growth, and $\mathcal{C}_\psi u = \psi * u$. In addition, $\widehat{\psi * u} = \hat\psi\hat u$, and $\psi * u = (2\pi)^{-1}(\hat\psi\hat u)^\vee$.*

Proof. By Exercise 13.4, $x \rightsquigarrow \tau_{-x}\psi$ is a continuous map of $\mathscr{S}(\mathbb{R};\mathbb{C})$ into itself and therefore that $\psi * u \in C(\mathbb{R};\mathbb{C})$. Also, the estimates given in that Exercise and Lemma 13.1 show that

$$|\psi * u(x)| \leq 2^m K_m(|x| \vee 1)^m \|\psi\|_{\mathscr{S}^{(m+3)}(\mathbb{R};\mathbb{C})}\|u\|_{\mathscr{S}^{(-m-3)}(\mathbb{R};\mathbb{C})},$$

and therefore $\psi * u$ has at most polynomial growth.

Turing to the proof that $\mathcal{C}_\psi u = \psi * u$, suppose that $u \in \mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$ and set $u_n = \sum_{k=0}^n \overline{\langle \tilde h_k, u \rangle}\tilde h_k$. Then $u_n \in \mathscr{S}(\mathbb{R};\mathbb{C})$ and $u_n \longrightarrow u$ in $\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})$. Since $\mathcal{C}_\psi u_n = \psi * u_n$, we will know that $\mathcal{C}_\psi u = \psi * u$ once we show that that $\psi * u_n \longrightarrow \psi * u$ in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$. To that end, note that, by Theorem 13.2 and that Exercise 13.4,

$$|\psi * (u_n - u)(x)| \leq \|\tau_{-x}\psi\|_{\mathscr{S}^{(m)}(\mathbb{R};\mathbb{C})}\|u_n - u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$$

$$\leq K_m\|\tau_{-x}\psi\|_{\mathrm{u}}^{(m+1)}\|u_n - u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})}$$

$$\leq 2^{(m+1)}K_m(|x| \vee 1)^{m+1}\|\psi\|_{\mathrm{u}}^{(m+1)}\|u_n - u\|_{\mathscr{S}^{(-m)}(\mathbb{R};\mathbb{C})},$$

and so $\psi * u_n \longrightarrow \psi * u$ in $\mathscr{S}(\mathbb{R};\mathbb{C})^*$.

Finally, since $\widehat{\psi * u} = \hat\psi\hat u$ and $\psi * u = (2\pi)^{-1}(\hat\psi\hat u)^\vee$ when $u \in \mathscr{S}(\mathbb{R};\mathbb{C})$, the $\mathscr{S}(\mathbb{R};\mathbb{C})^*$-continuity of $u \rightsquigarrow \psi * u$ guarantees that these continue to hold for all $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$. $\qquad\square$

A simple, but typical, application of these results is to the ordinary differential equation $\lambda u - u'' = \mu$, where $\lambda > 0$ and μ is a finite Borel measure on $\mathbb{R}$. The solution u to this equation describes the electric potential along a wire produced by a charge distribution μ when the wire has resistance that is a linear function of the potential. To solve this equation, assume that $u \in \mathscr{S}(\mathbb{R};\mathbb{C})^*$, and take the Fourier transform of both sides. Then $\lambda\hat u + \xi^2\hat u = \hat\mu$, and so $\hat u = \frac{\hat\mu}{\lambda+\xi^2}$. Next observe (cf. (7.5)) that $\frac{1}{\lambda+\xi^2} = \widehat{G_\lambda}$, where

$$G_\lambda(x) = \frac{1}{2\lambda^{\frac{1}{2}}}e^{-\lambda^{\frac{1}{2}}|x|}.$$

Even though $G_\lambda \notin \mathscr{S}(\mathbb{R};\mathbb{C})$, it and the function $x \rightsquigarrow G_\lambda * \mu(x) \equiv \int G_\lambda(x-y)\,\mu(dy)$ are elements of $L^1(\lambda_\mathbb{R};\mathbb{C})$ and therefore of $\mathscr{S}(\mathbb{R};\mathbb{C})^*$. In addition, by Fubini's theorem, $\widehat{G_\lambda * \mu} = \widehat{G_\lambda}\hat{\mu}$, and so

$$u(x) = \frac{1}{2\lambda^{\frac{1}{2}}} \int e^{-\lambda^{\frac{1}{2}}|x-y|}\,\mu(dy).$$

It is an instructive exercise to check that this u is a solution. To this end, first use Exercise 14.1 to see that u' is the function

$$u'(x) = \frac{\lambda^{\frac{1}{2}}}{2} \int \operatorname{sgn}(y-x)e^{-\lambda^{\frac{1}{2}}|x-y|}\,dy.$$

Thus

$$\langle \varphi, u'' \rangle = -\langle \varphi', u' \rangle = \int \varphi'(x)\left(\frac{1}{2}\int \operatorname{sgn}(x-y)e^{-\lambda^{\frac{1}{2}}|x-y|}\varphi'(y)\,\mu(dy)\right)dx$$

$$= \int \left(\frac{1}{2}\int \operatorname{sgn}(x-y)e^{-\lambda^{\frac{1}{2}}|x-y|}\varphi'(x)\,dx\right)\mu(dy).$$

Next note that

$$\int \operatorname{sgn}(x-y)e^{-\lambda^{\frac{1}{2}}|x-y|}\varphi'(x)\,dx = \int_y^\infty e^{\lambda^{\frac{1}{2}}(y-x)}\varphi'(x)\,dx - \int_{-\infty}^y e^{\lambda^{\frac{1}{2}}(x-y)}\varphi'(x)\,dx$$

$$= -\varphi(y) + \lambda^{\frac{1}{2}}\int_y^\infty e^{\lambda^{\frac{1}{2}}(y-x)}\,dx - \varphi(y) + \lambda^{\frac{1}{2}}\int_{-\infty}^y e^{\lambda^{\frac{1}{2}}(x-y)}\,dx = -2\varphi(y) + 2\lambda u(y),$$

and therefore $\langle \varphi, u'' \rangle = -\langle \varphi, \mu \rangle + \lambda\langle \varphi, u \rangle$, which means that $\lambda u - u'' = \mu$.

Exercise 15.1. This exercise deals with the special case when an element of $\mathscr{S}(\mathbb{R};\mathbb{C})^*$ is given by a Borel measure μ.

(**i**) Show that $\psi * \mu$ equals the function

$$x \in \mathbb{R} \longmapsto \int \psi(x-y)\,\mu(dy) \in \mathbb{C}.$$

(**ii**) If μ is finite, show that $\hat{\mu}$ equals the function

$$\xi \in \mathbb{R} \longmapsto \hat{\mu}(\xi) \equiv \int e^{i\xi x}\,\mu(dx) \in \mathbb{C}$$

and that $\hat{\mu} \in C_b(\mathbb{R};\mathbb{C})$ with norm $\|\hat{u}\|_u = \mu(\mathbb{R})$.

(**iii**) If $\int (1+x^2)^{\frac{m}{2}}\,\mu(dx) < \infty$ for some $m \geq 0$, show that $\hat{\mu} \in C_b^m(\mathbb{R};\mathbb{C})$ and that

$$\|\partial^k \hat{\mu}\|_u \leq \int |x|^k\,\mu(dx) \text{ for } 0 \leq k \leq m.$$

(**iv**) Assume that $\int |x|^k\,\mu(dx) < \infty$ for all $k \in \mathbb{N}$, and show that $\psi * \mu$ is an element of $\mathscr{S}(\mathbb{R};\mathbb{C})$ for all $\psi \in \mathscr{S}(\mathbb{R};\mathbb{C})$.

Hint: Show that $\widehat{\psi * \mu}$ is an element of $\mathscr{S}(\mathbb{R};\mathbb{C})$.

16. Moving to $\mathbb{R}^N$

With essentially no new ideas and the introduction of only slightly uglier notation, we will transfer most of the contents of §§7–15 to $\mathbb{R}^N$.

If $f \in L^1(\mathbb{R}^N; \mathbb{C})$, its Fourier transform is the function

$$\hat{f}(\boldsymbol{\xi}) = \int e^{\imath(\boldsymbol{\xi},\mathbf{x})_{\mathbb{R}^N}} f(\mathbf{x}) \, d\mathbf{x},$$

and, using exactly the same arguments as we did when $N = 1$, one can easily show that $\|\hat{f}\|_{\mathrm{u}} \leq \|f\|_{L^1(\lambda_{\mathbb{R}^N};\mathbb{C})}$, $\hat{f}$ is continuous and that, if $f \in C^1(\mathbb{R}^N;\mathbb{C}) \cap L^1(\lambda_{\mathbb{R}^N};\mathbb{C})$ and $f' \in L^1(\lambda_{\mathbb{R}^N};\mathbb{C})$, then $\widehat{\partial_{x_j} f}(\boldsymbol{\xi}) = -\imath \xi_j \hat{f}(\boldsymbol{\xi})$ for $1 \leq j \leq N$, from which it follows that $\hat{f}(\boldsymbol{\xi}) \longrightarrow 0$ as $|\boldsymbol{\xi}| \to \infty$.

To develop an inversion formula, one introduces the functions

$$g_t(\mathbf{x}) = (2\pi t)^{-\frac{N}{2}} e^{-\frac{|\mathbf{x}|^2}{2t}},$$

uses Fubini's theorem to check that $\widehat{g_t}(\boldsymbol{\xi}) = e^{-\frac{t|\boldsymbol{\xi}|^2}{2}}$, and proceeds as before to see first that

$$\int g_t(\mathbf{x} - \mathbf{y}) f(\mathbf{y}) \, d\mathbf{y} = (2\pi)^{-N} \int e^{-\frac{t|\boldsymbol{\xi}|^2}{2}} e^{-\imath(\boldsymbol{\xi},\mathbf{x})_{\mathbb{R}^N}} \hat{f}(\boldsymbol{\xi}) \, d\boldsymbol{\xi},$$

and then that, as $t \searrow 0$,

$$(2\pi)^{-N} \int e^{-\frac{t|\boldsymbol{\xi}|^2}{2}} e^{-\imath(\boldsymbol{\xi},\mathbf{x})_{\mathbb{R}^N}} \hat{f}(\boldsymbol{\xi}) \, d\mathbf{x} \text{ converges to } \begin{cases} f & \text{in } L^1(\lambda_{\mathbb{R}^N};\mathbb{C}) \\ f(\mathbf{x}) & \text{if } f \text{ is continuous at } \mathbf{x}. \end{cases}$$

The normalized Hermite functions on $\mathbb{R}^N$ are indexed by $\mathbf{m} = (m_1, \ldots, m_N) \in \mathbb{N}^N$ and defined by

$$\tilde{h}_{\mathbf{m}}(\mathbf{x}) = \tilde{h}_{n_1}(x_1) \cdots \tilde{h}_{n_N}(x_N).$$

By standard results about products of Hilbert spaces, one knows that they form an orthonormal basis in $L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$. In addition, if

$$\mathcal{H} = |\mathbf{x}|^2 - \Delta = \sum_{j=1}^{N} (x_j^2 - \partial_{x_j}^2),$$

then

$$\mathcal{H}\tilde{h}_{\mathbf{m}} = \mu_{\mathbf{m}} \tilde{h}_{\mathbf{m}} \text{ where } \mu_{\mathbf{m}} = \sum_{j=1}^{N} \mu_{m_j}$$

and

$$(\tilde{h}_{\mathbf{m}})^{\wedge} = \imath^{\|\mathbf{m}\|_1} (2\pi)^{\frac{N}{2}} \tilde{h}_{\mathbf{m}} \text{ where } \|\mathbf{m}\|_1 = \sum_{j=1}^{N} m_j.$$

Finally, the estimates in (11.2) can be used to show that

$$\|\tilde{h}_{\mathbf{m}}\|_{L^1(\lambda_{\mathbb{R}^N};\mathbb{C})} \leq \left(\prod_{j=1}^{N} \big(2\pi(m_j + 1)\big) \right)^{\frac{1}{2}}, \ \|\tilde{h}_{\mathbf{m}}\|_{\mathrm{u}} \leq \left(\prod_{j=1}^{N} (m_j + 1) \right)^{\frac{1}{2}} \text{ and}$$

$$(16.1)$$

$$\|x_j \tilde{h}_{\mathbf{m}}\|_{\mathrm{u}} \vee \|\partial_{x_j} \tilde{h}_{\mathbf{m}}\|_{\mathrm{u}} \leq 2^N \prod_{j=1}^{N} (m_j + 1).$$

The Schwartz test function space $\mathscr{S}(\mathbb{R}^N; \mathbb{C})$ for $\mathbb{R}^N$ is defined as the space of $\varphi \in C^\infty(\mathbb{R}^N; \mathbb{C})$ with the property that $\|x_i^k \partial_{x_j}^\ell \varphi\|_{\mathrm{u}} < \infty$ for all $1 \leq i, j \leq N$ and $k, \ell \in \mathbb{N}$. Again one introduces the operators

$$\mathcal{H}^s \varphi = \sum_{\mathbf{k} \in \mathbb{N}^N} \mu_{\mathbf{k}}^s (\varphi, \tilde{h}_{\mathbf{k}})_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})} \tilde{h}_{\mathbf{k}}$$

and defines the norms

$$\|\varphi\|_{\mathrm{u}}^{(m)} = \max_{\substack{1 \leq i, j \leq N \\ k + \ell \leq m}} \|x_i \partial_{x_j} \varphi\|_{\mathrm{u}}$$

and

$$\|\varphi\|_{\mathscr{S}(\mathbb{R}; \mathbb{C})^{(m)}(\mathbb{R}^N; \mathbb{C})} = \sum_{\mathbf{k} \in \mathbb{N}^N} \mu_{\mathbf{k}}^m |(\varphi, \tilde{h}_{\mathbf{k}})_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})}|^2,$$

and the spaces

$$\mathscr{S}^{(m)}(\mathbb{R}^N; \mathbb{C}) = \{\varphi \in C^\infty(\mathbb{R}^N; \mathbb{C}) : \|\varphi\|_{\mathscr{S}(\mathbb{R}; \mathbb{C})^{(m)}} < \infty\}.$$

Clearly, if $\varphi \in \mathscr{S}^{(m)}(\mathbb{R}; \mathbb{C})$, then $\|\varphi\|_{\mathscr{S}(\mathbb{R}; \mathbb{C})^{(m)}} = \|\mathcal{H}^{\frac{m}{2}} \varphi\|_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})}$.

Using the estimates in (16.1) and the reasoning in Lemma 13.1 and Theorem 13.2, one sees that, for each m there is a $K_m \in (0, \infty)$ such that

$$\|\varphi\|_{\mathscr{S}^{(m)}(\mathbb{R}^N; \mathbb{C})} \leq K_m \|\varphi\|_{\mathrm{u}}^{(m+N)}$$

and

$$\|\varphi\|_{\mathrm{u}}^{(m)} \leq K_m \|\varphi\|_{\mathscr{S}^{(m+3N)}(\mathbb{R}^N; \mathbb{C})}.$$

Hence, $\mathscr{S}(\mathbb{R}^N; \mathbb{C}) = \bigcap_{m=0}^\infty \mathscr{S}^{(m)}(\mathbb{R}^N; \mathbb{C})$ and $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$ can be identified as the union $\bigcup_{m=0}^\infty \mathscr{S}^{(-m)}(\mathbb{R}^N; \mathbb{C})$ where $\mathscr{S}^{(-m)}(\mathbb{R}^N; \mathbb{C})$ is the analog for $N \geq 2$ of $\mathscr{S}^{(-m)}(\mathbb{R}; \mathbb{C})$ for $N = 1$. Further, the obvious analogs of Theorems 14.3 and 14.5 hold. In proving the analogs of Theorems 14.5 and 14.7, one needs to use the $\mathbb{R}^N$ version of Taylor's theorem which says that

$$\varphi(\mathbf{x}) = \sum_{m=0}^n \sum_{\|\mathbf{k}\|_1 = m} \frac{\partial^{\mathbf{k}} \varphi(0)}{\mathbf{k}!} \mathbf{x}^{\mathbf{k}} + \frac{1}{n!} \sum_{\|\mathbf{k}\|_1 = n+1} \binom{n+1}{\mathbf{k}} \mathbf{x}^{\mathbf{k}} \int_0^1 (1-t)^n \partial^{\mathbf{k}} \varphi(t\mathbf{x}) \, dt,$$

where $\mathbf{k}! = \prod_{j=1}^N k_j$, $\mathbf{x}^{\mathbf{k}} = \prod_{j=1}^N x_j^{k_j}$, $\partial^{\mathbf{k}} = \prod_{j=1}^N \partial_{x_j}^{k_j}$, and $\binom{n+1}{\mathbf{k}}$ is the multinomial coefficient $\frac{(n+1)!}{\mathbf{k}!}$.

Once one has the preceding, it should be clear how to extend a continuous map $A : \mathscr{S}(\mathbb{R}^N; \mathbb{C}) \longrightarrow \mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$ to continuous operators on $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$ if A admits an adjoint A^* which is a continuous operator on $\mathscr{S}(\mathbb{R}^N; \mathbb{C})$. In particular, both the Fourier transform and convolution have such extensions.

The extension of the Fourier transform to $L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})$ can be done as follows. First note that if $\varphi \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})$, then

$$\left(\varphi, \tilde{h}_{\mathbf{m}}\right)_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})} = \mu_{\mathbf{m}}^{-n} \left(\mathcal{H}^n \varphi, \tilde{h}_{\mathbf{m}}\right)_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})},$$

and therefore, using the first estimate in (16.1), one see that

$$\varphi_n \equiv \sum_{\|\mathbf{m}\|_1 \leq n} (\varphi, \tilde{h}_{\mathbf{m}})_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})} \tilde{h}_{\mathbf{m}} \longrightarrow \varphi$$

in $L^1(\lambda_{\mathbb{R}^N}; \mathbb{C})$ as well as $L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})$. Thus, $\|\hat{\varphi}_n - \hat{\varphi}\|_{\mathrm{u}} \longrightarrow 0$, and so

$$\hat{\varphi} = (2\pi)^{\frac{N}{2}} \sum_{m=0}^\infty \imath^{\|\mathbf{m}\|_1} (\varphi, \tilde{h}_{\mathbf{m}})_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})} \tilde{h}_{\mathbf{m}}$$

for $\varphi \in \mathscr{S}(\mathbb{R};\mathbb{C})$. Next suppose that $f \in L^1(\lambda_{\mathbb{R}^N};\mathbb{C}) \cap L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$, and set

$$\varphi_n = \sum_{\|m\|_1 \leq n} (f, \tilde{h}_{\mathbf{m}})_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})} \tilde{h}_{\mathbf{m}} \in \mathscr{S}(\mathbb{R}^N;\mathbb{C}).$$

Then $\varphi_n \longrightarrow f$ in $L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$, and so, by Fatou's lemma, one sees that $\|\hat{f}\|_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})} \leq (2\pi)^{\frac{N}{2}}\|f\|_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})}$. Hence the Fourier transform on $L^1(\lambda_{\mathbb{R}^N};\mathbb{C}) \cap L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$ admits a unique extension as a continuous operator on $L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$. In particular, for all $f \in L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$,

$$\hat{f}(\boldsymbol{\xi}) = \lim_{R\to\infty} \int_{|\mathbf{x}|\leq R} e^{\imath(\boldsymbol{\xi},\mathbf{x})_{\mathbb{R}^N}} f(\mathbf{x})\,d\mathbf{x},$$

where the convergence is in $L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$. Also

$$\hat{f} = (2\pi)^{\frac{N}{2}} \sum_{\mathbf{m}\in\mathbb{N}^N} \imath^{\|\mathbf{m}\|_1} \left(f, \tilde{h}_{\mathbf{m}}\right)_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})} \tilde{h}_{\mathbf{m}},$$

and so the Parseval identity

$$(\hat{f}, \hat{g})_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})} = (2\pi)^N (f, g)_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})}$$

holds for all $f, g \in L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$, from which the Fourier inversion formula $(\hat{f})^\vee = (2\pi)^N f = (\check{f})^\vee$ follows in the same way that it did when $N = 1$. Finally, by the same argument used when $N = 1$, one can show that $\widehat{\partial_{x_j} f} = -\imath\xi_j\hat{h}$ if $f \in L^2(\lambda_{\mathbb{R}^N};\mathbb{C}) \cap C^1(\mathbb{R}^N;\mathbb{C})$ and $\partial_{x_j} f \in L^2(\lambda_{\mathbb{R}^N};\mathbb{C})$.

To demonstrate the use these considerations, consider again the example discussed at the end of §15, only now its analog $\lambda u - \Delta u = \mu$ in $\mathbb{R}^N$, where $\lambda > 0$ and μ is a finite Borel measure on $\mathbb{R}^N$. Just as before, the Fourier transform of this equation lead to the conclusion that $\hat{u} = \frac{\hat{\mu}}{\lambda+|\boldsymbol{\xi}|^2}$. To find the function G_λ of which $(\lambda + |\boldsymbol{\xi}|^2)^{-1}$ is the Fourier transform, note that

$$\frac{1}{\lambda + |\boldsymbol{\xi}|^2} = \int_0^\infty e^{-t(\lambda+|\boldsymbol{\xi}|^2)}\,dt = \int_0^\infty e^{-\lambda t}\widehat{g_{2t}}(\boldsymbol{\xi})\,dt,$$

from which it follows that

$$G_\lambda(\mathbf{x}) = \int_0^\infty e^{-\lambda t} g_{2t}(\mathbf{x})\,dt = (4\pi)^{-\frac{1}{2}} \int_0^\infty t^{-\frac{N}{2}} e^{-\lambda t - \frac{|\mathbf{x}|^2}{4}}\,dt.$$

The function G_λ is a Bessel function, and a more explicit expression for it is easy to obtain only when N is odd. For example, when $N = 1$, we already knew that $G_\lambda(x) = \frac{1}{2\lambda^{-\frac{1}{2}}}e^{-\lambda^{\frac{1}{2}}|x|}$, and when $N = 3$, after differentiating (7.6) with respect to x, one sees that

$$G_\lambda(\mathbf{x}) = \frac{e^{-\lambda^{\frac{1}{2}}|\mathbf{x}|}}{2\pi|\mathbf{x}|}.$$

In any case, $G_\lambda \in L^1(\lambda_{\mathbb{R}};\mathbb{C})$, and it is clear that if a solution $u \in \mathscr{S}(\mathbb{R}^N;\mathbb{C})^*$ exists, then it is the function

$$\mathbf{x} \rightsquigarrow G_\lambda * \mu(\mathbf{x}) \equiv \int G_\lambda(\mathbf{x} - \mathbf{y})\,\mu(d\mathbf{y}). \tag{$*$}$$

Also, if the function $G_\lambda * \mu$ is an element of $\mathscr{S}(\mathbb{R}^N;\mathbb{C})^*$, for instance if

$$\int (1 + |\mathbf{x}|^2)^{-\frac{m}{2}} \left(\int G_\lambda(\mathbf{x} - \mathbf{y})\,\mu(d\mathbf{y}) \right) d\mathbf{x} < \infty,$$

then the function in $(*)$ determines the $u \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$ which is the one and only solution in $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$.

The *Poisson problem* $\Delta u = -\mu$ is a closely related to the preceding, but are two reasons why this problem is more difficult than the preceding one. The first reason is that, if one exists, then there is more than one solution. Indeed, if u is a solution and $\Delta v = 0$, then $u + v$ is again a solution. A v satisfying $\Delta v = 0$ is said to be *harmonic*, and there are lots of them. To see this, observe that $\Delta v = 0 \iff |\boldsymbol{\xi}|^2 \hat{v} =$ and that $|\boldsymbol{\xi}|^2 \hat{v} = 0$ implies that $\{\mathbf{0}\}$ is the support of $\hat{v}$, which by Theorem 14.5 means that it is a linear combination of $\delta_{\mathbf{0}}$ and its derivatives and therefore that v must be a polynomial. 14.5 means that $\hat{v}$ is a linear combination of derivatives of δ_0 and therefore that v is a polynomial. Thus, $v \in \mathscr{S}(\mathbb{R}; \mathbb{C})^*$ is harmonic harmonic if and only if $v = ax + b$. When $N \geq 2$, there are harmonic polynomials of all orders. For example, the real part of any complex polynomial will be a harmonic element of $\mathscr{S}(\mathbb{R}^2; \mathbb{C})$.

The second difficulty is that when $N \in \{1, 2\}$, $\frac{1}{|\boldsymbol{\xi}|^2} \notin \mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$, and therefore $\frac{\hat{\mu}}{|\boldsymbol{\xi}|^2}$ is not the Fourier transform of the convolution of u with an element of $L^1(\lambda_{\mathbb{R}}; \mathbb{C})$. Nonetheless, when $N = 1$ and $\int |y| \, \mu(dy) < \infty$, one can check by hand that if $G_0^{(1)}(x) = x^-$, then $u = G_0^{(1)} * \mu$ is an element of $\mathscr{S}(\mathbb{R}; \mathbb{C})^*$ which satisfies $\Delta u = -\mu$. When $N = 2$, one can use Green's formula and the divergence theorem to show that

$$\int \Delta \varphi(\mathbf{x}) \log |\mathbf{x} - \mathbf{y}| \, d\mathbf{x} = 2\pi \varphi(\mathbf{y})$$

for $\varphi \in \mathscr{S}(\mathbb{R}; \mathbb{C})$, and therefore, if $G_0^{(2)}(\mathbf{y}) \equiv -\frac{1}{2\pi} \log |\mathbf{y}|$ and there is an $m \geq 0$ for which

$$\int (1 + |\mathbf{x}|^2)^{-\frac{m}{2}} \left(\int \left| G_0^{(2)}(\mathbf{x} - \mathbf{y}) \right| \mu(d\mathbf{y}) \right) d\mathbf{x} < \infty,$$

then the function

$$\mathbf{x} \rightsquigarrow \int G_0^{(2)}(\mathbf{x} - \mathbf{y}) \, \mu(d\mathbf{y})$$

determines a solution $u \in \mathscr{S}(\mathbb{R}; \mathbb{C})(\mathbb{R}^2; \mathbb{C})^*$.

When $N \geq 3$, one should look for the tempered distribution of which $|\boldsymbol{\xi}|^{-2}$ is the Fourier transform. To that end, observe that

$$\frac{1}{|\boldsymbol{\xi}|^2} = \int_0^\infty e^{-t|\boldsymbol{\xi}|^2} \, dt = \int_0^\infty \widehat{g_{2t}}(\boldsymbol{\xi}) \, dt,$$

and so $|\boldsymbol{\xi}|^{-2}$ is the Fourier transform of

$$G_0(\mathbf{x}) = (4\pi)^{-\frac{N}{2}} \int_0^\infty t^{-\frac{N}{2}} e^{-\frac{|\mathbf{x}|^2}{4t}} \, dt = \frac{1}{4\pi^{\frac{N}{2}} |\mathbf{x}|^{N-2}} \int_0^\infty t^{\frac{N}{2}-2} e^{-t} \, dt = \frac{\Gamma\left(\frac{N-2}{2}\right)}{4\pi^{\frac{N}{2}} |\mathbf{x}|^{N-2}},$$

where Γ is Euler's gamma function. Because $\Gamma\left(\frac{N}{2}\right) = \frac{N-2}{2} \Gamma\left(\frac{N-2}{2}\right)$ and $\frac{2\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)}$ is the area ω_{N-1} of the unit sphere $\mathbb{S}^{N-1}$ in $\mathbb{R}^N$, we have that

$$G_0^{(N)}(\mathbf{x}) = \frac{1}{(N-2)\omega_{N-1} |\mathbf{x}|^{N-2}}.$$

Thus, if the function

$$\mathbf{x} \rightsquigarrow \int G_0^{(N)}(\mathbf{x} - \mathbf{y}) \, \mu(d\mathbf{y})$$

determines a $u \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$, then u is a solution.

The function $G_0^{(N)}$ is called the *Green's function* for the Laplacian in $\mathbb{R}^N$.

Exercise 16.1. Show that if f is an entire function on $\mathbb{C}$ (i.e., an analytic function there), then, as a function on $\mathbb{R}^2$ it is tempered distribution if and only if it is a polynomial. Conclude that if an entire function is not a polynomial, then it grows at infinity faster that any power of z.

17. Convergence of Probability Measures

Define $\mathbf{M}_1(\mathbb{R}^N)$ to be the set of Borel probability measures on $\mathbb{R}^N$. Clearly $\mathbf{M}_1(\mathbb{R}^N)$ is a convex subset of $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$, but it is a subset that possesses properties that are not shared by most other elements of $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$, and the topology of $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$ does not take full advantage of those properties. There are three stronger topologies that recommend themselves. Namely: the *uniform* topology, which is the one for which[8]

$$\|\mu - \nu\|_{\mathrm{var}} \equiv \sup\big\{ |\langle \varphi, \mu - \nu \rangle| : \varphi \text{ a Borel measurable function with } \|\varphi\|_{\mathrm{u}} = 1 \big\}$$

is the metric; the *strong* for which sets of the form

$$S(\mu, r; \varphi_1, \ldots, \varphi_n) = \big\{ \nu : |\langle \varphi_m, \nu - \mu \rangle| < r \text{ for } 1 \le m \le n \big\},$$

where φ_m's are bounded Borel measurable $\mathbb{R}$-valued functions on $\mathbb{R}^N$, are a neighborhood basis for μ; and the *weak* for which sets of the $S(\mu, r; \varphi_1, \ldots, \varphi_n)$ are a neighborhood basis for μ, only now with the restriction that φ_m's must be continuous as well as bounded.

Obviously, the strength of the uniform topology is greater than that of the strong topology, which is stronger than the weak topology, which, at first sight (cf. Exercise 17.1), looks stronger than the one which $\mathbf{M}_1(\mathbb{R}^N)$ inherits as a subset of $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$. Each of them has its virtues and flaws. The uniform topology admits a metric and is the strong topology on the dual space of the Banach space $C_0(\mathbb{R}^N; \mathbb{R})$ with the uniform topology; the strong topology is not separable and points don't have countable neighborhood bases; as we will show below, the weak topology is both separable and admits a metric. In addition, convergence of measures in the weak topology is intimately related to the convergence of their Fourier transforms.

In what follows, we will study some of the properties and applications of the weak topology.

Lemma 17.1. *The sets $S(\mu, r; \varphi_1, \ldots, \varphi_n)$ with $\varphi_1, \ldots, \varphi_n \in C_{\mathrm{c}}^\infty(\mathbb{R}^N; \mathbb{R})$ are a neighborhood basis at μ for the weak topology.*

Proof. We begin by showing if that $\varphi \in C_{\mathrm{b}}^\infty(\mathbb{R}; \mathbb{C})$ with $\|\varphi\|_{\mathrm{u}} = 1$ and $r > 0$, then there exist $\varphi_1, \varphi_2 \in C_{\mathrm{c}}^\infty(\mathbb{R}^N; \mathbb{C})$ such that

$$\big\{ \nu : |\langle \varphi_1, \nu - \mu \rangle| \vee |\langle \varphi_2, \nu - \mu \rangle| < \tfrac{r}{4} \big\} \subseteq \big\{ \nu : |\langle \varphi, \nu - \mu \rangle| < r \big\}.$$

To this end, choose $R > 0$ so that $\mu\big(B(\mathbf{0}, R)\big) > 1 - \tfrac{r}{4}$, and take $\eta \in C^\infty(\mathbb{R}^N; \mathbb{R})$ so that $\eta = 1$ on $\overline{B(\mathbf{0}, R)}$ and $\eta = 0$ off $B(\mathbf{0}, R + 1)$. Then

$$|\langle \varphi, \nu - \mu \rangle| \le |\langle \eta\varphi, \nu - \mu \rangle| + |\langle (1 - \eta)\varphi, \nu - \mu \rangle|$$

---

[8]We will continue to use $\langle \varphi, \mu \rangle$ to denote the integral with respect to μ of a function φ, even if $\varphi \notin \mathscr{S}(\mathbb{R}^N; \mathbb{C})$. Also, $\langle \varphi, \nu - \mu \rangle \equiv \langle \varphi, \nu \rangle - \langle \varphi, \mu \rangle$.

and

$$|\langle (1-\eta)\varphi, \nu - \mu \rangle| \leq \langle 1 - \eta, \mu \rangle + \langle 1 - \eta, \nu \rangle$$
$$\leq 2\langle 1 - \eta, \mu \rangle + |\langle 1 - \eta, \nu - \mu \rangle| = 2\langle 1 - \eta, \mu \rangle + |\langle \eta, \nu - \mu \rangle|.$$

Thus

$$|\langle (1-\eta)\varphi, \nu - \mu \rangle| \leq |\langle \eta\varphi, \nu - \mu \rangle| + 2\mu\big(B(\mathbf{0},R)^{\complement}\big) + |\langle \eta, \nu - \mu \rangle|,$$

and so

$$\big\{\nu : |\langle \eta\varphi, \nu - \mu \rangle| \vee |\langle \eta, \nu - \mu \rangle| < \tfrac{r}{4}\big\} \subseteq \{\nu : |\langle \varphi, \nu - \mu \rangle| < r\}.$$

In view of the preceding, it suffices to show that if $\varphi \in C_c(\mathbb{R}^N; \mathbb{C})$ with $\|\varphi\|_u = 1$ and $r > 0$, then there exists a $\psi \in C_c^\infty(\mathbb{R}^N; \mathbb{C})$ such

$$|\langle \psi, \nu - \mu \rangle| < \tfrac{r}{3} \implies |\langle \varphi, \nu - \mu \rangle| < r.$$

To this end, simply choose $\psi \in C_c^\infty(\mathbb{R}^N; \mathbb{C})$ so that $\|\varphi - \psi\|_u < \frac{r}{3}$, and check that this ψ will serve. $\qquad\square$

As Lemma 17.1 makes clear, what we are calling the weak topology on $\mathbf{M}_1(\mathbb{R}^N)$ is what a functional analyst would call the weak* topology on the dual space $C_0(\mathbb{R}^N; \mathbb{R})^*$ of the Banach space $C_0(\mathbb{R}^N; \mathbb{R})$ (the space of continuous functions that tend to 0 at infinity) with the uniform norm. Indeed, the Riesz representation theorem allows one to identify $C_0(\mathbb{R}^N; \mathbb{R})$ with the space of finite signed Borel measures on $\mathbb{R}^N$, and so $\mathbf{M}_1(\mathbb{R}^N)$ can be thought of as a convex subset of the unit ball in $C_0(\mathbb{R}^N; \mathbb{R})^*$, in which case Lemma 17.1 shows that the weak topology on $\mathbf{M}_1(\mathbb{R}^N)$ is the topology $\mathbf{M}_1(\mathbb{R}^N)$ inherits as a subset from the weak* topology on $C_0(\mathbb{R}^N; \mathbb{R})^*$.

Theorem 17.2. *The weak topology on* $\mathbf{M}_1(\mathbb{R}^N)$ *is a separable, metric topology.*

Proof. Let $\{\varphi_k : k \geq 1\}$ be a dense subset of $C_c(\mathbb{R}^N; \mathbb{R})$, and define

$$\rho(\mu, \nu) = \sum_{k=1}^\infty \frac{|\langle \varphi_k, \nu - \mu \rangle|}{2^k(1 + |\langle \varphi_k, \nu - \mu \rangle|)}.$$

Using Lemma 17.1, it is easy to check that φ is a metric for the weak topology on $\mathbf{M}_1(\mathbb{R}^N)$.

To prove separability, define D to be the set of measures $\sum_{m=1}^n a_m \delta_{\mathbf{x}_m}$, where $n \geq 1$, the a_m's are non-negative rational numbers whose sum is 1, and the $\mathbf{x}_m$'s are elements of $\mathbb{R}^N$ with rational coordinates. Clearly D is countable. Therefore it suffices to show that, for each $\mu \in \mathbf{M}_1(\mathbb{R}^N)$, each cellection $\{\varphi_1, \ldots, \varphi_\ell\} \subseteq C_b(\mathbb{R}^N; \mathbb{R})$, and $\epsilon > 0$, there is a $\nu \in D$ such that $\max_{1 \leq k \leq \ell} |\langle \varphi_k, \nu - \mu \rangle| < \epsilon$. Further, we need do so only for φ_k's and a μ which are supported on a ball $B(\mathbf{0}, R)$. Given such φ_k's and μ, choose $r > 0$ so that $\max_{1 \leq k \leq \ell} |\varphi_k(\mathbf{y}) - \varphi_k(\mathbf{x})| < \frac{\epsilon}{2}$ if $|\mathbf{y} - \mathbf{x}| < r$. Next, cover $\overline{B(\mathbf{0}, R)}$ with balls $B(\mathbf{x}_m, r)$, where $1 \leq m \leq n$, each $\mathbf{x}_m \in B(\mathbf{0}, R)$ and has rational coordinates, and define $A_1 = B(\mathbf{x}_1, r)$ and $A_m = B(\mathbf{x}_m, r) \setminus \bigcup_{k=1}^{m-1} A_k$ for $2 \leq m \leq n$. Finally, choose non-negative, rational numbers $a_1, \ldots, a_n$ so that

$$\max_{1 \leq k \leq \ell} \|\varphi_k\|_u \sum_{m=1}^n |a_m - \mu(A_m)| < \frac{\epsilon}{2}$$

and $\sum_{m=1}^{n} a_m = 1$, and take $\nu = \sum_{m=1}^{n} a_m \delta_{x_m}$. Then, for $1 \leq k \leq \ell$,

$$|\langle \varphi_k, \mu - \nu \rangle| \leq \sum_{m=1}^{n} \int_{A_m} |\varphi_k(\mathbf{x}) - \varphi_k(\mathbf{x}_m)| \, d\mu + \|\varphi_k\|_u \sum_{m=1}^{n} |\mu(A_m) - a_m| < \epsilon.$$

$\square$

We will use the notation $\mu_n \xrightarrow{\text{w}} \mu$ to mean that $\mu_n \longrightarrow$ in the weak topology on $\mathbf{M}_1(\mathbb{R}^N)$.

Theorem 17.3. *Given* $\{\mu_n : n \geq 1\} \cup \{\mu\} \subseteq \mathbf{M}_1(\mathbb{R}^N)$, *the following are equivalent:*

(i) $\mu_n \xrightarrow{\text{w}} \mu$.

(ii) $|\langle \varphi, \mu_n - \mu \rangle| \longrightarrow 0$ *for all* $\varphi \in C_c^\infty(\mathbb{R}^N; \mathbb{R})$.

(iii) *For all closed sets* $F \subseteq \mathbb{R}^N$, $\overline{\lim}_{n \to \infty} \mu_n(F) \leq \mu(F)$.

(iv) *For all open sets* $G \subseteq \mathbb{R}^N$, $\underline{\lim}_{n \to \infty} \mu_n(G) \geq \mu(G)$.

(v) *For all upper continuous functions* $f : \mathbb{R}^N \longrightarrow \mathbb{R}$ *that are bounded above,* $\overline{\lim}_{n \to \infty} \langle f, \mu_n \rangle \leq \langle f, \mu \rangle$.

(vi) *For all lower continuous functions* $f : \mathbb{R}^N \longrightarrow \mathbb{R}$ *that are bounded below,* $\underline{\lim}_{n \to \infty} \langle f, \mu_n \rangle \geq \langle f, \mu \rangle$.

Finally, if $\Gamma \in \mathcal{B}$ *and its boundary* $\partial \Gamma$ *has* μ-*measure* 0, *then* $\mu_n \xrightarrow{\text{w}} \mu \implies \mu(\Gamma) = \lim_{n \to \infty} \mu_n(\Gamma)$.

Proof. We already proved in Lemma 17.1 the equivalence of (i) and (ii), and the equivalence of (iii) and (iv) as well as that of (v) and (vi) is obvious. In addition, it is clear that (v) together with (vi) implies (i). Thus, we need only check that (i) implies (iii) and that (iv) implies (vi).

Assume that $\mu_n \xrightarrow{\text{w}} \mu$. Given a closed set F, define $\varphi_k(x) = 1 - \left(\frac{|x-F|}{1+|x-F|} \right)^{\frac{1}{k}}$. Then $\varphi_k \in C(\mathbb{R}^N; [0,1])$ and $\varphi_k \searrow \mathbf{1}_F$ as $k \to \infty$. Hence, for all k,

$$\langle \varphi_k, \mu \rangle = \lim_{n \to \infty} \langle \varphi_k, \mu_n \rangle \geq \overline{\lim}_{n \to \infty} \mu_n(F),$$

and so $\mu(F) = \lim_{k \to \infty} \langle \varphi_k, \mu \rangle \geq \overline{\lim}_{n \to \infty} \mu_n(F)$. Thus (i) $\implies$ (iii).

In proving that (iv) implies (vi), it suffices to handle f's which are positive as well as lower semicontinuous. Given such an f, define

$$f_k = \sum_{j=1}^{\infty} \frac{j \wedge 4^k}{2^k} \mathbf{1}_{I_{j,k}} \circ f = \frac{1}{2^k} \sum_{j=1}^{4^k} \mathbf{1}_{J_{j,k}} \circ f,$$

where

$$I_{j,k} = \left(\frac{j}{2^k}, \frac{j+1}{2^k} \right] \quad \text{and} \quad J_{j,k} = \left(\frac{j}{2^k}, \infty \right).$$

Then $0 \leq f_k \nearrow f$ as $k \to \infty$. In addition, because f is lower semicontinuous, the sets $G_{j,k} = \{x : f(x) \in J_{j,k}\}$ are open. Hence, if (iv) holds, then, for all k,

$$\langle f_k, \mu \rangle \leq \underline{\lim}_{n \to \infty} \langle f_k, \mu_n \rangle \leq \underline{\lim}_{n \to \infty} \langle f, \mu_n \rangle,$$

and so

$$\langle f, \mu \rangle = \lim_{k \to \infty} \langle f_k, \mu \rangle \leq \underline{\lim}_{n \to \infty} \langle f, \mu_n \rangle.$$

To prove the concluding assertion, assume $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$ and that $\mu(\partial\Gamma) = 0$. Set $G = \overset{\circ}{\Gamma}$ and $F = \bar{\Gamma}$. Then

$$\mu(\Gamma) = \mu(G) \leq \varliminf_{n\to\infty} \mu_n(G) \leq \varliminf_{n\to\infty} \mu_n(\Gamma)$$

and

$$\mu(\Gamma) = \mu(F) \geq \varlimsup_{n\to\infty} \mu_n(F) \geq \varlimsup_{n\to\infty} \mu_n(\Gamma),$$

and so $\mu(\Gamma) = \lim_{n\to\infty} \mu_n(\Gamma)$. $\qquad\square$

Another useful fact about weak convergence is the following.

Theorem 17.4. *Assume that $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$, let $\psi \in C\big(\mathbb{R}^N; [0,\infty)\big)$ be an element of $L^1(\mu; \mathbb{R})$ as well as of $\bigcap_{n=1}^\infty L^1(\mu_n; \mathbb{R})$. Then $\langle \psi, \mu \rangle \leq \varliminf_{n\to\infty} \langle \psi, \mu_n \rangle$. In addition, if $\{\varphi_n : n \geq 1\} \subseteq C(\mathbb{R}^N; \mathbb{R})$, $|\varphi_n| \leq \psi$ for all $n \geq 1$, and $\langle \psi, \mu_n \rangle \longrightarrow \langle \psi, \mu \rangle$, then $\langle \varphi_n, \mu_n \rangle \longrightarrow \langle \varphi, \mu \rangle$ if $\varphi_n \longrightarrow \varphi$ uniformly on compact subsets.*

Proof. Clearly,

$$\langle \psi \wedge R, \mu \rangle = \lim_{n\to\infty} \langle \psi \wedge R, \mu_n \rangle \leq \varliminf_{n\to\infty} \langle \psi, \mu_n \rangle$$

for all $R > 0$, and so $\langle \psi, \mu \rangle \leq \varliminf_{n\to\infty} \langle \psi, \mu_n \rangle$.

Now suppose that $\langle \psi, \mu_n \rangle \longrightarrow \langle \psi, \mu \rangle$, that $|\varphi_n| \leq \psi$, and that $\varphi_n \longrightarrow \varphi$ uniformly on compact subsets. Clearly

$$|\langle \varphi_n, \mu_n \rangle - \langle \varphi, \mu \rangle| \leq |\langle \varphi_n - \varphi, \mu_n \rangle| + |\langle \varphi, \mu - \mu_n \rangle|.$$

For each $R > 0$, choose $\eta_R \in C^\infty\big(\mathbb{R}^N; [0,1]\big)$ so that $\eta_R = 1$ on $\overline{B(\mathbf{0}, R)}$ and $\eta_R = 0$ off $B(\mathbf{0}, R+1)$. Then, for each $R > 0$,

$$\varlimsup_{n\to\infty} |\langle \varphi_n - \varphi, \mu_n \rangle|$$

$$\leq \varlimsup_{n\to\infty} \sup_{|x| \leq R+1} |\varphi_n(x) - \varphi(x)| \langle \eta_R, \mu_n \rangle + \varlimsup_{n\to\infty} |\langle (\mathbf{1} - \eta_R)(\varphi_n - \varphi), \mu_n \rangle|$$

$$\leq 2 \varlimsup_{n\to\infty} \langle (\mathbf{1} - \eta_R)\psi, \mu_n \rangle = 2\langle (\mathbf{1} - \eta_R)\psi, \mu \rangle,$$

and, by Lebesgue's dominated convergence theorem, the last expression tends to 0 as $R \to \infty$. Similarly, for all $R > 0$,

$$\varlimsup_{n\to\infty} |\langle \varphi, \mu_n - \mu \rangle|$$

$$\leq \varlimsup_{n\to\infty} |\langle \eta_R \varphi, \mu_n - \mu \rangle| + \varlimsup_{n\to\infty} \langle (\mathbf{1} - \eta_R)\psi, \mu_n \rangle + \langle (\mathbf{1} - \eta_R)\psi, \mu \rangle \leq 2\langle (\mathbf{1} - \eta_R)\psi, \mu \rangle,$$

and so $\varlimsup_{n\to\infty} |\langle \varphi, \mu_n - \mu \rangle| = 0$. $\qquad\square$

We will next investigate when a subset of $\mathbf{M}_1(\mathbb{R}^N)$ is relatively compact. Because the unit ball in the dual space of a Banach is compact in the weak* topology, a careless functional analyst might think that $\mathbf{M}_1(\mathbb{R}^N)$ is itself compact. However, although $\mathbf{M}_1(\mathbb{R}^N)$ is closed in the strong topology on $C_0(\mathbb{R}^N; \mathbb{R})^*$, it is *not* closed in the weak* topology. Indeed, the sequence $\{\delta_n : n \geq 1\} \subseteq \mathbf{M}_1(\mathbb{R})$ is weak* convergent to the measure whose total mass is 0, which is not an element of $\mathbf{M}_1(\mathbb{R})$. As this example indicates, in order for the weak* limit of a sequence $\{\mu_n : n \geq 1\} \subseteq \mathbf{M}_1(\mathbb{R}^N)$ to be in $\mathbf{M}_1(\mathbb{R}^N)$, one needs to know that the mass of the μ_n's is not

escaping to infinity. With that in mind, we will say that a subset A of $\mathbf{M}_1(\mathbb{R}^N)$ is *tight* if, for each $\epsilon \in (0, 1)$, there exists an $R \in [0, \infty)$ such that

$$\inf_{\mu \in A} \mu\big(\overline{B(\mathbf{0}, R)}\big) \geq 1 - \epsilon.$$

Theorem 17.5. *A subset $A \subseteq \mathbf{M}_1(\mathbb{R}^N)$ is relatively compact in the weak topology if and only if it is tight.*

Proof. Assume that A is tight, and let $\{\mu_n : n \geq 1\} \subseteq A$. As pointed out above, there is a subsequence of $\{\mu_n : n \geq 1\}$ which is weak* convergent in $C_0(\mathbb{R}^N; \mathbb{R})^*$ to a $\nu \in C_0(\mathbb{R}^N; \mathbb{R})^*$ which is a non-negative measure with total mass less than or equal to 1, and so, without loss in generality, we will assume that $\{\mu_n : n \geq 1\}$ is weak* convergent to ν. In order to check that $\nu(\mathbb{R}^N) = 1$, for any $\epsilon \in (0, 1)$ choose R so that $\inf_{n \geq 1} \mu_n\big(\overline{B(\mathbf{0}, R)}\big) \geq 1 - \epsilon$, and choose $\eta \in C\big(\mathbb{R}^N; [0, 1]\big)$ so that $\eta = 1$ on $\overline{B(\mathbf{0}, R)}$ and $\eta = 0$ off $B(\mathbf{0}, R + 1)$. Then

$$\nu(\mathbb{R}^N) \geq \nu\big(\overline{B(\mathbf{0}, R + 1)}\big) \geq \langle \eta, \nu \rangle = \lim_{n \to \infty} \langle \eta, \mu_n \rangle \geq \overline{\lim_{n \to \infty}} \, \mu_n\big(\overline{B(\mathbf{0}, R)}\big) \geq 1 - \epsilon,$$

and so $\nu(\mathbb{R}^N)$ must be 1.

Conversely, suppose that $A \subseteq \mathbf{M}_1(\mathbb{R}^N)$ is relatively compact in the weak topology. If A were not tight, then there would exist a $\theta \in [0, 1)$ and, for each $n \geq 1$, a $\mu_n \in A$ such that $\mu_n\big(\overline{B(\mathbf{0}, n)}\big) \leq \theta$, and, because A is relatively compact, we could assume that $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$ for some $\mu \in \mathbf{M}_1(\mathbb{R}^N)$. But if $\eta_m \in C\big(\mathbb{R}^N; [0, 1]\big)$ equals 1 on $\overline{B(\mathbf{0}, m)}$ and 0 off of $B(\mathbf{0}, m + 1)$, that would mean that, for all $m \geq 1$,

$$\mu\big(B(\mathbf{0}, m)\big) \leq \langle \eta_m, \mu \rangle = \lim_{n \to \infty} \langle \eta_m, \mu_n \rangle \leq \underline{\lim_{n \to \infty}} \, \mu_n\big(B(\mathbf{0}, n)\big) \leq \theta,$$

and so $\mu(\mathbb{R}^N)$ would have to be less than or equal to $\theta < 1$. $\square$

Exercise 17.1. Show that $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$ if and only if $\mu_n \longrightarrow \mu$ in $\mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$.

18. The Fourier Transform for $\mathbf{M}_1(\mathbb{R}^N)$

In many applications, it is important to know the relationship between the weak convergence of measures and convergence of their Fourier transforms, which are often called *characteristic functions* in the probability literature.

Theorem 18.1. *Given $\{\mu_n : n \geq 1\} \cup \{\mu\} \subseteq \mathbf{M}_1(\mathbb{R}^N)$, $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$ if and only if $\hat{\mu}_n(\boldsymbol{\xi}) \longrightarrow \hat{\mu}(\boldsymbol{\xi})$ for each $\boldsymbol{\xi} \in \mathbb{R}^N$. In fact, if $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$, then $\hat{\mu}_n \longrightarrow \hat{\mu}$ uniformly on compact subsets.*

Proof. Suppose that $\hat{\mu}_n \longrightarrow \hat{\mu}$ pointwise. Then, by Parseval's identity and Lebesgue's dominated convergence theorem, for each $\varphi \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})$,

$$(2\pi)^N \langle \varphi, \mu_n \rangle = \int \hat{\varphi}(\boldsymbol{\xi}) \hat{\mu}_n(-\boldsymbol{\xi}) \, d\boldsymbol{\xi} \longrightarrow \int \hat{\varphi}(\boldsymbol{\xi}) \hat{\mu}(-\boldsymbol{\xi}) \, d\boldsymbol{\xi} = (2\pi)^N \langle \varphi, \mu \rangle,$$

and so, by Theorem 17.3, $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$.

Now suppose that $\mu_n \overset{\mathrm{w}}{\longrightarrow} \mu$ and that $\boldsymbol{\xi}_n \longrightarrow \boldsymbol{\xi}$ in $\mathbb{R}^N$. Then the functions $\varphi_n(\mathbf{x}) = e^{\imath(\boldsymbol{\xi}_n, \mathbf{x})_{\mathbb{R}^N}}$ converge uniformly on compact subsets to the function $\varphi(\mathbf{x}) = e^{\imath(\boldsymbol{\xi}, \mathbf{x})}$, and therefore, by Theorem 17.4, $\hat{\mu}_n(\boldsymbol{\xi}_n) \longrightarrow \hat{\mu}(\boldsymbol{\xi})$. Hence $\hat{\mu}_n \longrightarrow \mu$ uniformly on compact subsets. $\square$

Undoubtedly the most famous application of Theorem 18.1 is to the derivation of the Central Limit Theorem in probability theory. The C.L.T. states that if $\{\mathbf{X}_n : n \geq 1\}$ is a sequence of $\mathbb{R}^N$-valued, mutually independent, uniformly square integrable random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ have the properties that their expected value is $\mathbf{0}$ and

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} \mathbb{E}\big[(\boldsymbol{\xi}, \mathbf{X}_m)^2_{\mathbb{R}^N}\big] = |\boldsymbol{\xi}|^2$$

for all $\boldsymbol{\xi} \in \mathbb{R}^N$, then the distribution σ_n of

$$\frac{\sum_{m=1}^{n} \mathbf{X}_m}{n^{\frac{1}{2}}}$$

converges weakly to γ^N, where $\gamma(dx) = (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} \, dx$ is the standard Gaussian measure on $\mathbb{R}$. To phrase this in analytic terms, let μ_m be the distribution of $\mathbf{X}_m$. Then the distribution of $\sum_{m=1}^{n} \mathbf{X}_m$ is the measure $\mu_1 * \cdots * \mu_n$, and so

$$\hat{\sigma}_n(\boldsymbol{\xi}) = \prod_{m=1}^{n} \hat{\mu}_m\big(\tfrac{\boldsymbol{\xi}}{n^{\frac{1}{2}}}\big)$$

is the Fourier transform of the distribution σ_n of $\frac{1}{n^{\frac{1}{2}}} \sum_{m=1}^{n} \mathbf{X}_m$. Next note that, by Taylor's theorem,

$$\hat{\mu}_m\big(\tfrac{\boldsymbol{\xi}}{n^{\frac{1}{2}}}\big)) = 1 + \tfrac{\imath}{n^{\frac{1}{2}}} \int (\boldsymbol{\xi}, \mathbf{x})_{\mathbb{R}^N} \, \mu_m(d\mathbf{x}) - \tfrac{1}{2n} \int (\boldsymbol{\xi}, \mathbf{x})^2 \, \mu_m(d\mathbf{x}) + o_m\big(\tfrac{1}{n}\big),$$

where, because the $\mathbf{X}_m$'s are uniformly square integrable,

$$\lim_{n \to \infty} n \sup_{m \geq 1} o_m\big(\tfrac{1}{n}\big) = 0.$$

Hence, because the $\mathbf{X}_m$ have expected value $\mathbf{0}$ and

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} \int (\boldsymbol{\xi}, \mathbf{x})^2_{\mathbb{R}^N} \, \mu_m(d\mathbf{x}) = |\boldsymbol{\xi}|^2,$$

one can use $|\log(1-t) - t| \leq t^2$ for $|t| \leq \frac{1}{2}$ to check that

$$\hat{\sigma}_n(\boldsymbol{\xi}) = \prod_{m=1}^{n} \left(1 - \frac{1}{2n} \int (\boldsymbol{\xi}, \mathbf{x})^2_{\mathbb{R}^N} \, \mu_m(d\mathbf{x}) + o_m\big(\tfrac{1}{n}\big)\right) \longrightarrow e^{-\frac{|\boldsymbol{\xi}|^2}{2}} = \widehat{\gamma^N}(\boldsymbol{\xi}).$$

In spite of Theorem 18.1, it is *not* true that a sequence of probability measures converges weakly just because their Fourier transform converge pointwise. The reason why is that if the sequence converges weakly, then it is relatively compact and therefore must be tight. The following theorem of P. Lévy shows how one can use Fourier transforms to test for tightness.

Theorem 18.2. (**Lévy's Continuity Theorem**) *If $A \subseteq \mathbf{M}_1(\mathbb{R}^N)$, then A is tight if and only if for each $\epsilon > 0$ there exists an $r > 0$ such that*

(18.1) $$\sup_{\substack{\mu \in A \\ |\boldsymbol{\xi}| \leq r}} \big|1 - \hat{\mu}(\boldsymbol{\xi})\big| \leq \epsilon.$$

Hence, $\{\mu_n : n \geq 1\} \subseteq \mathbf{M}_1(\mathbb{R}^N)$ is weakly convergent in $\mathbf{M}_1(\mathbb{R}^N)$ if and only if $\hat{\mu}_n$ converges uniformly in a neighborhood of $\mathbf{0}$, in which case there is a $\mu \in \mathbf{M}_1(\mathbb{R}^N)$ to which $\{\mu_n : n \geq 1\}$ is converging weakly.

Proof. Assume that A is tight and therefore relatively compact. To see that (18.1) hold, suppose it did not. Then there would be an $\epsilon > 0$ such that, for each $n \geq 1$, $\left|1 - \hat{\mu}_n(\boldsymbol{\xi}_n)\right| \geq \epsilon$ for some $\mu_n \in A$ and $\boldsymbol{\xi}_n \in B\left(\mathbf{0}, \frac{1}{n}\right)$. Because A is relatively compact, we could choose these μ_n so that they converge weakly to some $\mu \in \mathbf{M}_1(\mathbb{R}^N)$, in which case there would exist an $m \geq 1$ for which

$$|1 - \hat{\mu}(\boldsymbol{\xi})| \vee |\hat{\mu}_n(\boldsymbol{\xi}) - \hat{\mu}(\boldsymbol{\xi})| < \frac{\epsilon}{2}$$

when $n \geq m$ and $|\boldsymbol{\xi}| \leq \frac{1}{n}$, which would mean that $\epsilon \leq \left|1 - \hat{\mu}_n(\boldsymbol{\xi}_n)\right| < \epsilon$.

Now assume that (18.1) holds. To show that A must be tight, begin by observing that

$$|1 - \hat{\mu}(\boldsymbol{\xi})| \geq \int \left(1 - \cos(\boldsymbol{\xi}, \mathbf{y})_{\mathbb{R}^N}\right) \mu(d\mathbf{y}).$$

Therefore, if [9] $\mathbf{e} \in \mathbb{S}^{N-1}$, for all $r > 0$,

$$\frac{1}{r} \int_0^r \left|1 - \hat{\mu}(t\mathbf{e})\right| dt \geq \int_{\mathbb{R}^N \setminus \{\mathbf{0}\}} \left(1 - \frac{\sin\left(r(\mathbf{e}, \mathbf{y})_{\mathbb{R}^N}\right)}{r(\mathbf{e}, \mathbf{y})_{\mathbb{R}^N}}\right) \mu(d\mathbf{y}).$$

Now set

$$s(t) = \inf \left\{1 - \frac{\sin \tau}{\tau} : \tau \geq t\right\} \text{ for } t > 0.$$

Then $s(t) > 0$ for all $t > 0$, and, for all $R > 0$ and $\mathbf{e} \in \mathbb{S}^{N-1}$,

$$\sup_{t \in (0,r]} \left|1 - \hat{\mu}(t\mathbf{e})\right| \geq \frac{1}{r} \int_0^r \left|1 - \hat{\mu}(t\mathbf{e})\right| dt \geq s(rR)\mu\left(\{\mathbf{y} : |(\mathbf{e}, \mathbf{y})_{\mathbb{R}^N}| \geq R\}\right).$$

Since

$$\mu\left(\{\mathbf{y} : |\mathbf{y}| \geq R\}\right) \leq N \sup_{\mathbf{e} \in \mathbb{S}^{N-1}} \mu\left(\{\mathbf{y} : |(\mathbf{e}, \mathbf{y})_{\mathbb{R}^N}| \geq N^{-\frac{1}{2}}R\}\right),$$

we have the estimate

$$(18.2) \qquad \mu\left(\{\mathbf{y} : |\mathbf{y}| \geq R\}\right) \leq \frac{N}{s(rN^{-\frac{1}{2}}R)} \sup_{|\boldsymbol{\xi}| \leq r} |1 - \hat{\mu}(\boldsymbol{\xi})|.$$

Now let $\epsilon > 0$ be given, choose $r > 0$ so that $\sup_{|\boldsymbol{\xi}| \leq r} |1 - \hat{\mu}(\boldsymbol{\xi})| < \frac{s(1)}{N}$ for $\mu \in A$, and take $R = \frac{N^{\frac{1}{2}}}{r}$. Then

$$\sup_{\mu \in A} \mu\left(\{\mathbf{y} : |\mathbf{y}| \geq R\}\right) \leq \epsilon.$$

$\square$

Bochner found an interesting characterization of characteristic functions, one which is intimately related to Lévy's continuity theorem. To describe his result, say that a function $f : \mathbb{R}^N \longrightarrow \mathbb{C}$ is *non-negative definite* if the matrix

$$\left(\left(f(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k)\right)\right)_{1 \leq j,k \leq n}$$

is non-negative definite for all $n \geq 2$ and $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n \in \mathbb{R}^N$, which is equivalent to saying

$$\sum_{j,k=1}^n f(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k)\alpha_j \overline{\alpha_k} \geq 0$$

for all $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$.

---

[9]$\mathbb{S}^{N-1}$ is the unit sphere in $\mathbb{R}^N$.

Theorem 18.3. *A function $f : \mathbb{R}^N \longrightarrow \mathbb{C}$ is a characteristic function if and only if f is continuous, $f(0) = 1$, and f is non-negative definite.*

Proof. Assume that $f = \hat{\mu}$ for some $\mu \in M_1(\mathbb{R}^N)$. Then it is obvious that f is continuous and that $f(0) = 1$. To see that it is non-negative definite, observe that

$$\sum_{j,k=1}^{n} f(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k)\alpha_j\overline{\alpha_k} = \int \left(\sum_{j,k=1}^{n} e^{i(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k, \mathbf{x})_{\mathbb{R}^N}} \alpha_j\overline{\alpha_k} \right) \mu(d\mathbf{x})$$

$$= \int \left| \sum_{j,k=1}^{n} e^{i\boldsymbol{\xi}_j x}\alpha_j \right|^2 \mu(d\mathbf{x}) \geq 0.$$

Now assume that f is a continuous, non-negative definite function with $f(0) = 1$. Because

$$A \equiv \begin{pmatrix} 1 & f(\boldsymbol{\xi}) \\ f(-\boldsymbol{\xi}) & 1 \end{pmatrix}$$

is non-negative definite, $\mathfrak{Im}\big(f(\boldsymbol{\xi}) + f(-\boldsymbol{\xi})\big)$ and $\mathfrak{Im}\big(if(\boldsymbol{\xi}) - if(-\boldsymbol{\xi})\big)$ are both 0, and therefore $f(\boldsymbol{\xi}) = \overline{f(-\boldsymbol{\xi})}$. Thus A is Hermitian, and because it is non-negative definite, $1 - |f(\boldsymbol{\xi})|^2 \geq 0$. Therefore $|f(\boldsymbol{\xi})| \leq 1$. Next, let $\psi \in \mathscr{S}(\mathbb{R}^N; \mathbb{R})$, and use Riemann approximations to see that

$$\iint f(\boldsymbol{\xi} - \boldsymbol{\eta})\hat{\psi}(\boldsymbol{\xi})\overline{\hat{\psi}(\boldsymbol{\eta})} \, d\boldsymbol{\xi}d\boldsymbol{\eta} \geq 0.$$

Assume for the moment that f is in $L^1(\lambda_{\mathbb{R}^N}; \mathbb{C})$, and set

$$h(\mathbf{x}) = (2\pi)^{-N} \int e^{-i(\boldsymbol{\xi}, \mathbf{x})_{\mathbb{R}^N}} f(\boldsymbol{\xi}) \, d\boldsymbol{\xi}.$$

By Parseval's identity, Fubini's Theorem and the fact that $\overline{\hat{\psi}(\boldsymbol{\xi})} = \hat{\psi}(-\boldsymbol{\xi})$,

$$(2\pi)^N \int h(\mathbf{x})\psi(\mathbf{x})^2 \, d\mathbf{x} = \int f(\boldsymbol{\xi})\widehat{\psi^2}(-\boldsymbol{\xi}) \, d\boldsymbol{\xi} = \int f(\boldsymbol{\xi})(\hat{\psi} * \hat{\psi})(-\boldsymbol{\xi}) \, d\boldsymbol{\xi}$$

$$= \iint f(\boldsymbol{\xi})\overline{\hat{\psi}(\boldsymbol{\xi} + \boldsymbol{\eta})}\hat{\psi}(\boldsymbol{\eta}) \, d\boldsymbol{\xi}d\boldsymbol{\eta} = \iint f(\boldsymbol{\xi} - \boldsymbol{\eta})\overline{\hat{\psi}(\boldsymbol{\xi})}\hat{\psi}(\boldsymbol{\eta}) \, d\boldsymbol{\xi}d\boldsymbol{\eta} \geq 0.$$

Hence, since h is continuous, it follows that $h \geq 0$. In addition, by the Fourier inversion formula for $L^1(\lambda_{\mathbb{R}^N}; \mathbb{C})$,

$$1 = f(0) = \lim_{t \searrow 0} g_t * f(0) = \int e^{-\frac{t|\boldsymbol{\xi}|^2}{2}} h(\boldsymbol{\xi}) \, dx = \int h \, d\lambda_{\mathbb{R}^N},$$

and so f is the Fourier transform of the probability measure $d\mu = h \, d\lambda_{\mathbb{R}^N}$.

To remove the assumption that f is integrable, set $g_t(\mathbf{x}) = (2\pi t)^{-\frac{N}{2}} e^{-\frac{|\mathbf{x}|^2}{2t}}$ and define $\gamma_t(d\mathbf{x}) = g_t(\mathbf{x}) \, d\mathbf{x}$. Then $\widehat{\gamma_t}(\boldsymbol{\xi}) = e^{-\frac{t|\boldsymbol{\xi}|^2}{2}}$ and therefore $f_t \equiv \widehat{\gamma_t}f$ is a continuous, $\lambda_{\mathbb{R}^N}$-integrable function that is 1 at 0. To see that f_t is non-negative definite, note that

$$\sum_{j,k=1}^{n} f_t(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k)\alpha_j\overline{\alpha_k} = \sum_{j,k=1}^{n} f(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k)\alpha_j\overline{\alpha_k} \int e^{i(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k, \mathbf{x})_{\mathbb{R}^N}} \gamma_t(d\mathbf{x})$$

$$= \int \left(\sum_{j,k=1}^{n} f(\boldsymbol{\xi}_j - \boldsymbol{\xi}_k)\big(\alpha_j e^{i(\boldsymbol{\xi}_j, \mathbf{x})_{\mathbb{R}^N}}\big)\overline{\big(\alpha_k e^{i(\boldsymbol{\xi}_k, \mathbf{x})_{\mathbb{R}^N}}\big)} \right) \gamma_t(d\mathbf{x}) \geq 0.$$

Thus $f_t = \widehat{\mu_t}$ for some $\mu_t \in M_1(\mathbb{R}^N)$, and so, since $f_t \longrightarrow f$ uniformly on compact subsets, Lévy's continuity theorem implies that μ_t tends weakly to a $\mu \in M_1(\mathbb{R}^N)$ for which $f = \hat{\mu}$. $\qquad\square$

Because it is difficult to check whether a function is non-negative definite, it is the more or less trivial necessity part of Bochner's Theorem that turns out in practice to be more useful than the sufficiency conditions.

Exercise 18.1. Given $f \in C_{\mathrm{b}}(\mathbb{R}^N; \mathbb{C})$ with $f(0) = 1$, define the quadratic form

$$(\varphi, \psi)_f = \iint_{\mathbb{R}^N \times \mathbb{R}^N} \varphi(\boldsymbol{\xi}) f(\boldsymbol{\xi} - \boldsymbol{\eta}) \overline{\psi(\boldsymbol{\eta})} \, d\boldsymbol{\xi} d\boldsymbol{\eta}$$

for $\varphi, \psi \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})$. Show that this quadratic form is non-negative (i.e., $(\varphi, \varphi)_f \geq 0$) if and only if f is a characteristic function. Further, if $f = \hat{\mu}$, show that $(\varphi, \psi)_f = (\hat{\varphi}, \hat{\psi})_{L^2(\mu; \mathbb{C})}$ and therefore that $(\,\cdot\,, \,\cdot\,)_f$ is non-degenerate (i.e., $(\varphi, \varphi)_f = 0 \implies \varphi = 0$) if and only if $\mu(G) > 0$ for all non-empty open sets G.

Exercise 18.2. Here are some interesting facts about characteristic functions.

(**i**) It is easy to check that if $\mu \in \mathbf{M}_1(\mathbb{R}^N)$, then

$$|\hat{\mu}(\boldsymbol{\eta}) - \hat{\mu}(\boldsymbol{\xi})|^2 \leq 2\mathfrak{Re}\big(1 - \hat{\mu}(\boldsymbol{\eta} - \boldsymbol{\xi})\big),$$

and so, by Theorem 18.3, one sees that if f is a continuous, non-negative definite function for which $f(\mathbf{0}) = 1$, then $|f(\boldsymbol{\xi})| \leq 1$ and $|f(\boldsymbol{\eta}) - f(\boldsymbol{\xi})|^2 \leq 2\big(1 - \mathfrak{Re} f(\boldsymbol{\eta} - \boldsymbol{\xi})\big)$. Show that these inequalities hold even if one drops the continuity assumption.

Hint: Use the non-negative definiteness of the matrices

$$\begin{pmatrix} 1 & f(-\boldsymbol{\xi}) \\ f(\boldsymbol{\xi}) & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & f(-\boldsymbol{\xi}) & f(-\boldsymbol{\eta}) \\ f(\boldsymbol{\xi}) & 1 & f(\boldsymbol{\xi} - \boldsymbol{\eta}) \\ f(\boldsymbol{\eta}) & f(\boldsymbol{\eta} - \boldsymbol{\xi}) & 1 \end{pmatrix}$$

to see that $f(-\boldsymbol{\xi}) = \overline{f(\boldsymbol{\xi})}$ and that

$$|z|^2 - 2\bar{z}|f(\boldsymbol{\eta}) - f(\boldsymbol{\xi})| + 2\big(1 - \mathfrak{Re} f(\boldsymbol{\eta} - \boldsymbol{\xi})\big) \geq 0.$$

(**ii**) Without using Bochner's theorem, show that if f_1 and f_2 are non-negative definite functions, then so are $f_1 f_2$ and, for any $a, b \geq 0$, $af_1 + bf_2$ is also.

Hint: Show that if A and B are non-negative definite, Hermitian $N \times N$ matrices, then $\big((A_{k,\ell} B_{k,\ell})\big)_{1 \leq k, \ell \leq N}$ is also. One way to see this is to use the fact that B admits a square root.

(**iii**) Suppose that $f : \mathbb{R}^N \longrightarrow \mathbb{C}$ is a non-constant function for which $f(\mathbf{0}) = 1$. Show that if $\lim_{|\mathbf{x}| \searrow 0} \frac{1 - f(\mathbf{x})}{|\mathbf{x}|^2} = 0$, then f cannot be a characteristic function. In particular, if $\alpha > 2$, then $e^{-|\boldsymbol{\xi}|^\alpha}$ is not a characteristic function.

(**iv**) Given a finite signed Borel measure μ on $\mathbb{R}^N$, define

$$\hat{\mu}(\boldsymbol{\xi}) = \int e^{i(\boldsymbol{\xi}, \mathbf{x})_{\mathbb{R}^N}} \, \mu(d\mathbf{x}),$$

and show that $\hat{\mu} = 0$ if and only if $\mu = 0$.

Hint: Use the Hahn Decomposition Theorem to write μ as the difference of two, mutually singular, non-negative Borel measures on $\mathbb{R}^N$.

(**v**) Suppose that $f : \mathbb{R} \longrightarrow \mathbb{C}$ is a non-constant, twice continuously differentiable characteristic function. Show that $f''(0) < 0$ and that $\frac{f''}{f''(0)}$ is again a characteristic function. In addition, show that $\|f'\|_{\mathrm{u}}^2 \vee \|f''\|_{\mathrm{u}} \leq |f''(0)|$ and that $|f(\eta) - f(\xi)| \leq |f''(0)|^{\frac{1}{2}} |\eta - \xi|$.

(**vi**) Suppose that $\{\mu_n : n \geq 1\} \subseteq M_1(\mathbb{R})$ and that $f(\boldsymbol{\xi}) = \lim_{n \to \infty} \widehat{\mu_n}(\boldsymbol{\xi})$ exists for each $\boldsymbol{\xi} \in \mathbb{R}$. Show that f is a characteristic function if and only if it is continuous at $\mathbf{0}$, and notice that this provides an alternative proof of Theorem 18.2.

(**vii**) Let $\mu_n \in M_1(\mathbb{R})$ be the measure for which $\frac{d\mu_n}{d\lambda_{\mathbb{R}}} = (2n)^{-1}\mathbf{1}_{[-n,n]}$. Show that $\widehat{\mu_n} \longrightarrow \mathbf{1}_{\{0\}}$ pointwise, and conclude that $\{\mu_n : n \geq 1\}$ has no weak limits. This example demonstrates the essential role that continuity plays in Bochner's and Lévy's theorems.

19. Infinitely Divisible Probability Measures

The convolution product turns $\mathbf{M}_1(\mathbb{R}^N)$ into a commutative ring in which $\delta_{\mathbf{0}}$ is the identity. A $\mu \in \mathbf{M}_1(\mathbb{R}^N)$ is said to be *infinitely divisible* in this ring if, for each $n \geq 1$, there exists a $\mu_{\frac{1}{n}} \in \mathbf{M}_1(\mathbb{R}^N)$ such that

$$\mu = \mu_{\frac{1}{n}}^{*n} \equiv \underbrace{\mu_{\frac{1}{n}} * \cdots * \mu_{\frac{1}{n}}}_{n \text{ times}},$$

and the set $\mathcal{I}(\mathbb{R}^N)$ of infinitely divisible measures is an important source of building blocks for constructions in probability theory.

For probabililists, an element of $\mathcal{I}(\mathbb{R}^N)$ is the distribution of a random variable which, for each $n \geq 1$, can be written as the sum of n identically distributed random variables. Using commutativity, it is easy to check that set $\mathcal{I}(\mathbb{R}^N)$ of infinitely divisible measures is a subring of $\mathbf{M}_1(\mathbb{R}^N)$.

A famous theorem of Lévy and A. Khinchine describes the characteristic function of every element of $\mathcal{I}(\mathbb{R}^N)$. Namely, $\mu \in \mathcal{I}(\mathbb{R}^N)$ if and only if

(19.1)
$$\begin{aligned}
\hat{\mu}(\boldsymbol{\xi}) = \exp\Big(& i(\mathbf{b}, \boldsymbol{\xi})_{\mathbb{R}^N} - \frac{1}{2}\big(\boldsymbol{\xi}, A\boldsymbol{\xi}\big)_{\mathbb{R}^N} \\
& + \int \Big(e^{i(\boldsymbol{\xi},y)_{\mathbb{R}^N}} - 1 - i\mathbf{1}_{B(0,1)}(y)(\boldsymbol{\xi}, y)_{\mathbb{R}^N} \Big) M(dy) \Big),
\end{aligned}$$

for some $\mathbf{b} \in \mathbb{R}^N$, non-negative definite, symmetric $A \in \mathrm{Hom}(\mathbb{R}^N; \mathbb{R}^N)$, and Borel measure M on $\mathbb{R}^N$ such that $M(\{0\}) = 0$ and $\int \frac{|\mathbf{y}|^2}{1+|\mathbf{y}|^2} M(d\mathbf{y}) < \infty$. The expression in (19.1) is called the *Lévy–Khinchine formula*, a measure M satisfying the stated conditions is called a *Lévy measure*, and the triple $(\mathbf{b}, A, M)$ is called a *Lévy system*. It is clear that if the right hand side of (19.1) is a characteristic function for every Lévy system, then these are characteristic functions of infinitely divisible laws. Indeed, if μ corresponds to (b, A, M) and $\mu_{\frac{1}{n}}$ corresponds to $\big(\frac{\mathbf{b}}{n}, \frac{A}{n}, \frac{M}{n}\big)$, then $\hat{\mu} = (\widehat{\mu_{\frac{1}{n}}})^n$.

Proving that the function $f_{(\mathbf{b},A,M)}$ on the right hand side of (19.1) is a characteristic function is a relatively easy. To wit, $f_{(0,\mathbf{I},0)} = \hat{\gamma}$, where γ is the standard Gaussian measure on $\mathbb{R}^N$, and so it is easy to check that $f_{\mathbf{b},A,0}$ is the characteristic function of the distribution of $\mathbf{x} \rightsquigarrow \mathbf{b} + A^{\frac{1}{2}}\mathbf{x}$ under γ. Also, if the Lévy measure M

is finite and π_M is the Poisson measure given by

$$(19.2) \qquad\qquad \pi_M = e^{-M(\mathbb{R}^N)} \sum_{n=0}^{\infty} \frac{M^{*n}}{n!},$$

then

$$\widehat{\pi_M}(\boldsymbol{\xi}) = e^{-M(\mathbb{R}^N)} \sum_{n=0}^{\infty} \frac{\hat{M}(\boldsymbol{\xi})^n}{n!} = e^{-M(\mathbb{R}^N)+\hat{M}(\boldsymbol{\xi})} = \exp\left(\int \left(e^{\imath(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}} - 1\right) M(d\mathbf{y}),$$

and so $\widehat{\pi_M} = f_{(\mathbf{b}_M,0,M)}$, where $b_M = \int_{B(0,1)} \mathbf{y}\, M(d\mathbf{y})$. Hence, when M is finite, $f_{(\mathbf{b},A,M)}$ is the characteristic function of $\gamma_{\mathbf{b}-\mathbf{b}_M,A} * \pi_M$. Finally, for general Lévy measures M, set $M_k(d\mathbf{y}) = \mathbf{1}_{[\frac{1}{k},\infty)}(|\mathbf{y}|)M(dy)$. Then M_k is finite, and so $f_{(\mathbf{b},A,M_k)}$ is a characteristic function. Therefore, since $f_{(\mathbf{b},A,M_k)} \longrightarrow f_{(\mathbf{b},A,M)}$ uniformly on compact subsets, Theorem 18.2 says that $f_{(\mathbf{b},A,M)}$ is a characteristic function.

There are no easy proofs that the characteristic function of any $\mu \in \mathcal{I}(\mathbb{R}^N)$ is given by (19.1). The first step is to show that if $\mu \in \mathcal{I}(\mathbb{R}^N)$, then there is a unique $\ell \in C(\mathbb{R}^N;\mathbb{C})$ such that $\ell(\mathbf{0}) = 0$, $\frac{|\ell(\boldsymbol{\xi})|}{1+|\boldsymbol{\xi}|^2}$ is bounded, and $\hat{\mu}(\boldsymbol{\xi}) = e^{\ell(\boldsymbol{\xi})}$. Showing that ℓ exists and is unique comes down to showing that $\hat{\mu}$ never vanishes. To do that, choose $r > 0$ so that $|1 - \hat{\mu}(\boldsymbol{\xi})| \leq \frac{1}{2}$ when $|\boldsymbol{\xi}| \leq r$. Then there is an ℓ for which $\ell(0) = 0$, $|\ell(\boldsymbol{\xi})| \leq 2$, and $\hat{\mu}(\boldsymbol{\xi}) = e^{\ell(\boldsymbol{\xi})}$ if $|\boldsymbol{\xi}| \leq r$. Using $\log z = -\sum_{n=1}^{\infty} \frac{(1-z)^n}{n}$ when $|1 - z| < 1$, one sees that $|\ell(\boldsymbol{\xi})| \leq 2$ for $|\boldsymbol{\xi}| < r$.

Since $\widehat{\mu_{\frac{1}{n}}}(\xi)^n = \hat{\mu}(\boldsymbol{\xi})$, $\widehat{\mu_{\frac{1}{n}}}(\boldsymbol{\xi}) \neq 0$ when $|\boldsymbol{\xi}| \leq r$, and so, by uniqueness, it must be that $\widehat{\mu_{\frac{1}{n}}}(\boldsymbol{\xi}) = e^{\frac{\ell(\boldsymbol{\xi})}{n}}$ for $|\boldsymbol{\xi}| \leq r$, and therefore $|1 - \widehat{\mu_{\frac{1}{n}}}(\boldsymbol{\xi})| \leq \frac{2}{n}$ when $|\boldsymbol{\xi}| \leq r$. Hence, by (18.2), for any $R > 0$,

$$\mu_{\frac{1}{n}}\left(\{\mathbf{y} : |\mathbf{y}| \geq R\}\right) \leq \frac{2N}{ns(rN^{-\frac{1}{2}}R)},$$

and so

$$\left|1 - \widehat{\mu_{\frac{1}{n}}}(\boldsymbol{\xi})\right| \leq \int \left|1 - e^{\imath(\boldsymbol{\xi},\mathbf{y})}\right| \mu_{\frac{1}{n}}(d\mathbf{y}) \leq |\boldsymbol{\xi}|R + 2\mu_{\frac{1}{n}}\left(\{\mathbf{y} : |\mathbf{y}| \geq R\}\right) \leq |\boldsymbol{\xi}|R + \frac{2N}{ns(rN^{-\frac{1}{2}}R)}.$$

Given $\boldsymbol{\xi} \neq \mathbf{0}$, take $R = \frac{1}{4|\boldsymbol{\xi}|}$, choose n so that $\frac{2N}{ns(rN^{-\frac{1}{2}}R)} \leq \frac{1}{4}$, and conclude that $|1 - \widehat{\mu_{\frac{1}{n}}}(\boldsymbol{\xi})| \leq \frac{1}{2}$ and therefore $|\hat{\mu}(\boldsymbol{\xi})| \geq 2^{-n}$. This proves that $\hat{\mu}$ never vanishes and therefore that $\hat{\mu} = e^{\ell}$. In addition, by using the fact that $\lim_{t \searrow} \frac{s(t)}{t^2} = \frac{1}{6}$, the preceding line of reasoning shows that there is a $C < \infty$ such that $\left|1 - e^{\frac{\ell(\boldsymbol{\xi})}{n}}\right| \leq \frac{1}{2}$ when $n \geq C|\boldsymbol{\xi}|^2$, and therefore $\frac{|\ell(\boldsymbol{\xi})|}{1+|\boldsymbol{\xi}|^2}$ is bounded.

Knowing that $\widehat{\mu_{\frac{1}{n}}} = e^{\frac{\ell}{n}}$, one knows that

$$\ell(\xi) = \lim_{n\to\infty} n\left(\widehat{\mu_{\frac{1}{n}}}(\xi) - 1\right).$$

Thinking of ℓ as a tempered distribution, the challenge is to describe the distribution of which it is the Fourier transform. Thus, set $u = \check{\ell}$. Then, since ℓ has at most

quadratic growth,

$$(2\pi)^N\langle\varphi,u\rangle = \langle\hat{\varphi},\ell\rangle = \lim_{n\to\infty} n\int\hat{\varphi}(\boldsymbol{\xi})\left(\int\left(e^{-\imath(\boldsymbol{\xi},\mathbf{x})_{\mathbb{R}^N}}-1\right)\mu_{\frac{1}{n}}(d\mathbf{x})\right)d\boldsymbol{\xi}$$

$$= \lim_{n\to\infty} n\int\left(\int\left(e^{-\imath(\boldsymbol{\xi},\mathbf{x})_{\mathbb{R}^N}}-1\right)\hat{\varphi}(\boldsymbol{\xi}))d\boldsymbol{\xi}\right)\mu_{\frac{1}{n}}(d\mathbf{x})$$

$$= (2\pi)^N\lim_{n\to\infty} n\int\left(\varphi(\mathbf{x})-\varphi(\mathbf{0})\right)\mu_{\frac{1}{n}}(d\mathbf{x}),$$

and so

$$\langle\varphi,u\rangle = \lim_{n\to\infty} n\int\left(\varphi(\mathbf{x})-\varphi(\mathbf{0})\right)\mu_{\frac{1}{n}}(d\mathbf{x}).$$

In particular, u satisfies the obvious $\mathbb{R}^N$ analog of the minimum principle in (14.4). In addition, because $\ell(\mathbf{0})=0$ and $\bar{\tilde{\ell}}=\hat{\ell}$,

$$\langle\varphi_R,u\rangle = \int\varphi_R(\boldsymbol{\xi})\hat{\ell}(\boldsymbol{\xi})\,d\boldsymbol{\xi} = (2\pi)^N R\int\check{\varphi}(R\boldsymbol{\xi})\ell(\boldsymbol{\xi})\,d\boldsymbol{\xi}$$

$$= (2\pi)^N\int\check{\varphi}(\boldsymbol{\xi})\ell\left(R^{-1}\boldsymbol{\xi}\right)d\boldsymbol{\xi}\longrightarrow 0$$

as $R\to\infty$. Thus u satisfies the $\mathbb{R}^N$-analog of (14.5), and therefore, by the $\mathbb{R}^N$-analog of Theorem 14.7, we know that

$$\langle\varphi,u\rangle = \frac{1}{2}\sum_{i,j=1}^N A_{i,j}\partial_{x_i}\partial_{x_j}\varphi(\mathbf{0}) + \sum_{i=1}^N b_i\partial_{x_i}\varphi(0)$$

$$+ \int\left(\varphi(y)-\varphi(\mathbf{0})-\mathbf{1}_{B(0,1)}(y)\big(y,\nabla\varphi(\mathbf{0})\big)_{\mathbb{R}^N}\right)M(dy),$$

where $(\mathbf{b},A,M)$ is a Lévy system.

To compute the Fourier transform of u, introduce the operator

$$\mathcal{L}_{(\mathbf{b},A,M)}\varphi(\mathbf{x}) = \frac{1}{2}\sum_{i,j=1}^N A_{i,j}\partial_{x_i}\partial_{x_j}\varphi(\mathbf{x}) + \sum_{i=1}^N b_i\partial_{x_i}\varphi(\mathbf{x})$$

$$+ \int\left(\varphi(\mathbf{x}+\mathbf{y})-\varphi(\mathbf{x})-\big(\mathbf{b},\nabla\varphi(\mathbf{x})\big)_{\mathbb{R}^N}\right)M(d\mathbf{y}).$$

What we have shown is that $\langle\varphi,u\rangle = \mathcal{L}_{(\mathbf{b},A,M)}\varphi(\mathbf{0})$. Using $\widehat{\partial_{x_j}\varphi}(\boldsymbol{\xi})=-\imath\xi_j\hat{\varphi}(\boldsymbol{\xi})$ and Fubini's theorem, one sees that

$$\widehat{\mathcal{L}_{(\mathbf{b},A,M)}\varphi}(\boldsymbol{\xi}) = \hat{\varphi}(\boldsymbol{\xi})\ell_{(\mathbf{b},A,M)}(-\boldsymbol{\xi}),$$

where

$$\ell_{(\mathbf{b},A,M)}(\boldsymbol{\xi}) = \log f_{(\mathbf{b},A,M)}$$

$$= -\tfrac{1}{2}\big(\boldsymbol{\xi},A\boldsymbol{\xi}\big)_{\mathbb{R}^N} + \imath(\mathbf{b},\boldsymbol{\xi})_{\mathbb{R}^N} + \int\left(e^{\imath(\boldsymbol{\xi},\mathbf{y})}-1-\imath\mathbf{1}_{B(\mathbf{0},1)}(\mathbf{y})(\boldsymbol{\xi},\mathbf{y})_{\mathbb{R}^N}\right)M(d\mathbf{y}).$$

Hence, by Parseval's indentity,

$$\langle\hat{\varphi},\ell\rangle = (2\pi)^N\langle\varphi,u\rangle = (2\pi)^N\mathcal{L}_{(\mathbf{b},A,M)}(\mathbf{0}) = \langle\hat{\varphi},\ell_{(\mathbf{b},A,M)}(\boldsymbol{\xi})\rangle,$$

and so $\ell=\ell_{(\mathbf{b},A,M)}$.

We will now use (19.1) to prove some properties of the associated measures based on properties of the Lévy system. Use $\mu_{(\mathbf{b},A,M)}\in\mathscr{S}(\mathbb{R}^N;\mathbb{C})^*$ to denote

the probability measure of which $f_{(\mathbf{b},A,M)}$ is the Fourier transform, and set $\mu_t = \mu_{(t\mathbf{b},tA,tM)}$ for $t > 0$. Then

$$(2\pi)^N \partial_t \langle \varphi, \mu_t \rangle = \langle \hat{\varphi}, \ell_{(\mathbf{b},A,M)} f_{(t\mathbf{b},tA,tM)} \rangle = (2\pi)^N \langle \mathcal{L}_{(\mathbf{b},A,M)} \varphi, \mu_t \rangle.$$

That is, we have shown that

$$(19.3) \qquad \partial_t \langle \varphi, \mu_{(t\mathbf{b},tA,tM)} \rangle = \langle \mathcal{L}_{(\mathbf{b},A,M)} \varphi, \mu_{(t\mathbf{b},tA,tM)} \rangle.$$

Theorem 19.1. *If either A is non-degenerate or $M(G) > 0$ for all non-empty open sets $G \subseteq \mathbb{R}^N \setminus \{\mathbf{0}\}$, then $\mu_{(\mathbf{b},A,M)}(G) > 0$ for all non-empty open sets $G \subseteq \mathbb{R}^N$.*

Proof. First observe that $\mu_{(\mathbf{b},A,M)} = \delta_{\mathbf{b}} * \mu_{(\mathbf{0},A,M)}$, and therefore we can assume that $\mathbf{b} = \mathbf{0}$. Next note that $\mu_{(\mathbf{0},A,M)} = \gamma_A * \mu_{(\mathbf{0},0,M)}$ where γ_A is the distribution of $x \rightsquigarrow A^{\frac{1}{2}}x$ under γ, and so, if A is non-degenerate and therefore γ_A has a strictly positive density, $\mu_{(\mathbf{0},A,M)}$ does also.

Now assume that $\mathbf{b} = 0$, $A = 0$, and $M(G) > 0$ for all open $\emptyset \neq G \subseteq \mathbb{R}^N \setminus \{0\}$. Given an open $G \neq \emptyset$, choose an $\eta \in C^\infty(\mathbb{R}^N; [0,1])$ which is strictly positive on G and vanishes off of G. Then

$$\mathcal{L}_{(\mathbf{0},0,M)} \eta(\mathbf{x}) = \int \left(\eta(\mathbf{x}+\mathbf{y}) - \eta(\mathbf{x}) - \mathbf{1}_{B(\mathbf{0},1)}(\mathbf{y}) (\nabla \eta(\mathbf{x}), \mathbf{y})_{\mathbb{R}^N} \right) M(d\mathbf{y})$$

$$= \int \eta(\mathbf{x}+\mathbf{y}) \, M(d\mathbf{y}) > 0$$

if $\mathbf{x} \notin G$. Hence, if $f(t) = \langle \eta, \mu_{(\mathbf{0},0,tM)} \rangle$, then $f \geq 0$ and, by (19.3), $\mu_{(\mathbf{0},0,tM)}(G) = 0 \implies f'(t) > 0$. But $\mu_{(\mathbf{0},0,tM)}(G) = 0$ also implies that $f(t) = 0$, which, by the first derivative test, is possible only if $f'(t) = 0$. Hence $f(t) > 0$ for all $t > 0$, and so $\mu_{(\mathbf{0},0,M)}(G) > 0$. $\qquad \square$

Theorem 19.2. *If $N = 1$, then $\mu_{(b,A,M)}\big((-\infty,0)\big) = 0$ if and only if*

$$(19.4) \qquad A = 0, \; M\big((-\infty,0)\big) = 0, \; and \; \int_{|y|<1} y \, M(dy) \leq b.$$

Proof. Observe that, for $n \geq 1$,

$$\{\mathbf{x} \in \mathbb{R}^n : x_j < 0 \text{ for } 1 \leq j \leq n|\} \subseteq \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{j=1}^n x_j < 0 \right\},$$

and therefore $\mu_{\frac{1}{n}}\big((-\infty,0)\big)^n \leq \mu^{*n}\big((-\infty,0)\big)$ for any $\mu \in \mathbf{M}_1(\mathbb{R})$.

Now assume that $\mu_{(\mathbf{b},A,M)}\big((-\infty,0)\big) = 0$. Since $\mu_{(\mathbf{b},A,M)} = \gamma_A * \mu_{(b,0,M)}$ and $\gamma_A(G) > 0$ for all open $G \neq \emptyset$ unless $A = 0$, it follows that $A = 0$. Next observe that $f_{(b,0,M)}$ has a bounded analytic extension to $\{\zeta \in \mathbb{C} : \mathfrak{Re}\zeta < 0\}$, and therefore $M\big((-\infty,0)\big)$ must be 0. Finally, to prove the inequality in (19.4), set $\mu_{\frac{1}{n}} = \mu_{(\frac{b}{n},0,\frac{M}{n})}$. Since $\mu_1 = \mu_{\frac{1}{n}}^{*n}$, the observation above shows that $\mu_{\frac{1}{n}}\big((-\infty,0)\big) = 0$, and therefore, if $\varphi \geq 0$ on $[0,\infty)$ and $\varphi(0) = 0$, then, by (19.3),

$$\mathcal{L}_{(b,0,M)} \varphi(0) = \lim_{n \to \infty} n\big(\langle \varphi, \mu_{\frac{1}{n}} \rangle - \varphi(0)\big) \geq 0,$$

and so

$$b\varphi'(0) + \int \big(\varphi(y) - \mathbf{1}_{(-1,1)}(y) y \varphi'(0)\big) M(dy) \geq 0.$$

Now choose $\eta \in C^\infty\big(\mathbb{R}; [0,1]\big)$ so that $\eta = 1$ on $\big[-\frac{1}{2}, \frac{1}{2}\big]$ and $\eta = 0$ off $(-1,1)$, and, for $r \in (0,1)$, set $\varphi_r(x) = y\eta_r(y)$ where $\eta_r(y) = \eta\big(\frac{y}{r}\big)$. By the preceding applied to φ_r,

$$b - \int \big(\mathbf{1}_{(-1,1)}(y) - \eta_r(y)\big)y\, M(dy) \geq 0,$$

and so

$$\int_{(r,1)} y M(dy) \leq b \text{ for all } r \in (0,1).$$

Finally, assume that (19.4) holds, and set $M_r(dy) = \mathbf{1}_{[r,\infty)}(y)\, M(dy)$ and $b_r = b - \int y\, M_r(dy)$ for $r > 0$. Then (19.4) holds for $(b, 0, M_r)$ and (cf. (19.2)) $\mu_{(b,0,M_r)} = \delta_{b_r} * \pi_{M_r}$, from which it is clear that $\mu_{(b,0,M_r)}\big((-\infty, 0)\big) = 0$. Therefore, since $\mu_{(b,0,M_r)} \overset{\mathrm{w}}{\longrightarrow} \mu_{(b,0,M)}$, $\mu_{(b,0,M)}\big((-\infty, 0)\big) = 0$. $\qquad\square$

Exercise 19.1. If M is symmetric, show that the integral in (19.1) can be replaced by

$$\int \big(\cos(\boldsymbol{\xi}, \mathbf{y})_{\mathbb{R}^N} - 1\big) M(d\mathbf{y}).$$

If $M(\mathbf{y}) = |\mathbf{y}|^{-1-\alpha}$ for some $\alpha \in (0,2)$, show that

$$\int_{\mathbb{S}^{N-1}} \big(\cos(\boldsymbol{\xi}, \mathbf{y})_{\mathbb{R}^N} - 1\big) M(d\mathbf{y}) = |\boldsymbol{\xi}|^\alpha \int \big(\cos(\mathbf{e}, \mathbf{y})_{\mathbb{R}^N} - 1\big) d\mathbf{y},$$

for every $\mathbf{e} \in \mathbb{S}^{N-1}$. In particular, by combining this with part (**iii**) of Exercise 18.2, conclude that $e^{-|\xi|^\alpha}$ is a characteristic function if and only if $\alpha \in [0,2]$.

20. Singular Integral Operators

The classic *Poisson problem* is that of finding, for a given a function φ, a solution u to the equation $\Delta u = -\varphi$ in $\mathbb{R}^N$, and one of the questions that arises is determining how properties of the function φ are reflected by the solution u. In particular, one wants to know whether second order derivatives of u can be estimated in terms of φ. When $N = 1$, this problem doesn't arise because $-\varphi$ is the second derivative of u. However, when $N \geq 2$, it is not at all clear to what extent the entire Hessian matrix of u is controlled by its trace.

To address this question, it is best to begin by giving an integral representation of the solution u. Depending on dimension, u is given by

$$u(\mathbf{x}) = \int G_0^{(N)}(\mathbf{x} - \mathbf{y})\varphi(\mathbf{y})\, d\mathbf{y},$$

where $G_0^{(N)}$ is the (cf. §16) Green's function for the Laplacian in $\mathbb{R}^N$:

$$G_0^{(N)}(\mathbf{x}) = \begin{cases} \frac{1}{\pi}\log|\mathbf{x}| & \text{if } N = 2 \\ \frac{1}{(N-2)\omega_{N-1}|\mathbf{x}|^{N-2}} & \text{if } N \geq 3. \end{cases}$$

Thus

$$\partial_{x_i}\partial_{x_j} u(x) = \int G_{i,j}^{(N)}(x - y)\varphi(y)\, dy$$

where

(20.1) $$G_{i,j}^{(N)}(\mathbf{x}) = \frac{1}{\omega_{N-1}|\mathbf{x}|^N}\int\left(-\delta_{i,j} + N\frac{x_i x_j}{|\mathbf{x}|^2}\right).$$

Because $G_{i,j}^{(N)}$ is not an integrable function, one has take care when interpreting convolution with it. On the other hand, since $G_0^{(N)} \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})^*$, so is is $G_{i,j}^{(N)}$, and therefore $\varphi * G_{i,j}^{(N)}$ makes perfectly good sense when $\varphi \in \mathscr{S}(\mathbb{R}^N; \mathbb{C})$. The question then is whether, using this interpretation, one can derive estimates.

Before getting into the details, it is important to know what sort of estimates are possible. In particular, because $G_{i,j}^{(N)}$ is neither integrable nor bounded, one should not expect that convolution with it will map either $L^1(\lambda_{\mathbb{R}^N}; \mathbb{C})$ or $L^\infty(\lambda_{\mathbb{R}^N}; \mathbb{C})$ into itself. Even so, it turns out (cf. (24.2) below) that it maps $L^p(\lambda_{\mathbb{R}^N}; \mathbb{C})$ boundedly into itself when $p \in (1, \infty)$, and what follows is one way to prove that.

21. The Hilbert Transform

A key fact about $G_{i,j}^{(N)}$ is that it is a Borel measurable, homogeneous function of order N whose integral over $\mathbb{S}^{N-1}$ is 0. That is, it is a function of the form

$$k(\mathbf{x}) = \frac{\Omega(\mathbf{x})}{|\mathbf{x}|^N}$$

where $\Omega \restriction \mathbb{S}^{N-1} \in L^1(\lambda_{\mathbb{S}^{N-1}}; \mathbb{C})$ satisfies $\Omega(r\mathbf{x}) = \Omega(\mathbf{x})$ for all $r > 0$ and

$$\int_{\mathbb{S}^{N-1}} \Omega(\boldsymbol{\omega}) \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}) = 0.$$

A Calderòn–Zygmund kernel k determines a tempered distribution by the prescription

$$\langle \varphi, k \rangle = \lim_{r \searrow 0} \int_{|\mathbf{y}| \geq r} \varphi(\mathbf{y}) \bar{k}(\mathbf{y}) \, d\mathbf{y}$$

$$= \lim_{r \searrow 0} \int_{|\mathbf{y}| \geq r} \big(\varphi(\mathbf{y}) - \varphi(0) \mathbf{1}_{[-1,1]}(\mathbf{y}) \big) \bar{k}(\mathbf{y}) \, d\mathbf{y} = \int \big(\varphi(\mathbf{y}) - \varphi(0) \mathbf{1}_{[-1,1]}(\mathbf{y}) \big) \bar{k}(\mathbf{y}) \, d\mathbf{y}.$$

Such functions k are called *Calderòn–Zygmund* kernels because Calderòn and Zygmund were able to prove a large number of deep results about convolution with respect to them. In particular (cf. (23.2) below), they showed that, in great generality, for each $p \in (1, \infty)$ there is a constant C_p, depending on N and Ω, such that $\|\varphi * k\|_{L^p(\lambda_{\mathbb{R}^N}; \mathbb{C})} \leq C_p \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N}; \mathbb{C})}$.

When $N = 1$ there is, up to a multiple constant, only one C-K kernel, namely, the function $h(x) = \frac{1}{\pi x}$. Convolution with respect to h was studied originally by Hilbert and has been known as the *Hilbert transform* ever since. A seminal observation made by Hilbert is that, even though $h \notin L^1(\lambda_{\mathbb{R}}; \mathbb{C})$, this transform is a bounded mapping of $L^2(\lambda_{\mathbb{R}}; \mathbb{C})$ into itself. Indeed, thinking of h as a tempered distribution, we showed in (6.2) that $\hat{h}(\xi) = \imath \operatorname{sgn}(\xi)$. Thus, we know that $\|\varphi * h\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})} \leq \|\varphi\|_{L^2(\lambda_{\mathbb{R}}; \mathbb{C})}$.

In order to prove the estimate for $p \neq 2$, I will use an beautiful approach that I think was introduced by M. Riesz and is closely related to the ideas we used to compute $\hat{h}$. Recall the functions $p_y(x) = \frac{1}{\pi} \frac{y}{x^2 + y^2}$ and $q_y = \frac{1}{\pi} \frac{x}{x^2 + y^2}$ which are, respectively, the real and imaginary parts of $\frac{\imath}{z}$ when $z = x + \imath y$. Next, set $h_y(x) = \mathbf{1}_{[y, \infty)}(x) h(x)$, and observe that $\|h_y - q_y\|_{L^1(\lambda_{\mathbb{R}}; \mathbb{C})} = \|h_1 - q_1\|_{L^1(\lambda_{\mathbb{R}}; \mathbb{C})} \leq \frac{2}{\pi}$, and therefore $\|\varphi * h_y - \varphi * q_y\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})} \leq \frac{2}{\pi} \|\varphi\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})}$. Thus, showing that $\sup_{y > 0} \|\varphi * q_y\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})} \leq C_p \|\varphi\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})}$ for some $C_p < \infty$ will show that

$$\sup_{y > 0} \|\varphi * h_y\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})} \leq C_p \|\varphi\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})} \text{ for some other } C_p < \infty.$$

The advantage that q_y has over h_y is its connection to analytic functions. Namely, since $\frac{\imath}{z} = p_y(x) + \imath q_y(x)$ when $z = x + \imath y$,

$$f(z) \equiv \varphi * p_y(x) + \imath \varphi * q_y(x) = \frac{\imath}{\pi} \int \frac{\varphi(\xi)}{x + \imath y - \xi}\, d\xi.$$

Further, because $\|p_y\|_{L^p(\lambda_R;\mathbb{C})} = 1$, $\|\varphi * p_y\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \leq \|\varphi\|_{L^p(\lambda_R;\mathbb{C})}$, and Riesz's idea was to use these observations to control $\|\varphi * q_y\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})}$ in terms of $\|\varphi * p_y\|_{L^p(\lambda_R;\mathbb{C})}$. To do so he needed the fact that, for each $n \geq 1$ there exist finite constants A_n and B_n such that

$$(\mathfrak{Im}\zeta)^{2n} \leq A_n \mathfrak{Re}\zeta^{2n} + B_n(\mathfrak{Re}\zeta)^{2n} \text{ for } \zeta \in \mathbb{C}. \tag{$*$}$$

Proving $(*)$ comes down to showing that $\cos^{2n}\theta \leq A_n \cos 2n\theta + B_n \sin^{2n}\theta$ for $\theta \in [-\pi, \pi]$. Clearly, if $\theta \in \left[-\frac{\pi}{8n}, \frac{\pi}{8n}\right] \cup \left[\frac{7\pi}{8}, \frac{9\pi}{8}\right]$, A_n can be chosen so the $A_n \cos 2n\theta$ dominates $\cos^{2n}\theta$; and for θ not in those intervals, B_n can be chosen so that $B_n \sin^{2n}\theta$ dominates $\cos^{2n}\theta - A_n \cos 2n\theta$.

With the preceding at hand, we know that

$$\int \left(\varphi * q_y(x)\right)^{2n} dx \leq A_n \mathfrak{Re}\left(\int f(x + \imath y)^{2n}\, dx\right) dx + B_n \int \left(\varphi * p_y(x)\right)^{2n} dx.$$

What Riesz saw is that he could use Cauchy's theorem to prove that the integral of $x \rightsquigarrow f(x + \imath y)^{2n}$ is independent of $y > 0$. Indeed, consider the rectangle $\{z = x + \imath y : |x| \leq R \ \& \ y_1 \leq y \leq y_2\}$. Cauchy's theorem says that the contour integral of f^{2n} around the boundary is 0. In addition, since $\varphi \in \mathscr{S}(\mathbb{R}^2; \mathbb{C})$, as $R \to \infty$ the contribution to the integral from the vertical parts of the boundary tends to 0, and so the integrals over the horizontal parts are equal. Finally, as $y \nearrow \infty$, $\int f(x + \imath y)^{2n}\, dx \longrightarrow 0$, and so we now know that

$$\|\varphi * q_y\|_{L^{2n}(\lambda_{\mathbb{R}};\mathbb{C})} \leq B_n^{\frac{1}{2n}} \|\varphi\|_{L^{2n}(\lambda_{\mathbb{R}};\mathbb{C})}.$$

Hence, we have proved that, for each $n \geq 1$ there is a $C_{2n} < \infty$ such that

$$\sup_{y>0} \|\varphi * h_y\|_{L^{2n}(\lambda_{\mathbb{R}};\mathbb{C})} \leq C_{2n}\|\varphi\|_{L^{2n}(\lambda_{\mathbb{R}};\mathbb{C})}. \tag{21.1}$$

22. INTERPOLATION

Although (21.1) is already significant, one should suspect that a similar estimate holds for all $p \in (0, \infty)$, not just even integers. However, because Riesz needed f^p to be an analytic function, he needed p to be an integer; and because he needed $(\mathfrak{Re}f)^p$ to be non-negative, he needed it to be an even integer. It was to overcome this problem that he proved a powerful general result, known as an *interpolation* theorem, that can be viewed as an operator theoretic analog of Hölder's equality. The following version and proof of his result is due to G. Thorin.

Theorem 22.1. (**Riesz–Thorin**) *Given a σ-finite measure space $(E, \mathcal{F}, \mu)$ and numbers*

$$1 \leq p_0,\, p_1,\, q_0, q_1 \leq \infty \ \text{ with } p_0 \wedge p_1 < \infty,$$

assume that T is a linear operator on $L^{p_0}(\mu; \mathbb{C}) \cap L^{p_1}(\mu; \mathbb{C})$ into $L^{q_0}(\mu; \mathbb{C}) \cap L^{q_1}(\mu; \mathbb{C})$ satisfying

$$\|Tf\|_{L^{q_j}(\mu;\mathbb{C})} \leq M_j \|f\|_{L^{p_j}(\mu;\mathbb{C})} \ \text{ for } j \in \{0, 1\},$$

where $M_0 \vee M_1 < \infty$. Then, for each $\theta \in [0, 1]$

$$\|Tf\|_{L^{q_\theta}(\mu;\mathbb{C})} \leq M_1^{1-\theta} M_2^\theta \|f\|_{L^{p_\theta}(\mu;\mathbb{C})},$$

where $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$.

Thorin's proof of Theorem 22.1 requires to following simple version, due to Hadamard and known as the *three lines theorem*, of the Phragmen–Lindelöf theorem.

Lemma 22.2. *Suppose that F is a bounded continuous function on the closed strip $S = \{z \in \mathbb{C} : \mathfrak{Re}\, z \in [0,1]\}$ which is analytic on the interior of S. If $|F(\imath y)| \leq m_0$ and $|F(1 + \imath y)| \leq m_1$ for all $y \in \mathbb{R}$, then $|F(z)| \leq m_0^{1-x} m_1^x$ for $z = x + \imath y \in S$.*

Proof. By replacing F with $\frac{F(z)}{m_0^{1-z} m_1^z}$, one can reduce to the case when $m_0 = m_1 = 1$, in which case one needs to show that $|F(z)| \leq 1$ for $z \in S$. Thus we will assume that $m_0 = m_1 = 1$ and will prove that $|F| \leq 1$.

If $\lim_{|y| \to \infty} \sup_{x \in [0,1]} |F(x + \imath y)| = 0$, then the maximum principle for analytic functions says that

$$\sup_{\substack{z \in S \\ |\mathfrak{Im}\, z| \leq R}} |F(z)| = \sup\big\{|F(x + \imath y)| : (x,y) \in \big(\{0,1\} \times [-R,R]\big) \cup \big((0,1) \times \{-R,R\}\big)\big\}$$

$$\longrightarrow \sup_{y \in \mathbb{R}}\{|F(\imath y) \vee |F(1 + \imath y)|\} \leq 1.$$

Even if $F(x + \imath y)$ doesn't tend to 0 as $|y| \to \infty$, for each $n \geq 1$, the function $F_n(z) = e^{\frac{z^2-1}{n}} F(z)$ does. In addition, $|F_n(\imath y)| \vee |F_n(1 + \imath y)| \leq 1$, and so $|F_n(z)| \leq 1$. Now let $n \to \infty$. $\qquad\square$

Proof of Theorem 22.1. Without loss in generality, we will assume that $p_0 \leq p_1$. Also, q' will be used to denote the Hölder conjugate of $q \in [1, \infty]$.

The first step is to check that it suffices to prove that

$$\left|\int g(\xi) Tf(\xi)\, \mu(d\xi)\right| \leq M_0^{1-\theta} M_1^\theta \tag{$*$}$$

for simple functions f and g satisfying $\|f\|_{L^{p_\theta}(\mu;\mathbb{C})} = 1$ and $\|g\|_{L^{q_\theta'}(\mu;\mathbb{C})} = 1$. Indeed, $\|Tf\|_{L^{q_\theta}(\mu;\mathbb{C})}$ equals the supremum of $\left|\int gTf\, d\mu\right|$ over simple functions g with $\|g\|_{L^{q_\theta'}(\mu;\mathbb{C})} = 1$, and, if $p_1 < \infty$, then, for any $f \in L^{p_0}(\mu;\mathbb{C}) \cap L^{p_1}(\mu;\mathbb{C})$, we can choose simple function f_n such that $f_n \longrightarrow f$ both in $L^{p_0}(\mu;\mathbb{C})$ and in $L^{p_1}(\mu;\mathbb{C})$. Hence, if $(*)$ holds for simple functions, then, by Hölder's inequality,

$$\|Tf\|_{L^{q_\theta}(\mu;\mathbb{C})} \leq \|T(f_n - f)\|_{L^{q_\theta}(\mu;\mathbb{C})} + \|Tf_n\|_{L^{q_\theta}(\mu;\mathbb{C})}$$

$$\leq \|T(f_n - f)\|_{L^{q_0}(\mu;\mathbb{C})}^{1-\theta} \|T(f_n - f)\|_{L^{q_1}(\mu;\mathbb{C})}^{\theta} + M_0^{1-\theta} M_2^\theta \|f_n\|_{L^{p_\theta}(\mu;\mathbb{C})}$$

$$\leq M_0^{1-\theta} M_1^\theta \big(2\|f_n - f\|_{L^{p_0}(\mu;\mathbb{C})}^{1-\theta} \|f_n - f\|_{L^{p_1}(\mu;\mathbb{C})}^{\theta} + \|f\|_{L^{p_\theta}(\mu;\mathbb{C})}\big),$$

from which the required estimate follows when $n \to \infty$. When $p_1 = \infty$, one can choose the f_n's so that they converge to f in $L^{p_1}(\mu;\mathbb{C})$ and are uniformly bounded and thereby use the preceding argument to get the desired result.

Turning to the proof of $(*)$, let $\theta \in (0,1)$ and determine p and q by $\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$ and $\frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$. Next, define $p(z)$ and $q(z)$ for (cf. Lemma 22.2) $z \in S$ so that $\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}$ and $\frac{1}{q'(z)} = \frac{1-z}{q_0'} + \frac{z}{q_1'}$. Given simple functions

$$f = \sum_{m=1}^n a_m \mathbf{1}_{\Gamma_m} \quad \text{and} \quad g = \sum_{m=1}^n b_m \mathbf{1}_{\Delta_m} \text{ with } \|f\|_{L^p(\mu;\mathbb{C})} = 1 \text{ and } \|g\|_{L^{q'}(\mu;\mathbb{C})} = 1,$$

define $f_z = |f|^{\frac{p}{p(z)}} \frac{f}{|f|}$ and $g_z = |g|^{\frac{q'}{q'(z)}} \frac{g}{|g|}$, where $\frac{h(\boldsymbol{\xi})}{|h(\boldsymbol{\xi})|}$ is taken to be equal 0 if $h(\boldsymbol{\xi}) = 0$. Then

$$f_z = \sum_{m=1}^{n} |a_m|^{\frac{p}{p(z)}} \frac{a_m}{|a_m|} \mathbf{1}_{\Gamma_m} \text{ and } g_z = \sum_{m=1}^{n} |b_m|^{\frac{q'}{q'(z)}} \frac{b_m}{|b_m|} \mathbf{1}_{\Delta_m}.$$

Now define

$$F(z) = \int g_z(\boldsymbol{\xi}) T f_z(\boldsymbol{\xi}) \, \mu(d\boldsymbol{\xi}) = \sum_{k,\ell=1}^{n} |a_k|^{\frac{p}{p(z)}} \frac{a_k}{|a_k|} |b_\ell|^{\frac{q'}{q'(z)}} \frac{b_\ell}{|b_\ell|} \int_{\Delta_\ell} T\mathbf{1}_{\Gamma_k}(\boldsymbol{\xi}) \, \mu(d\boldsymbol{\xi}).$$

Then F is a bounded continuous function on S that is analytic function on the interior of S, and so, by Lemma 22.2,

$$|F(\theta)| \le m_0^{1-\theta} m_1^{\theta} \text{ where } m_0 = \sup_{y \in \mathbb{R}} |F(\imath y)| \text{ and } m_1 = \sup_{y \in \mathbb{R}} |F(1 + \imath y)|.$$

Thus, what remains is to check that $m_0 \le M_0$ and $m_1 \le M_1$. But, by Hölder's inequality,

$$|F(\imath y)| \le \|g_{\imath y}\|_{L^{q'_0}(\mu;\mathbb{C})} \|T f_{\imath y}\|_{L^{q_0}(\mu;\mathbb{C})} \le M_0 \|g_{\imath y}\|_{L^{q'_0}(\mu;\mathbb{C})} \|f_{\imath y}\|_{L^{p_0}(\mu;\mathbb{C})},$$

and

$$\|f_{\imath y}\|_{L^{p_0}(\mu;\mathbb{C})}^{p_0} = \sum_{m=1}^{n} \left| |a_m|^{\frac{p}{p(\imath y)}} \right|^{p_0} \mu(\Gamma_m) = \sum_{m=1}^{n} |a_m|^p \mu(\Gamma_m) = 1$$

Similarly

$$\|f_{1+\imath y}\|_{L^{p_1}(\mu;\mathbb{C})}^{p_1} = 1, \quad \|g_{\imath y}\|_{L^{q'_0}(\mu;\mathbb{C})}^{q'_0} = 1, \text{ and } \|g_{1+\imath y}\|_{L^{q'_1}(\mu;\mathbb{C})}^{q'_1} = 1.$$

$\square$

By combining (21.1) and Theorem 22.1, we know that there is a $C_p < \infty$ such $\sup_{y>0} \|\varphi * h_y\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \le C_p \|\varphi\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})}$ for each $p \in [2, \infty)$. To extend this result to $p \in (1, 2)$, observe that if $p \in (1, 2)$, then $p' \in (2, \infty)$. Hence, since

$$(\psi, \varphi * h_y)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})} = -(\psi * h_y, \varphi)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})},$$

we have that

$$|(\psi, \varphi * h_y)_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}| \le C_{p'} \|\psi\|_{L^{p'}(\lambda_{\mathbb{R}};\mathbb{C})} \|\varphi\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})}$$

and therefore that, for all $p \in (1, \infty)$,

$$(22.1) \qquad \sup_{y>0} \|\varphi * h_y\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})} \le C_p \|\varphi\|_{L^p(\lambda_{\mathbb{R}};\mathbb{C})},$$

where $C_p = C_{p'}$ when $p \in (1, 2)$.

Exercise 22.1. Note that $\|\hat{\varphi}\|_{L^2(\lambda_{\mathbb{R}^N};\mathbb{C})} = (2\pi)^{\frac{N}{2}} \|\varphi\|_{L^2(\lambda_{\mathbb{R}};\mathbb{C})}$ and $\|\hat{\varphi}\|_{L^\infty(\lambda_{\mathbb{R}^N};\mathbb{C})} \le \|\varphi\|_{L^1(\lambda_{\mathbb{R}^N};\mathbb{C})}$, and use Theorem 22.1 to prove that $\|\hat{\varphi}\|_{L^{p'}(\lambda_{\mathbb{R}^N};\mathbb{C})} \le (2\pi)^{\frac{N}{p'}} \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})}$ for $p \in [1, 2]$. Next, let $\psi \in L^p(\lambda_{\mathbb{R}^N};\mathbb{C})$ for some $p \in [1, \infty)$, and define $T\varphi = \varphi * \psi$. Remember that $\|T\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \le \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \|\psi\|_{L^1(\lambda_{\mathbb{R}^N};\mathbb{C})}$ and $\|T\varphi\|_{L^\infty(\lambda_{\mathbb{R}^N};\mathbb{C})} \le \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \|\psi\|_{L^{p'}(\lambda_{\mathbb{R}^N};\mathbb{C})}$, and use Theorem 22.1 to prove *Young's inequality*

$$\|\psi * \varphi\|_{L^r(\lambda_{\mathbb{R}^N};\mathbb{C})} \le \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \|\psi\|_{L^q(\lambda_{\mathbb{R}^N};\mathbb{C})} \text{ if } \frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1 \ge 0.$$

23. The Method of Rotations

Calderòn and Zygmund noticed that the Hilbert transform, and especially (22.1), can be used to prove the L^p boundedness of their kernels when $\Omega \in L^1(\lambda_{\mathbb{S}^{N-1}}; \mathbb{C})$ is odd (i.e., $\Omega(-\boldsymbol{\omega}) = -\Omega(\boldsymbol{\omega})$ for $\boldsymbol{\omega} \in \mathbb{S}^{N-1}$). For example, set $k_y(\mathbf{x}) = \mathbf{1}_{[y,\infty)}(|\mathbf{x}|)k(\mathbf{x})$ for $(\mathbf{x}, y) \in \mathbb{R}^N \times (0, \infty)$. Then because

$$\widehat{k_y}(\boldsymbol{\xi}) = \lim_{R \to \infty} \int_{y < |\mathbf{x}| \leq R} e^{\imath(\boldsymbol{\xi}, \mathbf{x})} k(\mathbf{x}) \, d\mathbf{x}$$

$$= \lim_{R \to \infty} \int_{\mathbb{S}^{N-1}} \Omega(\boldsymbol{\omega}) \left(\int_{(y,R]} e^{\imath r(\boldsymbol{\xi}, \boldsymbol{\omega})} \frac{1}{r} \, dr \right) \lambda_{\mathbb{S}^{N-1}} d\boldsymbol{\omega},$$

if Ω is odd, one has that

$$\widehat{k_y}(\boldsymbol{\xi}) = \lim_{R \to \infty} \frac{1}{2} \int_{\mathbb{S}^{N-1}} \Omega(\boldsymbol{\omega}) \left(\int_{y < |r| \leq R} e^{\imath r(\boldsymbol{\xi}, \boldsymbol{\omega})} \frac{1}{r} \, dr \right) \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega})$$

$$= \frac{\pi}{2} \int_{\mathbb{S}^{N-1}} \Omega(\boldsymbol{\omega}) \widehat{h_y}\big((\boldsymbol{\xi}, \boldsymbol{\omega})\big) \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}).$$

Hence,

(23.1) $$\widehat{k_y}(\boldsymbol{\xi}) = \frac{\imath\pi}{2} \int_{\mathbb{S}^{N-1}} \Omega(\boldsymbol{\omega}) \widehat{h_y}\big((\boldsymbol{\xi}, \boldsymbol{\omega})\big) \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}),$$

and so

$$\|\widehat{k_y}\|_{\mathrm{u}} \leq \frac{\pi \|\Omega\|_{L^1(\lambda_{\mathbb{S}^{N-1}}; \mathbb{C})} \|\widehat{h_y}\|_{\mathrm{u}}}{2}.$$

In particular, we already know that

$$\|\varphi * k\|_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})} \leq \frac{\pi \|\Omega\|_{L^1(\lambda_{\mathbb{S}^{N-1}}; \mathbb{C})}}{2} \|\varphi\|_{L^2(\lambda_{\mathbb{R}^N}; \mathbb{C})}.$$

The same trick as we just used allows us to prove estimates for general $p \in (1, \infty)$. Namely, again using the oddness of k, one can first write

$$\varphi * k_y(\mathbf{x}) = \frac{1}{2} \int_{\mathbb{S}^{N-1}} \Omega(\boldsymbol{\omega}) \left(\int_{|r| > y} \varphi(\mathbf{x} - r\boldsymbol{\omega}) \frac{dr}{r} \right) \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}),$$

and then, after applying Minkowski's inequality,

$$\|\varphi * k_\epsilon\|_{L^p(\lambda_{\mathbb{R}^N}; \mathbb{C})} \leq \frac{1}{2} \int_{\mathbb{S}^{N-1}} |\Omega(\boldsymbol{\omega})| \left(\int_{\mathbb{R}^N} \left| \int_{|r| > y} \varphi(\mathbf{x} - r\boldsymbol{\omega}) \frac{dr}{r} \right|^p d\mathbf{x} \right)^{\frac{1}{p}} \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}).$$

Finally, for fixed $\mathbf{e} \in \mathbb{S}^{N-1}$, choose Euclidean coordinates for $\mathbb{R}^N$ so that $\mathbf{e}$ points in the direction of the first coordinate. Then

$$\int_{\mathbb{R}^N} \left| \int_{|r| > y} \varphi(\mathbf{x} - r\mathbf{e}) \frac{dr}{r} \right|^p d\mathbf{x}$$

$$= \pi^p \int_{\mathbb{R}^{N-1}} \cdots \int \left(\int_{\mathbb{R}} \left| [\varphi * h_y(\cdot, x_2, \ldots, x_N)](x_1) \right|^p dx_1 \right) dx_2 \cdots dx_N$$

$$\leq (\pi C_p)^p \int_{\mathbb{R}^{N-1}} \cdots \int \|\varphi(\cdot, x_2, \ldots, x_N)\|_{L^p(\lambda_{\mathbb{R}}; \mathbb{C})}^p dx_2 \cdots dx_N = (\pi C_p)^p \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N}; \mathbb{C})}^p,$$

which, together with the preceding, leads immediately to

(23.2) $$\|\varphi * k\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \le k_p \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \text{ for } p \in (1,\infty),$$

where $K_p = \pi C_p \|\Omega\|_{L^1(\lambda_{\mathbb{S}^{N-1}};\mathbb{C})}$.

24. The Riesz Kernels

In a sense which can be made very precise, the basic C-Z kernels for $\mathbb{R}^N$ are the *Riesz kernels* $r_i(\mathbf{x}) = c_N \frac{x_i}{|\mathbf{x}|^{N+1}}$, $1 \le i \le N$, where

$$c_N \equiv \left(\frac{\pi}{2} \int_{\mathbb{S}^{N-1}} |\omega_1| \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}) \right)^{-1}.$$

Obviously, the preceding applies to each of these. To get a feeling for how convolution with respect to r_i acts, apply (23.1) to see that

$$\widehat{r_i}(\boldsymbol{\xi}) = \frac{\imath \pi c_N}{2} \int_{\mathbb{S}^{N-1}} \omega_i \mathrm{sgn}\big((\boldsymbol{\xi}, \boldsymbol{\omega})\big) \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}).$$

Certainly, $\widehat{r_i}$ is homogeneous of degree 0, and so we need only worry about $\boldsymbol{\xi} \in \mathbb{S}^{N-1}$. Given $\boldsymbol{\xi} \in \mathbb{S}^{N-1}$, write $\boldsymbol{\omega} = (\boldsymbol{\omega}, \boldsymbol{\xi})\boldsymbol{\xi} + \boldsymbol{\omega}^{\perp \boldsymbol{\xi}}$. Then

$$\int_{\mathbb{S}^{N-1}} \omega_i \mathrm{sgn}\big((\boldsymbol{\xi}, \boldsymbol{\omega})\big) \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega})$$

$$= \xi_i \int_{\mathbb{S}^{N-1}} |(\boldsymbol{\omega}, \boldsymbol{\xi})| \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}) + \int_{\mathbb{S}^{N-1}} \big(\boldsymbol{\omega}^{\perp \boldsymbol{\xi}}\big)_i \mathrm{sgn}\big((\boldsymbol{\xi}, \boldsymbol{\omega})\big) \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}).$$

Because the integrand in the second term is an odd function of $\boldsymbol{\omega} \rightsquigarrow (\boldsymbol{\xi}, \boldsymbol{\omega})$, the second term vanishes. Hence,

(24.1) $$\widehat{r_i}(\boldsymbol{\xi}) = \frac{\imath \xi_i}{|\boldsymbol{\xi}|}, \quad \boldsymbol{\xi} \in \mathbb{R}^N \setminus \{0\}.$$

To evaluate c_N, observe that $c_1 = \frac{1}{\pi}$ is trivial. When $N \ge 2$, use

$$\int_{\mathbb{S}^{N-1}} |\omega_1| \, \lambda_{\mathbb{S}^{N-1}}(d\boldsymbol{\omega}) = \omega_{N-2} \int_{(-1,1)} |\rho| \big(1 - \rho^2\big)^{\frac{N-3}{2}} \, d\rho$$

$$= \omega_{N-2} \int_{(0,1)} (1-t)^{\frac{N-3}{2}} \, dt = \frac{2\omega_{N-2}}{N-1} = 2\Omega_{N-1},$$

where Ω_{N-1} is the volume to the unit ball in $\mathbb{R}^{N-1}$, and so $c_N = \frac{1}{\pi \Omega_{N-1}}$.

From the Riesz transforms one can build other kernels. For instance, recall the kernels in (20.1). Because $\partial_{x_i} \partial_{x_j} \varphi = -(\Delta\varphi) * G_{i,j}^{(N)}$, $-\xi_i \xi_j \hat{\varphi} = |\boldsymbol{\xi}|^2 \widehat{G_{i,j}^{(N)}} \hat{\varphi}$, and so

$$\widehat{G_{i,j}^{(N)}}(\boldsymbol{\xi}) = -\frac{\xi_i \xi_j}{|\boldsymbol{\xi}|^2} = -\widehat{r_i}(\boldsymbol{\xi})\widehat{r_j}(\boldsymbol{\xi}).$$

Hence, $\varphi * G_{i,j}^{(N)} = -(\varphi * r_i) * r_j$, and so

(24.2) $$\|\varphi * G_{i,j}^{(N)}\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \le K_p^2 \|\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \text{ for } p \in (1,\infty).$$

Equivalently, we now know that

$$\|\partial_{x_i} \partial_{x_j} \varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \le K_p^2 \|\Delta\varphi\|_{L^p(\lambda_{\mathbb{R}^N};\mathbb{C})} \text{ for } p \in (1,\infty).$$