VIII. Fourier Transform in Euclidean Space, 411-447

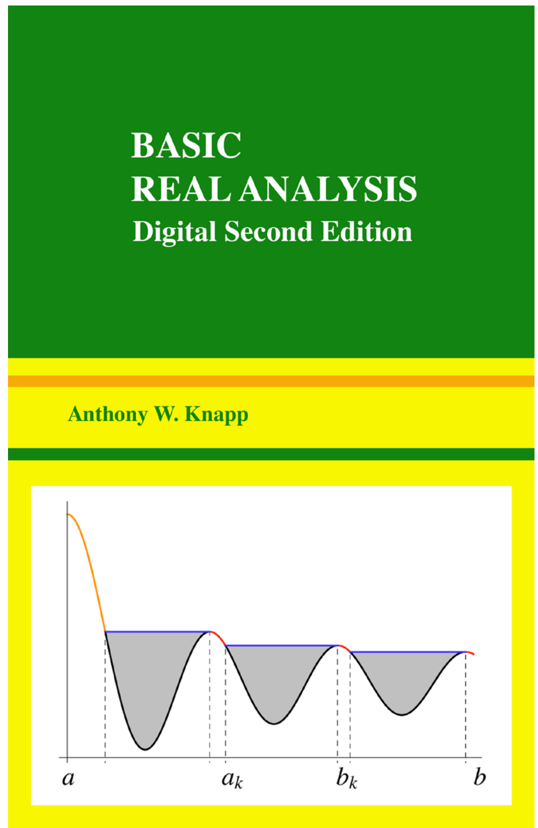
DOI: 10.3792/euclid/9781429799997-8

from

Basic Real Analysis
Digital Second Edition

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799997
ISBN: 978-1-4297-9999-7



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/~aknapp

Title: Basic Real Analysis, with an appendix "Elementary Complex Analysis"
Cover: An instance of the Rising Sun Lemma in Section VII.1.

Mathematics Subject Classification (2010): 28–01, 26–01, 42–01, 54–01, 34–01, 30–01, 32–01.

First Edition, ISBN-13 978-0-8176-3250-2

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the author

CHAPTER VIII

Fourier Transform in Euclidean Space

Abstract. This chapter develops some of the theory of the $\mathbb{R}^N$ Fourier transform as an operator that carries certain spaces of complex-valued functions on $\mathbb{R}^N$ to other spaces of such functions.

Sections 1–3 give the indispensable parts of the theory, beginning in Section 1 with the definition, the fact that integrable functions are mapped to bounded continuous functions, and various transformation rules. In Section 2 the main results concern L^1, chiefly the vanishing of the Fourier transforms of integrable functions at infinity, the fact that the Fourier transform is one-one, and the all-important Fourier inversion formula. The third section builds on these results to establish a theory for L^2. The Fourier transform carries functions in $L^1 \cap L^2$ to functions in L^2, preserving the L^2 norm; this is the Plancherel formula. The Fourier transform therefore extends by continuity to all of L^2, and the Riesz–Fischer Theorem says that this extended mapping is onto L^2. These results allow one to construct bounded linear operators on L^2 commuting with translations by multiplying by L^∞ functions on the Fourier transform side and then using Fourier inversion; a converse theorem is proved in the next section.

Section 4 discusses the Fourier transform on the Schwartz space, the subspace of L^1 consisting of smooth functions with the property that the product of any iterated partial derivative of the function with any polynomial is bounded. The Fourier transform carries the Schwartz space in one-one fashion onto itself, and this fact leads to the proof of the converse theorem mentioned above.

Section 5 applies the Schwartz space in $\mathbb{R}^1$ to obtain the Poisson Summation Formula, which relates Fourier series and the Fourier transform. A particular instance of this formula allows one to prove the functional equation of the Riemann zeta function.

Section 6 develops the Poisson integral formula, which transforms functions on $\mathbb{R}^N$ into harmonic functions on a half space in $\mathbb{R}^{N+1}$. A function on $\mathbb{R}^N$ can be recovered as boundary values of its Poisson integral in various ways.

Section 7 specializes the theory of the previous section to $\mathbb{R}^1$, where one can associate a "conjugate" harmonic function to any harmonic function in the upper half plane. There is an associated conjugate Poisson kernel that maps a boundary function to a harmonic function conjugate to the Poisson integral. The boundary values of the harmonic function and its conjugate are related by the Hilbert transform, which implements a "90° phase shift" on functions. The Hilbert transform is a bounded linear operator on L^2 and is of weak type (1, 1).

1. Elementary Properties

Although the Fourier transform in the one-variable case dates from the early nineteenth century, it was not until the introduction of the Lebesgue integral early in the twentieth century that the theory could advance very far. Fourier series in one variable have a standard physical interpretation as representing a resolution into component frequencies of a periodic signal that is given as a function of time. In the presence of the Riesz–Fischer Theorem, they are especially handy at analyzing time-independent operators on signals, such as those given by filters. An operator of this kind takes a function f with Fourier series $f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$ into the expression $\sum_{n=-\infty}^{\infty} m_n c_n e^{inx}$, where the constants m_n depend only on the filter. If the original function f is in L^2 and if the constants m_n are bounded, the Riesz–Fischer Theorem allows one to interpret the new series as the Fourier series of a new L^2 function $T(f)$, and thus the effect of the filter is to carry f to $T(f)$.

If one imagines that the period is allowed to increase without limit, one can hope to obtain convergence of some sort to a transform that handles aperiodic signals, and this was once a common attitude about how to view the Fourier transform. In the twentieth century the Fourier transform began to be developed as an object in its own right, and soon the theory was extended from one variable to several variables.

The Fourier transform in Euclidean space $\mathbb{R}^N$ is a mapping of suitable kinds of functions on $\mathbb{R}^N$ to other functions on $\mathbb{R}^N$. The functions will in all cases now be assumed to be complex valued. The underlying $\mathbb{R}^N$ is usually regarded as space, rather than time, and the Fourier transform is of great importance in studying operators that commute with translations, i.e., spatially homogeneous operators. One example of such an operator is a linear partial differential operator with constant coefficients, and another is convolution with a fixed function. In the latter case if $\mathcal{F}$ denotes the Fourier transform and h is a fixed function, the relevant formula is $\mathcal{F}(h * f) = \mathcal{F}(h)\mathcal{F}(f)$, the product on the right side being the pointwise product of two functions. Thus convolution can be understood in terms of the simpler operation of pointwise multiplication if we understand what $\mathcal{F}$ does and we understand how to invert $\mathcal{F}$.

In the actual definition of the Fourier transform, factors of 2π invariably pop up here and there, and there is no universally accepted place to put these factors. This ambiguity is not unlike the distinction between radians and cycles in connection with frequencies in physics; again the distinction is a factor of 2π. The definition that we shall use occurs quite commonly these days, namely

$$\boxed{\widehat{f}(y) = \mathcal{F}f(y) = \int_{\mathbb{R}^N} f(x) e^{-2\pi i x \cdot y}\, dx,}$$

with $x \cdot y$ referring to the dot product and with the 2π in the exponent. The formula for $\mathcal{F}^{-1}$ will turn out to be similar looking, except that the minus sign is changed to plus in the exponent. Some authors drop the 2π from the exponent, and then a factor of $(2\pi)^{-N}$ is needed in the inversion formula. Other authors who drop the 2π from the exponent also include a factor of $(2\pi)^{-N}$ in front of the integral; then the inversion formula requires no such factor. Still other authors who drop the 2π from the exponent insert a factor of $(2\pi)^{-N/2}$ in the formula for both $\mathcal{F}$ and its inverse. In all cases, there is a certain utility in adjusting the definition of convolution by an appropriate power of 2π so that the Fourier transform of a convolution is indeed the pointwise product of the Fourier transforms. The relationships among these alternative formulas are examined in Problem 1 at the end of the chapter.

At any rate, in this book we take the boxed formula above as the definition of the Fourier transform of a function f in $L^1(\mathbb{R}^N, dx)$. Convolution was defined in Section VI.2. Although there are many elementary functions for which one can compute the Fourier transform explicitly, there are precious few for which one can make the pair of calculations that compute the Fourier transform and verify the inversion formula. One example is $e^{-\pi|x|^2}$, which will be examined in the next section.

Recall from Section VI.1 that the translate $\tau_{x_0} f$ is defined by $\tau_{x_0} f(x) = f(x - x_0)$.

Proposition 8.1. The Fourier transform on $L^1(\mathbb{R}^N)$ has these properties:

(a) f in L^1 implies that $\widehat{f}$ is bounded and uniformly continuous with $\|\widehat{f}\|_{\sup} \le \|f\|_1$,

(b) f in L^1 implies that the translate $\tau_{x_0} f$ and the product $f(x)e^{-2\pi i x\cdot y_0}$ have $(\tau_{x_0} f)\widehat{\ }(y) = e^{-2\pi i x_0\cdot y}\widehat{f}(y)$ and $\mathcal{F}(f(x)e^{2\pi i x\cdot y_0})(y) = (\tau_{y_0}\widehat{f})(y)$,

(c) f and g in L^1 implies $\widehat{f * g} = \widehat{f}\,\widehat{g}$,

(d) f in L^1 implies $\widehat{f^*} = \overline{\widehat{f}}$, where $f^*(x) = \overline{f(-x)}$,

(e) **(multiplication formula)** f and φ in L^1 implies $\int_{\mathbb{R}^N} f\widehat{\varphi}\,dx = \int_{\mathbb{R}^N} \widehat{f}\varphi\,dx$,

(f) f in L^1 and $2\pi i x_j f$ in L^1 implies that $\frac{\partial\widehat{f}}{\partial y_j}$ exists in the ordinary sense everywhere and satisfies $\frac{\partial\widehat{f}}{\partial y_j} = \mathcal{F}(-2\pi i x_j f)$,

(g) f in L^1 and $\frac{\partial f}{\partial x_j}$ existing in the L^1 sense, i.e., $\lim_{h\to 0} h^{-1}(\tau_{-he_j} f - f)$ existing in L^1, implies $\mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y) = 2\pi i y_j \widehat{f}(y)$. This formula holds also when f is in $L^1 \cap C^1$, the ordinary $\frac{\partial f}{\partial x_j}$ is in L^1, and f vanishes at infinity.

PROOF. All the integrals will be over $\mathbb{R}^N$, and we drop $\mathbb{R}^N$ from the notation. For (a), we have $|\widehat{f}(y)| \le \int |f(x)|\,dx = \|f\|_1$, and hence $\|\widehat{f}\|_{\sup} \le \|f\|_1$. Also,

$$\begin{aligned}|\widehat{f}(y_1) - \widehat{f}(y_2)| &\le \int |f(x)|\,|e^{-2\pi i x\cdot y_1} - e^{-2\pi i x\cdot y_2}|\,dx \\ &= \int |f(x)|\,|e^{-2\pi i x\cdot(y_1-y_2)} - 1|\,dx.\end{aligned}$$

On the right side the second factor of the integrand is bounded by 2 and tends to 0 for each x as $y_1 - y_2$ tends to 0. Thus the right side tends to 0 by dominated convergence at a rate depending only on $y_1 - y_2$.

For (b), $(\tau_{x_0} f)\widehat{\ }(y) = \int f(x - x_0)e^{-2\pi i x\cdot y}\,dx = \int f(x)e^{-2\pi i(x+x_0)\cdot y}\,dx = e^{-2\pi i x_0\cdot y}\widehat{f}(y)$ and

$$\begin{aligned}\mathcal{F}(f(x)e^{2\pi i x\cdot y_0})(y) &= \int f(x)e^{2\pi i x\cdot y_0}e^{-2\pi i x\cdot y}\,dx\\ &= \int f(x)e^{-2\pi i x\cdot(y-y_0)}\,dx = (\tau_{y_0}\widehat{f}\,)(y).\end{aligned}$$

For (c), we use Fubini's Theorem. The standard technique for verifying the theorem's applicability was mentioned near the end of Section V.7. Let us see the technique in context this once. The procedure is to write out the computation, blindly making the interchange, and then to check the validity of the interchange by imagining that absolute value signs have been put in place. What needs to be verified is that the double or iterated integrals with the absolute value signs in place are finite. The computation here is

$$\begin{aligned}\widehat{f * g}(y) &= \iint f(x-t)g(t)e^{-2\pi i x\cdot y}\,dt\,dx = \iint f(x-t)g(t)e^{-2\pi i x\cdot y}\,dx\,dt\\ &= \iint f(x)g(t)e^{-2\pi i(x+t)\cdot y}\,dx\,dt = \widehat{f}(y)\widehat{g}(y).\end{aligned}$$

The steps with absolute value signs in place around the integrands are

$$\begin{aligned}\iint |f(x-t)g(t)e^{-2\pi i x\cdot y}|\,dt\,dx &= \iint |f(x-t)g(t)e^{-2\pi i x\cdot y}|\,dx\,dt\\ &= \iint |f(x)g(t)e^{-2\pi i(x+t)\cdot y}|\,dx\,dt.\end{aligned}$$

The first interchange is valid, but the first and second integrals are not so clearly finite. What is clear, because f and g are integrable, is that we have finiteness for the third integral, and the second and third integrals are equal by a translation in the inner integration. Thus the computation of $\widehat{f * g}(y)$ is justified.

For (d), we have $\widehat{f^*}(y) = \int \overline{f(-x)}e^{-2\pi i x\cdot y}\,dx = \overline{\int f(x)e^{-2\pi i x\cdot y}\,dx} = \overline{\widehat{f}(y)}$.

For (e), we use Fubini's Theorem, justifying the details in the same way as in (c). We obtain

$$\begin{aligned}\int f\widehat{\varphi}\,dx &= \iint f(x)\varphi(y)e^{-2\pi i y\cdot x}\,dy\,dx\\ &= \iint f(x)\varphi(y)e^{-2\pi i y\cdot x}\,dx\,dy = \int \widehat{f}\,\varphi\,dy,\end{aligned}$$

and the interchange is valid because f and φ have been assumed integrable.

For (f), we apply (b) and obtain

$$h^{-1}\big(\widehat{f}(y + he_j) - \widehat{f}(y)\big) = \mathcal{F}\big(f(x)\,h^{-1}(e^{-2\pi i h e_j\cdot x} - 1)\big)(y).$$

Application of the Mean Value Theorem to the real and imaginary parts of $h^{-1}(e^{-2\pi i h e_j\cdot x}-1)$ shows for $|h|\le 1$ that

$$\big|\operatorname{Re}(h^{-1}(e^{-2\pi i h e_j\cdot x}-1))\big| = |h^{-1}(1-\cos 2\pi h x_j)| \le 2\pi|x_j|$$

and $$\big|\operatorname{Im}(h^{-1}(e^{-2\pi i h e_j\cdot x}-1))\big| = |h^{-1}\sin 2\pi h x_j| \le 2\pi|x_j|,$$

hence that $$\big|h^{-1}(e^{-2\pi i h e_j\cdot x}-1)| \le 4\pi|x_j|.$$

Since $x_j f(x)$ is assumed integrable, we have dominated convergence in the computation of the limit of $\mathcal{F}\big(f(x)\ h^{-1}(e^{-2\pi i h e_j\cdot x}-1)\big)(y)$ as h tends to 0, and we get $\mathcal{F}\big(-2\pi i x_j f\big)(y) = \dfrac{\partial \widehat{f}}{\partial y_j}(y)$.

For the first part of (g), the assumptions and (a) give

$$\big|\mathcal{F}(h^{-1}(\tau_{-he_j}f - f))(y) - \mathcal{F}\big(\tfrac{\partial f}{\partial x_j}\big)(y)\big| \le \big\|h^{-1}(\tau_{-he_j}f - f) - \tfrac{\partial f}{\partial x_j}\big\|_1 \to 0.$$

The left side equals $\big|\widehat{f}(y)(h^{-1}(e^{2\pi i h e_j\cdot y}-1)) - \mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y)\big|$ by (b), and this tends to $\big|\widehat{f}(y)2\pi i y_j - \mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y)\big|$. Hence $\mathcal{F}\big(\frac{\partial f}{\partial x_j}\big)(y) = 2\pi i y_j\widehat{f}(y)$.

For the second part of (g), let x_j' denote the tuple of the $N-1$ variables other than x_j. Then integration by parts in the variable x_j gives

$$\begin{aligned}
\mathcal{F}\big(\tfrac{\partial f}{\partial x_j}\big)(y) &= \textstyle\int_{\mathbb{R}^{N-1}}\int_{-\infty}^{\infty}\frac{\partial f}{\partial x_j}(x)e^{-2\pi i x\cdot y}\,dx_j\,dx_j' \\
&= \textstyle\int_{\mathbb{R}^{N-1}}\lim_n\int_{-n}^{n}\frac{\partial f}{\partial x_j}(x)e^{-2\pi i x\cdot y}\,dx_j\,dx_j' \\
&= \textstyle\int_{\mathbb{R}^{N-1}}\lim_n\big[f(x)e^{-2\pi i x\cdot y}\big]_{x_j=-n}^{n}\,dx_j' \\
&\qquad - \textstyle\int_{\mathbb{R}^{N-1}}\lim_n\int_{-n}^{n}f(x)(-2\pi i y_j)e^{-2\pi i x\cdot y}\,dx_j\,dx_j' \\
&= 0 + 2\pi i y_j\widehat{f}(y),
\end{aligned}$$

as asserted. □

2. Fourier Transform on L^1, Inversion Formula

The main theorem of this section is the Fourier inversion formula for $L^1(\mathbb{R}^N)$. The Fourier transform for $\mathbb{R}^1$ is the analog for the line of the mapping that carries a function f on the circle to its doubly infinite sequence $\{c_k\}$ of Fourier coefficients. The inversion problem for the circle amounts to recovering f from the c_k's. We know that the procedure is to form the partial sums $s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}$ and to look for a sense in which $\{s_n\}$ converges to f. There is no problem for the case that f is itself a trigonometric polynomial; then s_n will be equal to f for large enough n, and no passage to the limit is necessary.

The situation with the Fourier transform is different. There is no readily available nonzero integrable function on the line analogous to an exponential on the circle for which we know an inversion formula with all constants in place. In order to obtain such an inversion formula for the Fourier transform on L^1, it is necessary to be able to invert the Fourier transform of some particular nonzero function explicitly. This step is carried out in Proposition 8.2 below, and then we can address the inversion problem of $L^1(\mathbb{R}^N)$ in general. The analog for the circle of what we shall prove for the line is a rather modest result: It would say that if $\sum |c_k|$ is finite, then the sequence of partial sums converges uniformly to a function that equals f almost everywhere. The uniform convergence is a relatively trivial conclusion, being an immediate consequence of the Weierstrass M test; but the conclusion that we recover f lies deeper and incorporates a version of the uniqueness theorem.

Proposition 8.2. $\mathcal{F}\big(e^{-\pi|x|^2}\big) = e^{-\pi|y|^2}$.

REMARKS. Readers who know about the Cauchy Integral Theorem from elementary complex analysis or else Green's Theorem in the theory of line integrals will recognize that the calculation below amounts to an application of one or the other of these theorems to the function $e^{-\pi z^2}$ over a long thin geometric rectangle next to the x axis in $\mathbb{C}$. However, the present application of either of these theorems is so simple that we can without difficulty substitute a proof of one of these theorems in the special case of interest, and hence neither of these other theorems needs to be assumed. As the proof below will show, matters come down to the Fundamental Theorem of Calculus in its traditional form (Theorem 1.32).

PROOF. The question is whether

$$\int_{\mathbb{R}^N} e^{-\pi(x_1^2+\cdots+x_N^2)} e^{-2\pi i(x_1y_1+\cdots+x_Ny_N)}\,dx_1\cdots dx_N \stackrel{?}{=} e^{-\pi(y_1^2+\cdots+y_N^2)},$$

and the integral on the left is the product of N integrals in one variable. Thus the question is whether

$$\int_{-\infty}^{\infty} e^{-\pi(x^2+2ixy)}\,dx \stackrel{?}{=} e^{-\pi y^2}.$$

We start by observing that

$$\int_{-\infty}^{\infty} e^{-\pi(x^2+2ixy)}\,dx = e^{-\pi y^2}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx. \qquad (*)$$

Write

$$e^{-\pi(x+iy)^2} = u(x,y) + iv(x,y) = e^{-\pi(x^2-y^2)}\cos 2\pi xy - ie^{-\pi(x^2-y^2)}\sin 2\pi xy.$$

Direct calculation gives[1]

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{$**$}$$

Regard n as positive and large. Then

$$\begin{aligned}
&\textstyle\int_{-n}^{n} u(s,0)\,ds - \int_{-n}^{n} u(s,y)\,ds \\
&\qquad = -\textstyle\int_{-n}^{n}\int_0^y \frac{\partial u}{\partial y}(s,t)\,dt\,ds && \text{by Theorem 1.32} \\
&\qquad = +\textstyle\int_{-n}^{n}\int_0^y \frac{\partial v}{\partial x}(s,t)\,dt\,ds && \text{by } (**) \\
&\qquad = \textstyle\int_0^y\int_{-n}^{n} \frac{\partial v}{\partial x}(s,t)\,ds\,dt && \text{by Fubini's Theorem} \\
&\qquad = \textstyle\int_0^y v(n,t)\,dt - \int_0^y v(-n,t)\,dt && \text{by Theorem 1.32.}
\end{aligned}$$

With y fixed we let n tend to infinity. Then $v(n,t)$ and $v(-n,t)$ tend to 0 uniformly for t between 0 and y by inspection of v, and hence the right side of our expression tends to 0. Thus

$$\textstyle\int_{-\infty}^{\infty} u(s,0)\,ds = \int_{-\infty}^{\infty} u(s,y)\,ds,$$

which says that

$$\operatorname{Re}\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = \operatorname{Re}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx. \tag{$\dagger$}$$

Similarly we calculate

$$\begin{aligned}
\textstyle\int_{-n}^{n} v(s,0)\,ds - \int_{-n}^{n} v(s,y)\,ds &= -\textstyle\int_{-n}^{n}\int_0^y \frac{\partial v}{\partial y}(s,t)\,dt\,ds \\
&= -\textstyle\int_{-n}^{n}\int_0^y \frac{\partial u}{\partial x}(s,t)\,dt\,ds && \text{by } (**) \\
&= -\textstyle\int_0^y\int_{-n}^{n} \frac{\partial u}{\partial x}(s,t)\,ds\,dt \\
&= -\textstyle\int_0^y u(n,t)\,dt + \int_0^y u(-n,t)\,dt.
\end{aligned}$$

Again we can see that the right side tends to 0, and thus

$$\textstyle\int_{-\infty}^{\infty} v(s,0)\,ds = \int_{-\infty}^{\infty} v(s,y)\,ds,$$

which says that

$$\operatorname{Im}\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = \operatorname{Im}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx. \tag{$\dagger\dagger$}$$

Taking $(*)$ into account and combining $(\dagger)$ and $(\dagger\dagger)$, we obtain

$$\int_{-\infty}^{\infty} e^{-\pi(x^2+2ixy)}\,dx = e^{-\pi y^2}\int_{-\infty}^{\infty} e^{-\pi(x+iy)^2}\,dx = e^{-\pi y^2}\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx,$$

and the proposition follows from the formula $\int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = 1$ given in Proposition 6.33. □

[1] The equations $(**)$ are called the **Cauchy–Riemann equations**. They occur again in Section 7.

We shall use dilations to create an approximate identity out of $e^{-\pi|x|^2}$ in the style of Section VI.2. Put $\varphi(x) = e^{-\pi|x|^2}$ and define $\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x)$ for $\varepsilon > 0$. Whenever φ in an integrable function and φ_ε is formed in this way, we have

$$\begin{aligned}\widehat{\varphi_\varepsilon}(y) &= \textstyle\int_{\mathbb{R}^N} \varphi_\varepsilon(x)e^{-2\pi i x\cdot y}\,dx = \varepsilon^{-N}\int_{\mathbb{R}^N} \varphi(\varepsilon^{-1}x)e^{-2\pi i x\cdot y}\,dx \\ &= \textstyle\int_{\mathbb{R}^N} \varphi(x)e^{-2\pi i x\cdot \varepsilon y}\,dx = \widehat{\varphi}(\varepsilon y),\end{aligned}$$

the next-to-last equality following from the change of variables $\varepsilon^{-1}x \mapsto x$.

For the particular function $\varphi(x) = e^{-\pi|x|^2}$, this calculation shows that $\widehat{\varphi_\varepsilon}(y) = e^{-\pi\varepsilon^2|y|^2}$. In particular, $\widehat{\varphi_\varepsilon}$ is ≥ 0 and vanishes at ∞ for each fixed $\varepsilon > 0$. As ε decreases to 0, $\widehat{\varphi_\varepsilon}$ increases pointwise to the constant function 1. The constant c in Theorem 6.20 for this φ is $c = \int_{\mathbb{R}^N}\varphi(x)\,dx = 1$ by Proposition 6.33. That theorem gives various convergence results for $\varphi_\varepsilon * f$, one of which is that $\varphi_\varepsilon * f$ converges to f in L^1 if f is in L^1.

Theorem 8.3 (Riemann–Lebesgue Lemma). If f is in $L^1(\mathbb{R}^N)$, then the continuous function $\widehat{f}$ vanishes at infinity.

PROOF. The continuity of $\widehat{f}$ is by Proposition 8.1a. Put $\varphi(x) = e^{-\pi|x|^2}$ and form φ_ε. Then parts (c) and (a) of Proposition 8.1 give

$$\|\widehat{\varphi_\varepsilon}\widehat{f} - \widehat{f}\|_{\sup} = \|\widehat{\varphi_\varepsilon * f} - \widehat{f}\|_{\sup} \leq \|\varphi_\varepsilon * f - f\|_1,$$

and Theorem 6.20 shows that the right side tends to 0 as ε decreases to 0. Hence $e^{-\pi\varepsilon^2|y|^2}\widehat{f}(y)$ tends to $\widehat{f}(y)$ uniformly in y. Since $\widehat{f}$ is bounded (Proposition 8.1a), $e^{-\pi\varepsilon^2|y|^2}\widehat{f}(y)$ vanishes at infinity. The uniform limit of functions vanishing at infinity vanishes at infinity, and the theorem follows. □

Theorem 8.4 (Fourier inversion formula). If f is in $L^1(\mathbb{R}^N)$ and $\widehat{f}$ is in $L^1(\mathbb{R}^N)$, then f can be redefined on a set of measure 0 so as to be continuous. After this adjustment,

$$f(x) = \int_{\mathbb{R}^N} \widehat{f}(y)e^{2\pi i x\cdot y}\,dy.$$

PROOF. By way of preliminaries, recall from Proposition 8.1e that the multiplication formula gives $\int f\,\widehat{g}\,dx = \int \widehat{f}\,g\,dx$ whenever f and g are both integrable. With ε fixed for the moment, let us apply this formula with $g(x) = e^{-\pi\varepsilon^2|x|^2}$. The remarks before Theorem 8.3 about how the Fourier transform interacts with dilations show that $\widehat{g}(y) = \varepsilon^{-N}e^{-\pi\varepsilon^{-2}|y|^2}$. In other words, if we take $\varphi(x) = e^{-\pi|x|^2}$, then

$$\int_{\mathbb{R}^N} f(x)\varphi_\varepsilon(x)\,dx = \int_{\mathbb{R}^N} \widehat{f}(y)e^{-\pi\varepsilon^2|y|^2}\,dy. \tag{$*$}$$

To prove the theorem, consider first the special case that f is bounded and continuous. If we let ε decrease to 0 in $(*)$, the left side tends to $f(0)$ by Theorem 6.20c, and the right side tends to $\int_{\mathbb{R}^N} \widehat{f}(y)\,dy$ by dominated convergence since $\widehat{f}$ is assumed integrable. Thus $f(0) = \int_{\mathbb{R}^N} \widehat{f}(y)\,dy$. Applying this conclusion to the translate $\tau_{-x} f$ and using Proposition 8.1b, we obtain

$$f(x) = (\tau_{-x} f)(0) = \int (\tau_{-x} f)\widehat{\ }(y)\,dy = \int \widehat{f}(y) e^{2\pi i x\cdot y}\,dy,$$

as required.

Without the special assumption on f, we adjust the above argument a little. Using the equality $\varphi_\varepsilon(-y) = \varphi_\varepsilon(y)$, we apply $(*)$ to the translate $\tau_{-x} f$ of f to get

$$\begin{aligned}\int \widehat{f}(y) e^{2\pi i x\cdot y} e^{-\pi\varepsilon^2|y|^2}\,dy &= \int f(x+y)\varphi_\varepsilon(y)\,dy \\ &= \int f(x-y)\varphi_\varepsilon(y)\,dy = (\varphi_\varepsilon * f)(x).\end{aligned}$$

As ε decreases to 0, the left side tends pointwise to $\int \widehat{f}(y) e^{2\pi i x\cdot y}\,dy$ by dominated convergence, and the result is a continuous function of x, by a version of Proposition 8.1a. The right side tends to f in L^1 by Theorem 6.20, and hence Theorem 5.59 shows that a subsequence of $\varphi_\varepsilon * f$ tends to f almost everywhere. Thus $f(x) = \int_{\mathbb{R}^N} \widehat{f}(y) e^{2\pi i x\cdot y}\,dy$ almost everywhere, with the right side continuous. □

Corollary 8.5. The Fourier transform is one-one on $L^1(\mathbb{R}^N)$.

PROOF. If f is in L^1 and $\widehat{f}$ is identically 0, then $\widehat{f}$ is in L^1, and the inversion formula (Theorem 8.4) applies. Hence f is 0 almost everywhere. □

3. Fourier Transform on L^2, Plancherel Formula

We mentioned in Section 1 that the Fourier transform is of great importance in analyzing operators that commute with translations. The initial analysis of such operators is done on $L^2(\mathbb{R}^N)$, and this section describes some of how that analysis comes about. The first result is the theorem for $\mathbb{R}^N$ that is the analog of Parseval's Theorem for the circle.

Theorem 8.6 (Plancherel formula). If f is in $L^1(\mathbb{R}^N)\cap L^2(\mathbb{R}^N)$, then $\|\widehat{f}\|_2 = \|f\|_2$.

REMARKS. There is a formal computation that is almost a proof, namely

$$\begin{aligned}\int |f(x)|^2\,dx &= \int f^*(-x) f(x)\,dx = (f^* * f)(0) \\ &= \int \widehat{f^* * f}(y)\,dy = \int \widehat{f^*}(y)\widehat{f}(y)\,dy = \int |\widehat{f}(y)|^2\,dy,\end{aligned}$$

the middle equality using the Fourier inversion formula (Theorem 8.4). What is needed in order to make this computation into a proof is a verification that the Fourier inversion formula actually applies. We know that $f^* * f$ is in L^1 since f^* and f are in L^1, and we know from Proposition 6.18 that $f^* * f$ is continuous, being in $L^2 * L^2$. But it is not immediately obvious that the Fourier transform to which the inversion formula is to be applied, namely $\widehat{f^* * f} = |\widehat{f}|^2$, is in L^1. We handle this question by proving a lemma that is a little more general than necessary.

Lemma 8.7. Suppose f is in $L^1(\mathbb{R}^N)$, is bounded on $\mathbb{R}^N$, and is continuous at 0. If $\widehat{f}(y) \geq 0$ for all y, then $\widehat{f}$ is in $L^1(\mathbb{R}^N)$.

PROOF. Put $\varphi(x) = e^{-\pi|x|^2}$ and $\varphi_\varepsilon(x) = \varepsilon^{-N}\varphi(\varepsilon^{-1}x)$. Then the function $\varphi_\varepsilon * f$ is continuous by Proposition 6.18 since φ_ε is in L^∞ and f is in L^1, and

$$\lim_{\varepsilon \downarrow 0}(\varphi_\varepsilon * f)(0) = f(0)$$

by Theorem 6.20c. The function $\widehat{\varphi_\varepsilon}$ is in L^1, and $\widehat{f}$ is bounded. Hence $\widehat{\varphi_\varepsilon * f} = \widehat{\varphi_\varepsilon}\widehat{f}$ is in L^1. By the Fourier inversion formula (Theorem 8.4),

$$(\varphi_\varepsilon * f)(0) = \int_{\mathbb{R}^N} \widehat{f}(y)e^{-\pi\varepsilon^2|y|^2}\,dy.$$

Letting ε decrease to 0 and taking into account the monotone convergence, we obtain $f(0) = \int_{\mathbb{R}^N} \widehat{f}(y)\,dy$. Therefore $\widehat{f}$ is integrable. □

PROOF OF THEOREM 8.6. The remarks after the statement of the theorem prove everything except that the Fourier transform $\widehat{f^* * f} = |\widehat{f}|^2$ is in L^1, and this step is carried out by Lemma 8.7. □

Abstract linear operators between normed linear spaces were introduced in Section V.9, and Proposition 5.57 showed that boundedness is equivalent to uniform continuity. Let us make use of such operators now.

Theorem 8.6 allows us to extend the Fourier transform for $\mathbb{R}^N$ from $L^1 \cap L^2$ to L^2. In fact, Proposition 5.56 shows that $L^1 \cap L^2$ is dense in L^2. The conclusion of Theorem 8.6 implies that the linear operator $\mathcal{F}$ is bounded relative to the L^2 norms on domain and range, and hence it is uniformly continuous. Since the range space L^2 is complete (Theorem 5.59), Proposition 2.47 shows that $\mathcal{F}$ extends to a continuous map $\mathcal{F} : L^2 \to L^2$. This extended map, also called $\mathcal{F}$, is readily checked to be linear and then is a bounded linear operator satisfying $\|\mathcal{F}f\|_2 = \|f\|_2$ for all f in L^2.

If f is in $L^2(\mathbb{R}^N)$, we can use any approximating sequence from $L^1 \cap L^2$ to obtain a formula for $\mathcal{F}f$. One such is $f I_{B(R;0)}$, as R increases to infinity through some sequence. Thus

$$\mathcal{F}f(y) = \lim_{\substack{(\text{in } L^2 \text{ sense}) \\ R\to\infty}} \int_{|x|<R} f(x) e^{-2\pi i x\cdot y}\, dx.$$

Corollary 8.8. If f is in $L^1(\mathbb{R}^N)$ and g is in $L^2(\mathbb{R}^N)$, then $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$ and $\mathcal{F}(g^*) = \overline{\mathcal{F}(g)}$.

PROOF. Set $g_n = g I_{B(n;0)}$, so that g_n is in $L^1 \cap L^2$ for all n and $g_n \to g$ in L^2. Then $f * g_n \to f * g$ in L^2 since $\|f * g_n - f * g\|_2 = \|f * (g_n - g)\|_2 \le \|f\|_1 \|g_n - g\|_2$. Therefore $\mathcal{F}(f)\mathcal{F}(g_n) = \mathcal{F}(f * g_n) \to \mathcal{F}(f * g)$ in L^2. Since $\mathcal{F}(f)$ is a bounded function and $\mathcal{F}(g_n) \to \mathcal{F}(g)$ in L^2, we see that $\mathcal{F}(f)\mathcal{F}(g_n) \to \mathcal{F}(f)\mathcal{F}(g)$ in L^2. Hence $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$. The identity $\mathcal{F}(g^*) = \overline{\mathcal{F}(g)}$ is proved similarly. $\square$

Theorem 8.9 (Riesz–Fischer Theorem). The Fourier transform operator $\mathcal{F}$ carries $L^2(\mathbb{R}^N)$ onto $L^2(\mathbb{R}^N)$.

PROOF. The operator $\mathcal{F}$ is built from the integral $\int_{\mathbb{R}^N} f(x)e^{-2\pi i x\cdot y}\, dx$. In a similar fashion, build an operator $\mathcal{I}$ from $\int_{\mathbb{R}^N} f(x)e^{2\pi i x\cdot y}\, dx$, or equivalently define $\mathcal{I}f(y) = \mathcal{F}f(-y)$. Then $\|\mathcal{I}f\|_2 = \|f\|_2$ for f in L^2. It is sufficient to prove that $\mathcal{F}\mathcal{I} = 1$ on L^2, since for any f in L^2, the equation $\mathcal{F}\mathcal{I} = 1$ implies that $\mathcal{I}f$ is a member of L^2 carried to f by $\mathcal{F}$. Moreover, $\mathcal{F}\mathcal{I}$ is continuous, being bounded. It is therefore enough to prove that $\mathcal{F}\mathcal{I}f = f$ for f in a dense subspace of L^2. We shall do so for the dense subspace $L^1 \cap L^2$.

For a function f in $L^1 \cap L^2$ with the additional property that $\widehat{f}$ is in L^1 (and also L^2 by Theorem 8.6), Theorem 8.4 for $\mathcal{I}$ (or Theorem 8.4 applied to the function $f(-x)$) shows that $\mathcal{F}\mathcal{I}f = f$.

For a general f in $L^1 \cap L^2$, form $\varphi_\varepsilon * f$, where $\varphi(x) = e^{-\pi|x|^2}$. Then $\widehat{\varphi_\varepsilon * f} = \widehat{\varphi_\varepsilon}\widehat{f}$ is in $L^1 \cap L^2$; in fact, it is in L^2 by Proposition 6.14 and Theorem 8.6, and it is in L^1 because $\widehat{f}$ is bounded and $\widehat{\varphi_\varepsilon}$ is in L^1. By the special case just proved, $\mathcal{F}\mathcal{I}(\varphi_\varepsilon * f) = \varphi_\varepsilon * f$. Since $\mathcal{F}\mathcal{I}$ is continuous and $\varphi_\varepsilon * f \to f$ in L^2 by Theorem 6.20a, $\mathcal{F}\mathcal{I}f = f$. Thus $\mathcal{F}\mathcal{I}f = f$ on the dense subspace $L^1 \cap L^2$, and the proof is complete. $\square$

We shall be interested especially in bounded linear operators A on $L^2(\mathbb{R}^N)$ that commute with translations, i.e., that satisfy $A(\tau_x f) = \tau_x(Af)$ for all x in $\mathbb{R}^N$ and all f in L^2. Recall that the operator norm $\|A\|$ of a bounded linear operator on L^2 is the least C such that $\|Af\|_2 \le C\|f\|_2$ for all f in L^2.

EXAMPLES.

(1) The translation τ_{x_0} is an example of a bounded linear operator on L^2 that commutes with translations; the commutativity in question follows from the commutativity of $\mathbb{R}^N$ as an additive group, and the equality $\|\tau_{x_0} f\|_2 = \|f\|_2$ shows that τ_{x_0} is bounded with $\|\tau_{x_0}\| = 1$. In terms of Fourier transforms, Proposition 8.1 shows that $(\tau_{x_0} f)\widehat{\ }(y) = e^{-2\pi i x_0 \cdot y}\widehat{f}(y)$.

(2) Another example of a bounded linear operator on L^2 that commutes with translations is the operator $Ag = f * g$ for fixed f in L^1. This commutes with translations by Proposition 6.15, and it is bounded with $\|A\| \leq \|f\|_1$ by Proposition 6.14. Proposition 8.1 shows that $\widehat{Ag} = \widehat{f}\,\widehat{g}$.

(3) Let $M(y)$ be any L^∞ function on $\mathbb{R}^N$, and for f in L^2, define Af by $\widehat{Af} = M\widehat{f}$. The function $\widehat{f}$ is in L^2 by the Plancherel Theorem, $M\widehat{f}$ is in L^2 since M is essentially bounded, and $M\widehat{f}$ is the Fourier transform of some L^2 function by the Riesz–Fischer Theorem. We take this L^2 function to be Af. The brief formula is $Af = \mathcal{F}^{-1}(M\mathcal{F}f)$. From the inequalities $\|Af\|_2 = \|M\mathcal{F}f\|_2 \leq \|M\|_\infty \|\mathcal{F}f\|_2 = \|M\|_\infty \|f\|_2$, we see that A is bounded with $\|A\| \leq \|M\|_\infty$. The bounded linear operator A commutes with translations, since

$$\mathcal{F}(A(\tau_x f))(y) = \mathcal{F}(\mathcal{F}^{-1}M\mathcal{F}\tau_x f)(y) = M\mathcal{F}\tau_x f(y) = M(y)e^{-2\pi i x\cdot y}\mathcal{F}f(y)$$

and $$\mathcal{F}(\tau_x(Af))(y) = e^{-2\pi i x\cdot y}\mathcal{F}(Af)(y) = e^{-2\pi i x\cdot y}M(y)\mathcal{F}f(y).$$

One speaks of the function M as a **multiplier** on L^2. The previous two examples are instances of this construction. Example 1 has $M(y) = e^{-2\pi i x_0\cdot y}$, and Example 2 has $M(y) = \widehat{f}(y)$. We shall see in Theorem 8.14 in the next section that *every* bounded linear operator A on L^2 commuting with translations arises from some such essentially bounded M and that $\|A\| = \|M\|_\infty$; for this reason a bounded linear operator on L^2 that commutes with translations is often called a "multiplier operator" or a "bounded multiplier operator" on L^2.

4. Schwartz Space

This section introduces the space $\mathcal{S}(\mathbb{R}^N)$ of Schwartz functions on $\mathbb{R}^N$. This space is a vector subspace of $L^1(\mathbb{R}^N)$, so that the Fourier transform is given on it by the usual concrete formula; $\mathcal{S}(\mathbb{R}^N)$ will turn out to be another space besides L^2 that is carried onto itself by the Fourier transform. Working with $\mathcal{S}(\mathbb{R}^N)$ provides a convenient way for using the Fourier transform and derivatives together, as becomes clearer when one studies partial differential equations.

If Q is a complex-valued polynomial on $\mathbb{R}^N$, define $Q(D)$ to be the partial differential operator with constant coefficients obtained by substituting, for each

j with $1 \le j \le N$, the operator $D_j = \frac{\partial}{\partial x_j}$ for x_j. A **Schwartz function** on $\mathbb{R}^N$ is a smooth function such that $P(x)Q(D)f$ is bounded for each pair of polynomials P and Q. An example is the function $e^{-\pi|x|^2}$, since its iterated partial derivatives are all of the form $R(x)e^{-\pi|x|^2}$ for some polynomial R. The **Schwartz space** $\mathcal{S} = \mathcal{S}(\mathbb{R}^N)$ is the set of all Schwartz functions.

The Schwartz space $\mathcal{S}$ is evidently a vector space, and it is closed under partial differentiation and under multiplication by polynomials. Closure under partial differentiation is in effect built into the definition. To see closure under multiplication by polynomials, it is enough to check closure under multiplication by each monomial x_j. This closure follows readily from the formula $Q(D)(x_j f) = Q^\#(D)f + x_j Q(D)f$, where $Q^\#$ is 0 or is a polynomial having degree strictly lower than Q has.

If f is a Schwartz function, then $P(x)Q(D)f$ is actually integrable, as well as bounded, for each pair of polynomials P and Q. In fact, $(1+|x|^2)^N P(x)Q(D)f$ is bounded, and therefore $P(x)Q(D)f$ is $\le$ a multiple of the integrable function $(1+|x|^2)^{-N}$. In particular, $\mathcal{S}$ is contained in L^1, L^2, and L^∞.

Finally the Fourier transform $\mathcal{F}$ carries $\mathcal{S}$ into itself. In fact, parts (f) and (g) of Proposition 8.1 give

$$P(x)Q(D)\widehat{f} = P(x)\mathcal{F}(Q(-2\pi i x)f) = \mathcal{F}(P((2\pi i)^{-1}D)Q(-2\pi i x)f),$$

and the right side is the Fourier transform of an L^1 function and therefore is bounded.

Proposition 8.10. The Fourier transform $\mathcal{F}$ is one-one from $\mathcal{S}(\mathbb{R}^N)$ onto $\mathcal{S}(\mathbb{R}^N)$, and the Fourier inversion formula holds on $\mathcal{S}(\mathbb{R}^N)$.

PROOF. Since $\mathcal{S} \subseteq L^1$, $\mathcal{F}$ is one-one on $\mathcal{S}$ as a consequence of Corollary 8.5. Since $\mathcal{F}(\mathcal{S}) \subseteq \mathcal{S} \subseteq L^1$, Theorem 8.4 shows that the Fourier inversion formula holds on $\mathcal{S}$. Let $(\mathcal{I}f)(x) = (\mathcal{F}f)(-x)$ for f in L^1. Then $\mathcal{I}(\mathcal{S}) \subseteq \mathcal{S}$. The Riesz–Fischer Theorem (Theorem 8.9) shows that $\mathcal{F}\mathcal{I} = 1$ on $L^1 \cap L^2$, and hence $\mathcal{F}\mathcal{I} = 1$ on $\mathcal{S}$ as well. Therefore if f is in $\mathcal{S}$, then $g = \mathcal{I}f$ is a member of $\mathcal{S}$ such that $\mathcal{F}g = f$, and we conclude that $\mathcal{F}$ carries $\mathcal{S}$ onto $\mathcal{S}$. $\square$

To make effective use of Proposition 8.10, we need to know that $\mathcal{S}(\mathbb{R}^N)$ is quite large, large enough so that we can shape functions suitably when we need them. For U open in $\mathbb{R}^N$, let $C^\infty_{\text{com}}(U)$ denote the vector space of smooth complex-valued functions on U whose support is a compact subset of U. It is apparent that $C^\infty_{\text{com}}(U)$ is closed under pointwise multiplication and that every member of $C^\infty_{\text{com}}(U)$ extends to a member of $C^\infty_{\text{com}}(\mathbb{R}^N)$ when set equal to 0 off U. But it is not apparent that $C^\infty_{\text{com}}(U)$ contains nonzero functions. We shall construct some.

Lemma 8.11. If δ_1 and δ_2 are given positive numbers with $\delta_1 < \delta_2$, then there exists ψ in $C^\infty_{\text{com}}(\mathbb{R}^N)$ such that $\psi(x)$ depends only on $|x|$, ψ is nonincreasing in $|x|$, ψ takes values in $[0, 1]$, $\psi(x) = 1$ for $|x| \le \delta_1$, and $\psi(x) = 0$ for $|x| \ge \delta_2$.

PROOF. We begin from the statement in Section I.7 that the function $f : \mathbb{R} \to \mathbb{R}$ with $f(t)$ equal to e^{-1/t^2} for $t > 0$ and equal to 0 for $t \le 0$ is smooth everywhere, including at $t = 0$. (The verification that f is smooth occurs in Problems 20–22 at the end of Chapter I.) If $\delta > 0$, then it follows that the function $g_\delta(t) = f(\delta + t) f(\delta - t)$ is smooth. Consequently the function $h_\delta(t) = \int_{-\infty}^t g_\delta(s)\, ds$ is smooth, is nondecreasing, is 0 for $t \le -\delta$, is some positive constant for $t \ge \delta$, and takes only values between 0 and that positive constant. Forming the function $h_{\delta,r}(t) = h_\delta(r + t) h_\delta(r - t)$ with r at least δ and dilating it suitably, we obtain a smooth even function $\psi_0(t)$ with values in $[0, c]$, the function being identically 0 for $|t| \ge \delta_2$ and being identically c for $|t| \le \delta_1$. Putting $\psi(x) = c^{-1}\psi_0(|x|)$, we obtain the desired function. □

Proposition 8.12. If K and U are subsets of $\mathbb{R}^N$ with K compact, U open, and $K \subseteq U$, then there exists $\varphi \in C^\infty_{\text{com}}(U)$ with values in $[0, 1]$ such that φ is identically 1 on K.

PROOF. There is no loss of generality in assuming that K is nonempty and U is bounded. The continuous distance function $D(x, U^c)$ is everywhere positive on the compact set K and hence assumes a positive minimum c. Define K' to be the set $\{x \in \mathbb{R}^N \mid D(x, K) \le \frac{1}{4}c\}$. This set is compact, contains K, and has nonempty interior. Since the interior is nonempty, K' has positive Lebesgue measure $|K'|$. Applying Lemma 8.11, let h be a nonnegative smooth function that vanishes identically for $|x| \ge \frac{1}{4}c$ and has total integral 1.

Define $\varphi = h * I_{K'}$, where $I_{K'}$ is the indicator function of K'. Corollary 6.19 shows that φ is smooth. The function φ is ≥ 0 and has $\sup|\varphi| \le \|h\|_1 \|I_{K'}\|_\infty = 1$.

We have $\varphi(x) = \int_{\mathbb{R}^N} h(x - y) I_{K'}(y)\, dy$. If x is in K and $h(x - y)$ is nonzero, then $|x - y| \le \frac{1}{4}c$. Then $D(y, K) \le |x - y| \le \frac{1}{4}c$, and y is in K'. Hence $I_{K'}(y) = 1$, and $\varphi(x) = \int_{\mathbb{R}^N} h(x - y)\, dy = 1$.

Next, suppose $D(x, U^c) \le \frac{1}{4}c$ and $h(x - y)$ is nonzero, so that again $|x - y| \le \frac{1}{4}c$. The claim is that y is not in K', i.e., that $D(y, K) > \frac{1}{4}c$. Assuming the contrary, we can find, because of the compactness of K, some $k \in K$ with $|y - k| \le \frac{1}{4}c$. Then every $u^c \in U^c$ satisfies $c \le |u^c - k| \le |u^c - x| + |x - y| + |y - k| \le |u^c - x| + \frac{1}{4}c + \frac{1}{4}c$, and we obtain $|u^c - x| \ge \frac{1}{2}c$. Taking the infimum over u^c, we obtain $D(x, U^c) \ge \frac{1}{2}c$, and this is a contradiction. Thus y is not in K', and the integrand is identically 0 whenever $D(x, U^c) \le \frac{1}{4}c$. Hence $\varphi(x) = 0$ if $D(x, U^c) \le \frac{1}{4}c$, and the support of φ is a compact subset of U. This completes the proof. □

Every function in $C^\infty_{\text{com}}(\mathbb{R}^N)$ is the Fourier transform of some Schwartz function by Proposition 8.10, and there are many such functions by Proposition 8.12. With this fact in hand, we can prove the theorem about operators commuting with translations that was promised in the previous section. We begin with a lemma.

Lemma 8.13. If A is a bounded linear operator on $L^2(\mathbb{R}^N)$ commuting with translations, then A commutes with convolution by any L^1 function.

PROOF. We are to show that $A(f*g) = f*(Ag)$ if f is in L^1 and g is in L^2. Let $\epsilon > 0$ be given. Corollary 6.17, with $g_1 = g$ and $g_2 = Ag$, shows that there exist $y_1, \dots, y_n$ in $\mathbb{R}^N$ and constants $c_1, \dots, c_n$ such that $\big\| f * g - \sum_{j=1}^n c_j \tau_{y_j} g \big\|_2 < \epsilon$ and $\big\| f * Ag - \sum_{j=1}^n c_j \tau_{y_j} Ag \big\|_2 < \epsilon$. Then we have

$$\begin{aligned}\|A(f*g) - f*Ag\|_2 \le{}& \big\| A\big(f*g - \textstyle\sum_{j=1}^n c_j \tau_{y_j} g\big)\big\|_2 \\ &+ \big\| A\big(\textstyle\sum_{j=1}^n c_j \tau_{y_j} g\big) - \textstyle\sum_{j=1}^n c_j \tau_{y_j} Ag \big\|_2 \\ &+ \big\| \textstyle\sum_{j=1}^n c_j \tau_{y_j} Ag - f*Ag \big\|_2.\end{aligned}$$

The first term on the right side is $\le \|A\| \, \big\| f*g - \sum_{j=1}^n c_j \tau_{y_j} g \big\| \le \epsilon \|A\|$, the second term is 0 since A commutes with translations, and the third term is $< \epsilon$ by construction. $\square$

Theorem 8.14. If A is a bounded linear operator on $L^2(\mathbb{R}^N)$ commuting with translations, then there exists an L^∞ function M such that $Af = \mathcal{F}^{-1}(M\mathcal{F}f)$ for all f in L^2. As a member of L^∞, M is unique and satisfies $\|M\|_\infty = \|A\|$.

REMARKS. The idea of the proof comes from the corresponding result for L^2 of the circle, where it is easy to define M. Call the operator T in the case of the circle. Each function e^{ikx} is in L^2, and the given operator T satisfies $\tau_{x_0}(T(e^{ikx})) = T(\tau_{x_0}(e^{ikx})) = T(e^{ik(x-x_0)}) = e^{-ikx_0}T(e^{ikx})$. If we write f for the L^2 function $T(e^{ikx})$ and form the Fourier series expansion $f(x) \sim \sum c_n e^{inx}$, then $\tau_{x_0} f$ has Fourier series $\tau_{x_0} f(x) \sim \sum c_n e^{-inx_0} e^{inx}$ by linearity and boundedness of τ_{x_0}. Since we have just seen that $\tau_{x_0} f = e^{-ikx_0} f$, we conclude that $\sum c_n e^{-inx_0} e^{inx} = \sum c_n e^{-ikx_0} e^{inx}$. If $c_n \ne 0$ for some n unequal to k, then we do not have the term-by-term match required by the uniqueness theorem. Hence only c_k can be nonzero, and we have $T(e^{ikx}) = c_k e^{ikx}$. The number c_k is the value of the multiplier M at the integer k. In the actual setting of the theorem, the circle is replaced by $\mathbb{R}^N$, and individual exponential functions are not in L^2. Thus this easy process for obtaining M is not available, and we are led to construct M by successive approximations.

PROOF. Choose, by Proposition 8.12, functions $\Phi_k \in C^\infty_{\text{com}}(\mathbb{R}^N)$ with

(i) $0 \le \Phi_k(y) \le 1$ for all y,

(ii) $\Phi_k(y) = 0$ for $|y| \geq k+1$,
(iii) $\Phi_k(y) = 1$ for $|y| \leq k$.

Then $\Phi_j \Phi_k = \Phi_{\min(j,k)}$ if $j \neq k$, and Φ_k is in $C^\infty_{\text{com}}(\mathbb{R}^N)$ and hence in the Schwartz space $\mathcal{S}(\mathbb{R}^N)$. Put $\varphi_k = \mathcal{F}^{-1}(\Phi_k)$. Proposition 8.10 shows that φ_k is in $\mathcal{S}$, and therefore φ_k is in $L^1 \cap L^2$. Since the Fourier transform carries convolution into pointwise product, we have $\varphi_j * \varphi_k = \varphi_{\min(j,k)}$ if $j \neq k$. Define

$$M_k = \mathcal{F}(A\varphi_k)$$

as an L^2 function. Lemma 8.13 shows that A commutes with convolution by an L^1 function, and thus $\varphi_k * A\varphi_{k+1} = A(\varphi_k * \varphi_{k+1}) = A\varphi_k = A(\varphi_k * \varphi_{k+2}) = \varphi_k * A\varphi_{k+2}$. Consequently

$$\Phi_k M_{k+1} = \Phi_k M_{k+2}$$

and

$$M_{k+1}(y) = M_{k+2}(y) \quad \text{for } |y| \leq k.$$

This equation shows that if we put

$$M(y) = M_{k+1}(y) \quad \text{for } |y| \leq k,$$

then M is consistently defined and is locally in L^2.

Let $\mathcal{S}_0 = \mathcal{F}^{-1}(C^\infty_{\text{com}}(\mathbb{R}^N)) \subseteq \mathcal{S}(\mathbb{R}^N)$. If a member f of $\mathcal{S}_0$ has $\widehat{f}(y) = 0$ for $|y| \geq k$, then $\widehat{f}\Phi_{k+1} = \widehat{f}$ and hence $f * \varphi_{k+1} = f$. Application of A gives $Af = A(f * \varphi_{k+1}) = f * A\varphi_{k+1}$. If we take the L^2 Fourier transform of both sides and use Corollary 8.8, we obtain $\mathcal{F}(Af) = M_{k+1}\widehat{f}$. The right side equals $M\widehat{f}$ since $\widehat{f}(y) = 0$ for $|y| \geq k$, and thus

$$\mathcal{F}(Af) = M\widehat{f}$$

whenever f is in $\mathcal{S}_0$ and $\widehat{f}(y) = 0$ for $|y| \geq k$.

The subspace $C^\infty_{\text{com}}(\mathbb{R}^N)$ of L^2 is dense by Corollary 6.19 and Theorem 6.20a. Since the L^2 Fourier transform carries L^2 onto L^2 and preserves norms (Theorems 8.6 and 8.9), $\mathcal{S}_0$ is dense in L^2. Let a general f in L^2 be given, and choose a sequence $\{f_j\}$ in $\mathcal{S}_0$ with $f_j \to f$ in L^2. Then $\mathcal{F}(Af_j) \to \mathcal{F}(Af)$ in L^2. By Theorem 5.59 we can pass to a subsequence, still written as $\{f_j\}$, so that $f_j \to f$ and $\mathcal{F}(f_j) \to \mathcal{F}(f)$ and $\mathcal{F}(Af_j) \to \mathcal{F}(Af)$ almost everywhere. Consequently

$$\begin{aligned}\mathcal{F}(Af)(y) &= \lim \mathcal{F}(Af_j)(y) = \lim M(y)\mathcal{F}(f_j)(y) \\ &= M(y) \lim \mathcal{F}(f_j)(y) = M(y)\mathcal{F}(f)(y)\end{aligned}$$

almost everywhere.

To see that M is in L^∞, suppose that $|M(y)| \geq C$ occurs at least on a set E of positive finite measure. Then I_E is in L^2. If we put $f = \mathcal{F}^{-1}(I_E)$, then we have $\|A\|\|f\|_2 \geq \|Af\|_2 = \|\mathcal{F}(Af)\|_2 = \|M\mathcal{F}(f)\|_2 = \|MI_E\|_2 \geq C\|I_E\|_2 = C\|f\|_2$, and hence $\|A\| \geq C$. Therefore $\|A\| \geq \|M\|_\infty$. In particular, M is in L^∞.

In the reverse direction we have $\|Af\|_2 = \|\mathcal{F}(Af)\|_2 = \|M\mathcal{F}(f)\|_2 \leq \|M\|_\infty\|\mathcal{F}(f)\|_2 = \|M\|_\infty\|f\|_2$ for all f in L^2, and thus $\|A\| \leq \|M\|_\infty$. We conclude that $\|M\|_\infty = \|A\|$. This completes the proof of existence.

If we have two candidates for the multiplier, say M and M_1, then subtraction of the equations $\mathcal{F}(Af) = M\mathcal{F}(f)$ and $\mathcal{F}(Af) = M_1\mathcal{F}(f)$ shows that $0 = (M - M_1)\mathcal{F}(f)$ for all f in L^2. Therefore $M = M_1$ almost everywhere. This proves uniqueness. $\square$

5. Poisson Summation Formula

The Poisson Summation Formula is a result combining Fourier series and the Fourier transform in a way that has remarkable applications, both pure and applied. Nowadays the formula is expressed as a result about Schwartz functions and therefore fits at this particular spot in the discussion of the Fourier transform.

Part of the power of the formula comes about because it applies to more settings than originally envisioned. The Euclidean version applies to the additive group $\mathbb{R}^N$, the discrete subgroup of points with integer coordinates, and the quotient group equal to the product of circle groups. In this section we shall take $N = 1$ simply because a theory of Fourier series has been developed in this book only in one variable.

We begin by stating and proving the 1-dimensional version of the theorem.

Theorem 8.15 (Poisson Summation Formula). If f is in the Schwartz space $\mathcal{S}(\mathbb{R}^1)$, then

$$\sum_{n=-\infty}^{\infty} f(x+n) = \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi inx}.$$

PROOF. Define $F(x) = \sum_{n=-\infty}^{\infty} f(x+n)$. From the definition of $\mathcal{S}$, it is easy to check that this series is uniformly convergent on any bounded interval and also the series of k^{th} derivatives is uniformly convergent on any bounded interval for each k. Consequently the function F is well defined and smooth, and it is periodic of period one. We form its Fourier series, taking into consideration that the period is 1 rather than 2π; the relevant formulas for Fourier series when the period is L rather than 2π are

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{2\pi inx/L} \qquad \text{with } c_n = \frac{1}{L}\int_{-L/2}^{L/2} f(t)e^{-2\pi int/L}\,dt.$$

A smooth periodic function is the sum of its Fourier series, and thus

$$F(x) = \sum_{n=-\infty}^{\infty} \Big(\int_0^1 F(t) e^{-2\pi i n t}\, dt \Big) e^{2\pi i n x}. \qquad (*)$$

The Fourier coefficient in parentheses in $(*)$ is

$$\begin{aligned} \int_0^1 F(t) e^{-2\pi i n t}\, dt &= \int_0^1 \sum_{k=-\infty}^{\infty} f(t+k) e^{-2\pi i n t}\, dt \\ &= \sum_{k=-\infty}^{\infty} \int_0^1 f(t+k) e^{-2\pi i n t}\, dt \\ &= \sum_{k=-\infty}^{\infty} \int_k^{k+1} f(t) e^{-2\pi i n t}\, dt \\ &= \int_{-\infty}^{\infty} f(t) e^{-2\pi i n t}\, dt \\ &= \widehat{f}(n), \end{aligned}$$

and the theorem follows by substituting this equality into $(*)$. □

Corollary 8.16. $\sum_{n=-\infty}^{\infty} e^{-\pi r^{-2} n^2} = r \sum_{n=-\infty}^{\infty} e^{-\pi r^2 n^2}$ for any $r > 0$.

PROOF. The remarks before Theorem 8.3 show that it we define $\varphi(x) = e^{-\pi x^2}$ and $\varphi_\varepsilon(x) = \varepsilon^{-1}\varphi(\varepsilon^{-1}x)$, then $\widehat{\varphi_\varepsilon}(y) = \widehat{\varphi}(\varepsilon y)$. If we put $f(x) = r\varphi_r(x) = e^{-\pi r^{-2} x^2}$, then it follows that $\widehat{f}(y) = r e^{-\pi r^2 x^2}$. Applying Theorem 8.15 to this f and setting $x = 0$ gives the asserted equality. □

In one especially significant application of the 1-dimensional Euclidean version of the Poisson Summation Formula to pure mathematics, the remarkable identity in Corollary 8.16 can be combined with some elementary complex analysis to obtain a functional equation for the **Riemann zeta function**, which is initially defined for complex s with $\operatorname{Re} s > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \Big(1 - \frac{1}{p^s}\Big)^{-1}.$$

The functional equation relates $\zeta(s)$ to $\zeta(1-s)$. More precisely the function $\zeta(s)$ extends to be defined in a natural way[2] for all s in $\mathbb{C} - \{1\}$, and the functional equation is

$$\Lambda(1-s) = \Lambda(s), \qquad \text{where } \Lambda(s) = \zeta(s)\Gamma\big(\tfrac{1}{2}s\big)\pi^{-\frac{1}{2}s}.$$

This implication is just the beginning of a deep theory in which Fourier analysis, elementary complex analysis, algebraic number theory, and algebraic geometry come together to yield a vast array of surprising results about prime numbers.

[2]The natural way is as an analytic function in $\mathbb{C} - \{1\}$. The function has a simple pole at $s = 1$.

The extension of the definition of the Riemann zeta function to $\mathbb{C} - \{1\}$ and the derivation of the functional equation of the Riemann zeta function use elementary complex analysis as in Appendix B, and they will be carried out in Problems 19–27 at the end of the chapter.

In real-world applications the 1-dimensional Fourier transform is of great significance because of its interpretation in signal processing. A given function $f(t)$ on $\mathbb{R}^1$ is interpreted as the voltage of some signal, written as a function of time, and the Fourier transform $\widehat{f}(\omega)$ gives the components of the signal at each frequency ω. The Plancherel formula states the comforting fact that energy can be computed either by summing the contributions over time or by summing the contributions over frequency, and the result is the same. Convolutions are of special significance in the theory because they represent the effects of time-independent operations on the signal—such as the passage of the signal through a filter.

To make numerical computations, one takes some discretized version of the signal, obtained for example by rapid sampling over a long interval of time. The discrete signal, which may well be obtained at 2^n points for some n, is then regarded as periodic. In other words, the signal is really a function on a cyclic group of order 2^n. Computing a convolution involves multiplying each translate of one function by the other function at 2^n points, adding, and assembling the results. The number of steps is on the order of 2^{2n}. Alternatively, a convolution can be computed using Fourier transforms: One computes the Fourier transform of each function, does a pointwise multiplication of the new functions, and then computes an inverse Fourier transform. The pointwise multiplication involves only 2^n steps, which is relatively trivial compared with 2^{2n} steps. How many steps are involved in the computation of a Fourier transform? Naively it would seem that an exponential depending on y has to be multiplied by the value of the function at each point x and the results added, hence 2^{2n} steps. However, the mechanism of the Poisson Summation Formula contains a better way of carrying out the computation of the Fourier transform that involves only about $n2^n$ steps. The algorithm in question is known as the **fast Fourier transform** and is discussed in more detail in Problems 13–18 at the end of the chapter. The upshot is that the Poisson Summation Formula leads to a practical device that cuts down enormously on the cost of analyzing signals mathematically.

Although the Poisson Summation Formula as stated in this section relates the real line, the subgroup of integers, and the quotient circle group, the fast Fourier transform iterates versions of the formula for settings that are different from this. The groups in question are cyclic of order 2^k with $k \leq n$. We can take the subgroup to have order 2, and the quotient group then has order 2^{k-1}. A still more general version of the Poisson Summation Formula applies to any "locally compact" abelian group with a discrete subgroup having compact quotient. This more general version of the formula is used in the full-fledged application to pure mathematics that combines Fourier analysis, elementary complex analysis, algebraic number theory, and algebraic geometry.

6. Poisson Integral Formula

Let $\mathbb{R}^{N+1}_+$ be the open half space $\{(x,t) \mid x \in \mathbb{R}^N \text{ and } t > 0\}$. We view the boundary $\{(x,0) \mid x \in \mathbb{R}^N\}$ as $\mathbb{R}^N$. For a function f in $L^p(\mathbb{R}^N)$ for p equal to 1, 2, or ∞, we consider the problem of finding $u(x,t)$ that is defined on $\mathbb{R}^{N+1}_+$, has f as boundary value in a suitable sense, and is **harmonic** in the sense of being a C^2 function satisfying the **Laplace equation** $\Delta u = 0$, where Δ is the **Laplacian**

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_N^2} + \frac{\partial^2}{\partial t^2}.$$

We studied the corresponding problem for the unit disk in a sequence of problems at the ends of Chapters I, III, IV, and VI. In that situation the open disk played the role of the open half space, and the circle played the role of the Euclidean-space boundary. We were able to see that the unique possible answer, at least if f is of class C^2, is given by the Poisson integral formula for the unit disk:

$$u(r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta - \varphi) P_r(\varphi)\, d\varphi,$$

where $P_r(\theta) = \frac{1-r^2}{1-2r\cos\theta + r^2} = \sum_{n=-\infty}^{\infty} r^{|n|} e^{in\theta}$.

The situation with $\mathbb{R}^{N+1}_+$ is different. One complication is that the boundary is not compact, and a discrete sum can no longer be expected. Another is that the harmonic function with given boundary values need not be unique; in fact, the function $u(x,t) = t$ is a nonzero harmonic function with boundary values given by $f = 0$, and thus we cannot expect to get a unique solution to a boundary-value problem unless we impose some further condition on u. In effect, the condition we impose will amount to a growth condition on the behavior of u at infinity. A partial compensation for these two complications is that the boundary is now the Euclidean space $\mathbb{R}^N$, and dilations are available as a tool.

Let us make a heuristic calculation to look for a harmonic function with given boundary values. Suppose $u(x,t)$ is the solution we seek that corresponds to f. Then we expect that the translate $\tau_{x_0} u(x,t)$ is the solution corresponding to $\tau_{x_0} f(x)$. We might further expect that the mapping $f \mapsto u(\,\cdot\,, t)$ is bounded on $L^2(\mathbb{R}^N)$. By Theorem 8.14 we would therefore have

$$\widehat{u}(y,t) = m_t(y)\widehat{f}(y),$$

for some multiplier $m_t(y)$; the Fourier transform is to be understood as occurring in the x variable only. If $t_1 > 0$ is fixed, then $u(x, t+t_1)$ is harmonic with boundary value $u(x, t_1)$, and so $\widehat{u}(y, t+t_1) = m_t(y)\widehat{u}(y, t_1) = m_t(y)m_{t_1}(y)\widehat{f}(y)$. The left side equals $m_{t+t_1}(y)\widehat{f}(y)$, and therefore

$$m_{t+t_1}(y) = m_t(y)m_{t_1}(y).$$

Since this is only a heuristic calculation anyway, we might as well assume that m is jointly measurable. Then we deduce that

$$m_t(y) = e^{tg(y)}$$

for some L^∞ function g. To compute g, we use the condition $\Delta u = 0$ more explicitly. Formally, as a result of the Fourier inversion formula, $u(x, t)$ is given as

$$\int_{\mathbb{R}^N} \widehat{u}(y, t)e^{2\pi i x\cdot y}\, dy = \int_{\mathbb{R}^N} m_t(y)\widehat{f}(y)e^{2\pi i x\cdot y}\, dy = \int_{\mathbb{R}^N} e^{tg(y)}\widehat{f}(y)e^{2\pi i x\cdot y}\, dy.$$

Without regard to the validity of the interchange of limits, we differentiate under the integral sign to obtain

$$\frac{\partial^2}{\partial x_j^2} u(x, t) = -4\pi^2 \int_{\mathbb{R}^N} y_j^2 e^{tg(y)}\widehat{f}(y)e^{2\pi i x\cdot y}\, dy$$

and
$$\frac{\partial^2}{\partial t^2} u(x, t) = \int_{\mathbb{R}^N} g(y)^2 e^{tg(y)}\widehat{f}(y)e^{2\pi i x\cdot y}\, dy.$$

Summing the derivatives and taking into account that $\widehat{f}(y)$ is rather arbitrary, we conclude that $g(y)^2 = 4\pi^2|y|^2$. Since $m_t(y)$ is an L^∞ function, we expect that the negative square root is to be used for all y. Thus $g(y) = -2\pi|y|$. Therefore our guess for the multiplier is

$$m_t(y) = e^{-2\pi t|y|}.$$

This is an L^1 function, and we begin our investigation of the validity of this answer by computing its "inverse Fourier transform," to see what to expect for the form of the bounded linear operator $f \mapsto u(\,\cdot\,, t)$.

Lemma 8.17. For $t > 0$,

$$\int_{\mathbb{R}^N} e^{-2\pi t|y|}e^{2\pi i x\cdot y}\, dy = \frac{c_N\, t}{(t^2 + |x|^2)^{\frac{1}{2}(N+1)}}\,,$$

where $c_N = \pi^{-\frac{1}{2}(N+1)}\Gamma\big(\frac{N+1}{2}\big)$.

REMARK. The idea is to handle $t = 1$ first and then to derive the formula for other t's by taking dilations into account. For $t = 1$, we express $e^{-2\pi|y|}$ as an integral of dilates of $e^{-\pi|y|^2}$, and then the integral in question will be computable in terms of the known inverse Fourier transforms of dilates of $e^{-\pi|y|^2}$.

PROOF. In one dimension, direct calculation using calculus methods on the intervals $[0, +\infty)$ and $(-\infty, 0]$ separately gives

$$\int_{-\infty}^{\infty} e^{-2\pi|u|} e^{-2\pi i u v}\, du = \frac{1}{\pi}\,\frac{1}{1+v^2}.$$

Since $(1+v^2)^{-1}$ is integrable, the Fourier inversion formula in $\mathbb{R}^1$ (Theorem 8.4) then shows that

$$\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+v^2}\, e^{2\pi i u v}\, dv = e^{-2\pi|u|}.$$

Putting $u = |y|$ with y in $\mathbb{R}^N$ yields

$$e^{-2\pi|y|} = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{2\pi i v|y|}\,(1+v^2)^{-1}\, dv. \tag{$*$}$$

Any $\beta > 0$ has $\beta^{-1} = \int_0^\infty e^{-\beta s}\, ds$, and hence $(1+v^2)^{-1} = \pi \int_0^\infty e^{-(1+v^2)\pi s}\, ds$. Substitution for $(1+v^2)^{-1}$ in $(*)$, interchange of integrals by Fubini's Theorem, and use of the formula in $\mathbb{R}^1$ for the inverse Fourier transform of a dilate of $e^{-\pi v^2}$ gives

$$e^{-2\pi|y|} = \int_0^\infty e^{-\pi s}\Big[\int_{-\infty}^{\infty} e^{2\pi i v|y|} e^{-\pi v^2 s}\, dv\Big]\, ds = \int_0^\infty e^{-\pi s} s^{-1/2} e^{-\pi|y|^2/s}\, ds,$$

and this is our formula for $e^{-2\pi|y|}$ as an integral of dilates of $e^{-\pi|y|^2}$.

We multiply both sides by $e^{2\pi i x\cdot y}$, integrate, interchange the order of integration, use the formula in $\mathbb{R}^N$ for the inverse Fourier transform of a dilate of $e^{-\pi|y|^2}$, and make a change of variables $\pi s(1+|x|^2) \to s$. The result is

$$\begin{aligned}
\int_{\mathbb{R}^N} e^{-2\pi|y|} e^{2\pi i x\cdot y}\, dy &= \int_0^\infty e^{-\pi s} s^{-1/2}\Big[\int_{\mathbb{R}^N} e^{-\pi|y|^2/s} e^{2\pi i x\cdot y}\, dy\Big]\, ds \\
&= \int_0^\infty e^{-\pi s} s^{\frac{1}{2}(N-1)} e^{-\pi s|x|^2}\, ds \\
&= \int_0^\infty e^{-\pi s(1+|x|^2)} s^{\frac{1}{2}(N-1)}\, ds \\
&= \pi^{-\frac{1}{2}(N-1)}(1+|x|^2)^{-\frac{1}{2}(N-1)}\pi^{-1}(1+|x|^2)^{-1}\int_0^\infty e^{-s} s^{\frac{1}{2}(N-1)}\, ds \\
&= \pi^{-\frac{1}{2}(N+1)}\Gamma\big(\tfrac{N+1}{2}\big)(1+|x|^2)^{-\frac{1}{2}(N+1)}.
\end{aligned}$$

The proof is completed by making use of the effect of the Fourier transform on dilations. We have just seen that the function $\varphi(x) = (1 + |x|^2)^{-\frac{1}{2}(N+1)}$ is integrable with Fourier transform $c_N^{-1}\widehat{\varphi}(y) = c_N^{-1}e^{-2\pi|y|}$. Then $\varphi_t(x) = t^{-N}\varphi(t^{-1}x) = t(t^2+|x|^2)^{-\frac{1}{2}(N+1)}$ has Fourier transform $c_N^{-1}\widehat{\varphi}(ty) = c_N^{-1}e^{-2\pi t|y|}$. □

We define

$$P(x,t) = P_t(x) = \frac{c_N\, t}{(t^2 + |x|^2)^{\frac{1}{2}(N+1)}}$$

for $t > 0$, with c_N as in Lemma 8.17, to be the **Poisson kernel** for $\mathbb{R}^{N+1}_+$. The **Poisson integral formula** for $\mathbb{R}^{N+1}_+$ is $u(x,t) = (P_t * f)(x)$, and the function u is called the **Poisson integral** of f.

Proposition 8.18. The Poisson kernel for $\mathbb{R}^{N+1}_+$ has the following properties:

(a) $P_t(x) = t^{-n}P_1(t^{-1}x)$,
(b) P_t is integrable with $\widehat{P_t}(y) = e^{-2\pi t|y|}$,
(c) $P_t \geq 0$ and $\int_{\mathbb{R}^N} P_t(x)\,dx = 1$,
(d) $P_t * P_{t'} = P_{t+t'}$,
(e) $P(x,t)$ is harmonic in $N+1$ variables.

PROOF. Conclusion (a) is by inspection. For (b), the formula for P_t shows that P_t, for fixed t, is continuous and is of order $|y|^{-(N+1)}$ as y tends to infinity. Therefore P_t is integrable. The formula for $\widehat{P_t}$ then follows from Lemma 8.17 and the Fourier inversion formula (Theorem 8.4). In (c), the first conclusion is by inspection of the formula, and the second conclusion follows from (b) by setting $y = 0$. Conclusion (d) follows from the corresponding formula on the Fourier transform side, namely $\widehat{P_t}\widehat{P_t'} = \widehat{P_{t+t'}}$, and conclusion (e) may be verified by a routine computation. □

Theorem 8.19. Let p be 1, 2, or ∞, let f be in $L^p(\mathbb{R}^N)$, and let $u(x,t) = (P_t * f)(x)$ be the Poisson integral of f. Then

(a) u is harmonic in $N+1$ variables,
(b) $\|u(\,\cdot\,,t)\|_p \leq \|f\|_p$,
(c) $u(\,\cdot\,,t)$ converges to f in L^p as t decreases to 0 provided $p < \infty$,
(d) $u(x,t)$ converges to $f(x)$ uniformly for x in E as t decreases to 0 provided f is in L^∞ and f is uniformly continuous at the points of E,
(e) the maximal function $f^{**}(x) = \sup_{t>0} |(P_t * f)(x)|$ satisfies an inequality $m\big(\{x \mid f^{**}(x) > \xi\}\big) \leq C\|f\|_1\big/\xi$ with C independent of f and ξ,
(f) **(Fatou's Theorem)** $u(x,t)$ converges to $f(x)$ for almost every x in $\mathbb{R}^N$.

REMARKS. The theorem says that u is harmonic and has boundary value f in various senses. The hypothesis for (f) is really that f is the sum of an L^1 function and an L^∞ function, and every L^2 function has this property, as will be observed in the proof below.

PROOF. Let us leave aside (a) for the moment. Conclusion (b) is immediate from Proposition 6.14 and parts (a) and (c) of Proposition 8.18. Conclusions (c) and (d) follow from parts (a) and (c) of Theorem 6.20. Conclusion (e) follows from Corollary 6.42 and the Hardy–Littlewood Maximal Theorem (Theorem 6.38), and conclusion (f) for L^1 functions f is part of Corollary 6.42. Now suppose that f is an L^∞ function. Fix a bounded interval $[a, b]$ and write $f = f_1 + f_2$ with f_1 equal to 0 off $[a, b]$ and f_2 equal to 0 on $[a, b]$. Then $P_t * f_1$ converges almost everywhere to f_1 since f_1 is integrable, and $P_t * f_2$ converges to 0 everywhere on (a, b) by (d). Hence $P_t * f$ converges almost everywhere on (a, b); since (a, b) is arbitrary, $P_t * f$ converges almost everywhere. This proves (f).

Now we return to (a). Since $P(x, t)$ is harmonic, conclusion (a) represents an interchange of differentiation and convolution. The prototype of the tool we need is Corollary 6.19, but that result does not apply here because P_t does not have compact support. If we break a function f into two pieces, one where $|f|$ is > 1 and one where $|f|$ is ≤ 1, we see that any L^2 function is the sum of an L^1 function and an L^∞ function. Thus it is enough to prove (a) when f is in L^1 or L^∞.

Let φ be P or one of P's iterated partial derivatives of some order, let $1 \le j \le N + 1$, and define D_j to be $\partial/\partial x_j$ if $j \le N$ or $\partial/\partial t$ if $j = N + 1$. It is sufficient to check that

$$h^{-1}\big((\varphi * f)((x, t) + he_j) - (\varphi * f)(x, t)\big) - ((D_j\varphi) * f)(x, t)$$

tends to 0 pointwise as h tends to 0. Taking Proposition 6.15 into account, we see that we are to check that

$$\Big(h^{-1}\big(\varphi((\,\cdot\,, t) + he_j) - \varphi(\,\cdot\,, t)\big) - (D_j\varphi)(\,\cdot\,, t)\Big) * f(x)$$

tends to 0 as h tends to 0. Proposition 6.18 shows that it is enough to have

$$h^{-1}\big(\varphi((x, t) + he_j) - \varphi(x, t)\big) - (D_j\varphi)(x, t)$$

tend to 0 in L^∞ of the x variable for each fixed t if f is in L^1, or in L^1 of the x variable for fixed t if f is in L^∞. The Mean Value Theorem shows that this expression is equal to

$$(D_j\varphi)((x, t) + h'e_j) - (D_j\varphi)(x, t)$$

for some h' between 0 and h, with h' depending on x and t, and a second application of the Mean Value Theorem shows that the expression is equal to

$$h'(D_j^2\varphi)((x,t)+h''e_j).$$

We are to show that this tends to 0, for fixed t, uniformly in x and in L^1 of the x variable as h tends to 0. It is enough to show for each fixed t that

$$(D_j^2\varphi)((x,t)+he_j) \tag{$*$}$$

is dominated in absolute value by a fixed bounded function of x and a fixed L^1 function of x when h satisfies $|h| \leq \frac{1}{2}\min\{1,t\}$.

An easy induction on the degree shows that any d^{th}-order partial derivative of $P(x,t)$ is of the form $Q(x,t)(t^2+|x|^2)^{-\frac{1}{2}(N+1)-d}$, where $Q(x,t)$ is a homogeneous polynomial in (x,t) of degree $d+1$. Since any monomial of degree 1 is bounded by a multiple of $(t^2+|x|^2)^{1/2}$, the d^{th}-order partial derivative is bounded by a multiple of

$$(t^2+|x|^2)^{-\frac{1}{2}(N+1)-\frac{1}{2}(d-1)}. \tag{$**$}$$

Thus the desired properties of the expression $(*)$ will follow if it is shown that $(**)$ has these properties. This is a routine matter for $d \geq 1$, and the proof of (a) is complete. $\square$

7. Hilbert Transform

This section concerns the **Hilbert transform**, the bounded linear operator H on $L^2(\mathbb{R}^N)$ given by

$$\mathcal{F}(Hf)(y) = -i(\operatorname{sgn} y)(\mathcal{F}f)(y).$$

Formally this operator has the effect, for $y > 0$, of mapping exponentials by

$$e^{2\pi ix\cdot y} \mapsto -ie^{2\pi ix\cdot y} \quad \text{and} \quad e^{-2\pi ix\cdot y} \mapsto ie^{-2\pi ix\cdot y},$$

and hence of mapping cosines and sines by

$$\cos(2\pi x\cdot y) \mapsto \sin(2\pi x\cdot y) \quad \text{and} \quad \sin(2\pi x\cdot y) \mapsto -\cos(2\pi x\cdot y).$$

For this reason, engineers sometimes call the Hilbert transform a "90° phase shift." The notion is of such importance that there is even a piece of hardware called a "Hilbert transformer" that takes an input signal and produces some kind of approximation to the Hilbert transform of the signal.[3]

[3]The delay in time that a Hilbert transformer requires in producing its output imposes a built-in theoretical limit for how good the approximation to the Hilbert transform can be. An exact result would require an infinite time delay.

We shall do some Fourier analysis in order to identify H more directly. To get an idea what H is, we begin by computing the effect on L^2 of composing the Hilbert transform and convolution with the Poisson kernel $P_\varepsilon(x)$. Then we examine what happens when ε decreases to 0.

Lemma 8.20. For $\varepsilon > 0$, $\displaystyle\int_{\mathbb{R}^1} (-i \operatorname{sgn} t) e^{-2\pi\varepsilon|t|} e^{2\pi i x\cdot t}\,dt = \frac{1}{\pi}\frac{x}{\varepsilon^2 + x^2}$.

PROOF. This result follows by direct calculation, using calculus methods on the intervals $[0, +\infty)$ and $(-\infty, 0]$ separately. □

If we define $Q(x) = \dfrac{1}{\pi}\dfrac{x}{1+x^2}$ for x in $\mathbb{R}^1$, then $Q_\varepsilon(x) = \varepsilon^{-1}Q(\varepsilon^{-1}x) = \dfrac{1}{\pi}\dfrac{x}{\varepsilon^2+x^2}$ is the function in the statement of Lemma 8.20. We define

$$Q(x, \varepsilon) = Q_\varepsilon(x) = \frac{1}{\pi}\frac{x}{\varepsilon^2 + x^2},$$

for $\varepsilon > 0$, to be the **conjugate Poisson kernel** on $\mathbb{R}^2_+$. The function Q_ε is *not* in $L^1(\mathbb{R}^1)$. However, it is in $L^2(\mathbb{R}^1)$, and therefore the convolution of Q_ε and any L^2 function is a well-defined bounded uniformly continuous function. For f in L^2, the function $v(x, \varepsilon) = (Q_\varepsilon * f)(x)$ is called the **conjugate Poisson integral** of f.

Proposition 8.21. The conjugate Poisson kernel for $\mathbb{R}^2_+$ has the following properties:

(a) the function $v(x, y) = Q_y(x)$ is harmonic in $\mathbb{R}^2_+$, and the pair of functions u and v with $u(x, y) = P_y(x)$ satisfies the **Cauchy–Riemann equations**

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

(b) the L^2 Fourier transform $\mathcal{F}(Q_\varepsilon)(y)$ equals $-i(\operatorname{sgn} y)e^{-2\pi\varepsilon|y|}$,

(c) $Q_\varepsilon * P_{\varepsilon'} = Q_{\varepsilon+\varepsilon'}$.

REMARKS. A fundamental result of elementary complex analysis is that if u and v are C^1 functions on an open subset of $\mathbb{C}$ satisfying the Cauchy–Riemann equations, then $f(z) = u(x, y) + iv(x, y)$ is an "analytic" function in the sense that in any open disk about any point in the open set, $f(z)$ equals the sum of a power series convergent in that disk.[4] We shall not make use of elementary complex analysis at this time, but the analyticity of $u + iv$ is the starting point for a great deal of analysis that will not be treated in this book. In the special case of the Poisson kernel and the conjugate Poisson kernel, the function f is $f(z) = i/(\pi z)$.

[4] This result amounts to a combination of Corollary B.2 and Theorem B.21 of Appendix B.

PROOF. Part (a) is a routine calculation.

For (b), we know that Q_ε is in L^2 and has an L^2 Fourier transform $g = \mathcal{F}(Q_\varepsilon)$. The inverse Fourier transform $\mathcal{F}^{-1}$ on L^2 satisfies $\mathcal{F}^{-1}(g) = Q_\varepsilon$, and (b) will follow if we show that $\mathcal{F}^{-1}(f) = Q_\varepsilon$, where $f(t) = -i(\operatorname{sgn} t)e^{-2\pi\varepsilon|t|}$. For each integer $n > 0$, let $f_n(t)$ be $f(t)$ for $|t| \le n$ and 0 for $|t| > n$. Then $f_n \to f$ in L^2 by dominated convergence, and hence $\mathcal{F}^{-1}(f_n) \to \mathcal{F}^{-1}(f)$ in L^2. By Theorem 5.59 a subsequence of $\mathcal{F}^{-1}(f_n)$ converges almost everywhere to $\mathcal{F}^{-1}(f)$. Since f is in L^1, Lemma 8.20 shows that $\mathcal{F}^{-1}(f_n)(t) = \int_{-n}^{n} f(t)e^{2\pi ixt}\,dt$ converges pointwise to $Q_\varepsilon(x)$, and therefore $\mathcal{F}^{-1}(f) = Q_\varepsilon$.

For (c), Corollary 8.8 shows that $\mathcal{F}(Q_\varepsilon * P_{\varepsilon'}) = \mathcal{F}(Q_\varepsilon)\mathcal{F}(P_{\varepsilon'})$. In combination with Proposition 8.18b, conclusion (b) of the present proposition gives $\mathcal{F}(Q_\varepsilon)(y)\mathcal{F}(P_{\varepsilon'})(y) = -i(\operatorname{sgn} y)e^{-2\pi(\varepsilon+\varepsilon')|y|}$ a.e., and this is $\mathcal{F}(Q_{\varepsilon+\varepsilon'})(y)$ a.e. by a second application of (b). $\square$

Theorem 8.22. Let f be in $L^2(\mathbb{R}^1)$, and let $u(x, y) = (P_y * f)(x)$ and $v(x, y) = (Q_y * f)(x)$ be the Poisson integral and conjugate Poisson integral of f. Then

(a) the function v is harmonic in $\mathbb{R}^2_+$, and the pair of functions u and v satisfies the Cauchy–Riemann equations,
(b) the function $Q_\varepsilon * f$ is in $L^2(\mathbb{R}^1)$ for every $\varepsilon > 0$, and its L^2 Fourier transform is $\mathcal{F}(Q_\varepsilon * f)(y) = -i(\operatorname{sgn} y)e^{-2\pi\varepsilon|y|}\mathcal{F}(f)(y)$,
(c) $\|Q_\varepsilon * f\|_2 = \|P_\varepsilon * f\|_2 \le \|f\|_2$ for every $\varepsilon > 0$,
(d) $Q_\varepsilon * f \to H(f)$ in L^2 as ε decreases to 0.

PROOF. Conclusion (a) is handled just like Theorem 8.19a. In the proof of Theorem 8.19a, the integrability of P_ε did not play a role; it was the integrability of the iterated partial derivatives of P_ε (i.e., the case $d > 0$) that was important. The estimates involving such derivatives here are the same as in that case.

For (b), put $g = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f)$. This is an L^2 function since $\mathcal{F}(Q_\varepsilon)$ is in L^∞ by inspection and since $\mathcal{F}(f)$ is in L^2 by the Plancherel formula. Define $f_n = I_{B(n;0)}f$, so that each f_n is in $L^1 \cap L^2$ and also $f_n \to f$ in L^2. Since $\mathcal{F}(Q_\varepsilon)$ is in L^∞, the Plancherel formula shows that $g_n = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f_n)$ is in L^2 for each n and converges to g in L^2. Since f_n is in L^1 and Q_ε is in L^2, Corollary 8.8 gives $\mathcal{F}(Q_\varepsilon * f_n) = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f_n) = g_n$ for all n. Thus $Q_\varepsilon * f_n = \mathcal{F}^{-1}(g_n)$. We now let n tend to infinity. We know that $\|Q_\varepsilon * f_n - Q_\varepsilon * f\|_{\sup} \le \|Q_\varepsilon\|_2\|f_n - f\|_2$. Since Q_ε is in L^2 and $f_n \to f$ in L^2, $Q_\varepsilon * f_n$ converges uniformly to $Q_\varepsilon * f$. On the other hand, $\mathcal{F}^{-1}(g_n)$ converges to $\mathcal{F}^{-1}(g)$ in L^2, and Theorem 5.59 shows that a subsequence converges almost everywhere. Therefore $\mathcal{F}^{-1}(g) = Q_\varepsilon * f$. Consequently $\mathcal{F}(Q_\varepsilon * f) = g = \mathcal{F}(Q_\varepsilon)\mathcal{F}(f)$, and we obtain $\mathcal{F}(Q_\varepsilon * f)(y) = -i(\operatorname{sgn} y)e^{-2\pi\varepsilon|y|}\mathcal{F}(f)(y)$.

Conclusions (c) and (d) follow by taking L^2 Fourier transforms and using (b), Proposition 8.18b, and the Plancherel Theorem. This completes the proof. $\square$

To get a more direct formula for the Hilbert transform, we introduce the functions

$$h(x) = \begin{cases} \frac{1}{\pi x} & \text{for } |x| \geq 1, \\ 0 & \text{for } |x| < 1, \end{cases}$$

and

$$h_\varepsilon(x) = \varepsilon^{-1} h(\varepsilon^{-1} x) = \begin{cases} \frac{1}{\pi x} & \text{for } |x| \geq \varepsilon, \\ 0 & \text{for } |x| < \varepsilon. \end{cases}$$

Let $\psi(x) = Q(x) - h(x)$, so that $\psi_\varepsilon(x) = \varepsilon^{-1}\psi(\varepsilon^{-1}x) = Q_\varepsilon(x) - h_\varepsilon(x)$.

Lemma 8.23. The function ψ on $\mathbb{R}^1$ is integrable, and $\int_{-\infty}^{\infty} \psi(x)\,dx = 0$.

PROOF. For $|x| < 1$, we have $\psi(x) = Q(x) = \pi^{-1}x\big/(1+x^2)$. This is a continuous odd function, and therefore it is integrable on $[-1, 1]$ with integral 0. For $|x| \geq 1$, we have $\psi(x) = \pi^{-1}\big(\frac{x}{1+x^2} - \frac{1}{x}\big) = -\pi^{-1}\big(\frac{1}{x(1+x^2)}\big)$. This is an integrable function for $|x| \geq 1$; since it is an odd function, its integral is 0. $\square$

Theorem 8.24. Let h_ε be defined as above. If f is in $L^2(\mathbb{R}^1)$, then $h_\varepsilon * f$ is in $L^2(\mathbb{R}^1)$ for every $\varepsilon > 0$, and $h_\varepsilon * f \to H(f)$ in L^2 as ε decreases to 0.

REMARKS. More concretely the limit relation in the theorem is that

$$Hf(x) = \lim_{\substack{(\text{in } L^2 \text{ sense}) \\ \varepsilon \downarrow 0}} \frac{1}{\pi} \int_{|t| \geq \varepsilon} \frac{f(x-t)}{t}\,dt.$$

The integrand on the right side is the product of two L^2 functions on the set where $|t| \geq \varepsilon$, and it is integrable by the Schwarz inequality.

PROOF. We have $h_\varepsilon * f = Q_\varepsilon * f - \psi_\varepsilon * f$. The term $Q_\varepsilon * f$ is in L^2 by Theorem 8.22b, and the term $\psi_\varepsilon * f$ is in L^2 by Lemma 8.23 and Proposition 6.14. As ε decreases to 0, $Q_\varepsilon * f$ tends to Hf in L^2 by Theorem 8.22d, and $\psi_\varepsilon * f$ tends to 0 in L^2 by Theorem 6.20a. This completes the proof. $\square$

Now that we have the concrete formula of Theorem 8.24 for the Hilbert transform on L^2 functions, we can ask whether the Hilbert transform is meaningful on other kinds of functions. For example, we could ask, If we have some other vector space V of functions and $V \cap L^2(\mathbb{R}^1)$ is dense in V, can we extend H to V? The answer for $V = L^1(\mathbb{R}^1)$ is unfortunately negative. In fact, if f is in $L^1 \cap L^2$, then the Fourier transform $\widehat{f}$ will be continuous and the Fourier transform of Hf will have to be $-i(\operatorname{sgn} y)\widehat{f}$. If $\widehat{f}(0)$ is nonzero, then $-i(\operatorname{sgn} y)\widehat{f}$ is not continuous and cannot be the Fourier transform of an L^1 function.

However, in Chapter IX we shall introduce L^p spaces for $1 \le p \le \infty$, thereby extending the definitions we have already made for p equal to $1, 2$, and ∞. Toward the end of the chapter, we shall see that the Hilbert transform makes sense as a bounded linear operator on $L^p(\mathbb{R}^1)$ for $1 < p < \infty$. This boundedness is an indication that the Hilbert transform is not a completely wild transformation, and the result in question will be used in the problems at the end of Chapter IX to prove that the partial sums of the Fourier series of an L^p function on the circle converge to the function in L^p as long as $1 < p < \infty$.

Actually, this boundedness on L^p will be a consequence of a substitute result about L^1 that we shall prove now. Although the Hilbert transform is not a bounded linear operator on L^1, its approximations in the statement of Theorem 8.24 are of weak type $(1, 1)$, in the same sense that the passage from a function to its Hardy–Littlewood maximal function in Chapter VI was of weak type $(1, 1)$.

Theorem 8.25. Let h_1 be the function on $\mathbb{R}^1$ equal to $1/(\pi x)$ for $|x| \ge 1$ and equal to 0 for $|x| < 1$. For f in $L^1(\mathbb{R}^1) + L^2(\mathbb{R}^1)$, define

$$H_1 f(x) = h_1 * f(x) = \frac{1}{\pi} \int_{|t| \ge 1} \frac{f(x-t)}{t}\, dt$$

as the convolution of the fixed function h_1 in L^2 with a function f that is the sum of an L^1 function and an L^2 function. Then

$$\|H_1 f\|_2 \le A \|f\|_2,$$

with the constant A independent of f, and

$$m\big\{x \in \mathbb{R}^1 \bigm| |H_1 f(x)| > \xi\big\} \le \frac{C\|f\|_1}{\xi}$$

for every $\xi > 0$, with the constant C independent of ξ and f.

REMARK. This result about the approximation H_1 to H on L^1 and L^2 will be enough for now. The result for L^1 is much more difficult than the result for L^2. In the next chapter we shall derive from Theorem 8.25 a boundedness theorem for all the other approximations $H_\varepsilon = h_\varepsilon * (\,\cdot\,)$ on $L^p(\mathbb{R}^1)$ for $1 < p < \infty$, with a bound independent of ε, and then it will be an easy matter to get the boundedness of the Hilbert transform H itself on L^p for these values of p.

PROOF. A preliminary fact is needed that involves a computation with the function h_1. We need to know that

$$\int_{|x|\geq 2r} |h_1(x+r') - h_1(x)|\,dx \leq 6 \tag{$*$}$$

whenever $0 < |r'| \leq r$. To see this, we break the region of integration into four sets—one where $|x| \geq 2r$, $|x| \geq 1$, and $|x+r'| \geq 1$; a second where $|x| \geq 2r$, $|x| < 1$, and $|x+r'| \geq 1$; a third where $|x| \geq 2r$, $|x| \geq 1$, and $|x+r'| < 1$; and a fourth where $|x| \geq 2r$, $|x| < 1$, and $|x+r'| < 1$. For the fourth piece the integrand is 0. For the second and third pieces, the integrand is ≤ 1 in absolute value, and the set has measure ≤ 2; hence each of these pieces contributes at most 2. For the first piece the absolute value of the integrand is $|r'|\big/|x(x+r')| \leq 2r/x^2$; thus the absolute value of the integral is $\leq \int_{|x|\geq 2r} 2r/x^2\,dx = 2$. This proves $(*)$.

It is an easy matter to prove that H_1 is a bounded linear operator on L^2. In fact, $h_1 = Q_1 - \psi$, and ψ is in L^1 by Lemma 8.23. Thus Theorem 8.22c gives $\|H_1 f\|_2 \leq \|Q_1 * f\|_2 + \|\psi * f\|_2 \leq \|f\|_2 + \|\psi\|_1\|f\|_2$. In other words, H_1 is bounded on L^2 with $\|H_1\| \leq 1 + \|\psi\|_1$. Put $A = 1 + \|\psi\|_1$.

The heart of the proof is the observation that if F is in L^1, vanishes off a bounded interval I with center y_0 and double[5] I^*, and has total integral $\int_{\mathbb{R}^1} F(y)\,dy$ equal to 0, then

$$\|H_1 F\|_{L^1(\mathbb{R}-I^*)} \leq 6\|F\|_1. \tag{$**$}$$

To see this, we use the fact that the total integral of F is 0 to write

$$H_1 F(x) = \int_I h_1(x-y)F(y)\,dy = \int_I [h_1(x-y) - h(x-y_0)]F(y)\,dy.$$

Taking the absolute value of both sides and integrating over $\mathbb{R} - I^*$, we obtain

$$\begin{aligned}\int_{x\notin I^*} |H_1 F(x)|\,dx &\leq \int_{x\notin I^*}\int_{y\in I} |h_1(x-y) - h(x-y_0)|\,|F(y)|\,dy\,dx \\ &= \int_{y\in I}\Big[\int_{x\notin I^*} |h_1(x-y) - h(x-y_0)|\,dx\Big]|F(y)|\,dy \\ &\leq 6\int_{y\in I} |F(y)|\,dy,\end{aligned}$$

the last step holding by $(*)$. This proves $(**)$.

Let the L^1 function f be given. Fix $\xi > 0$. We shall decompose the L^1 function f into the sum $f = g + b$ of a "good" function g and a "bad" function b, in a manner dependent on ξ. The good function will be in L^∞ and hence will be in $L^1 \cap L^\infty \subseteq L^2$; the effect of applying H_1 to it will be controlled by the bound of H_1 on L^2. The bad function will be nonzero on a set of rather small measure, and we shall be able to control the effect of H_1 on it by means of $(**)$.

[5] The "double" of a bounded interval I is an interval of twice the length of I and the same center.

We begin by constructing a disjoint countable system of bounded open intervals I_k such that

(i) $\sum_k m(I_k) \le 5\|f\|_1/\xi$,
(ii) $|f(x)| \le \xi$ almost everywhere off $\bigcup_k I_k$,
(iii) $\frac{1}{m(I_k)} \int_{I_k} |f(y)|\,dy \le 2\xi$ for each n.

Namely, let $f^*(x) = \sup_{h>0} \frac{1}{2h} \int_{[x-h,x+h]} |f(t)|\,dt$ be the Hardy–Littlewood maximal function of f, and let E be the set where $f^*(x) > \xi$. The set E is open. In fact, if $f^*(x) > \xi$, then $\frac{1}{2h} \int_{[x-h,x+h]} |f(t)|\,dt \ge \xi + \epsilon$ for some $\epsilon > 0$. For $\delta > 0$, the inequality $\frac{1}{2h} \int_{[x-h,x+h+2\delta]} |f(t)|\,dt \ge \xi + \epsilon$ shows that $f^*(x+\delta) \ge \frac{h}{h+\delta}(\xi+\epsilon)$. Hence $f^*(x+\delta) > \xi$ for δ sufficiently small. Similarly $f^*(x-\delta) > \xi$ for δ sufficiently small.

Since E is open, E is uniquely the disjoint union of countably many open intervals, and these intervals will be the sets I_k. The disjointness of the I_k's and the Hardy–Littlewood Maximal Theorem (Theorem 6.38) together give

$$\sum_k m(I_k) \le m(E) \le 5\|f\|_1/\xi,$$

and this proves (i) and the boundedness of the intervals. The a.e. differentiability of integrals (Corollary 6.39) shows that $|f(x)| \le f^*(x)$ a.e., and therefore $|f(x)| \le \xi$ a.e. off $E = \bigcup_k I_k$. This proves (ii). If $I = (a,b)$ is one of the I_k's, then a is not in E, and we must have $\frac{1}{2(b-a)} \int_{[b-2(b-a),b]} |f(t)|\,dt \le f^*(a) \le \xi$. Therefore $\frac{1}{b-a} \int_{[a,b]} |f(t)|\,dt \le 2\xi$. This proves (iii).

With the open intervals I_k in hand, we define the decomposition $f = g + b$ by

$$g(x) = \begin{cases} \dfrac{1}{m(I_k)} \displaystyle\int_{I_k} f(y)\,dy & \text{for } x \in I_k, \\ f(x) & \text{for } x \notin \bigcup_k I_k, \end{cases}$$

$$b(x) = \begin{cases} f(x) - \dfrac{1}{m(I_k)} \displaystyle\int_{I_k} f(y)\,dy & \text{for } x \in I_k, \\ 0 & \text{for } x \notin \bigcup_k I_k. \end{cases}$$

Since $\{x \mid |H_1 f(x)| > \xi\} \subseteq \{x \mid |H_1 g(x)| > \xi/2\} \cup \{x \mid |H_1 b(x)| > \xi/2\}$, it is enough to prove

- $m(\{x \mid |H_1 g(x)| > \xi/2\}) \le C'\|f\|_1/\xi$ and
- $m(\{x \mid |H_1 b(x)| > \xi/2\}) \le C'\|f\|_1/\xi$

for some constant C' independent of ξ and f.

The definition of g shows that $\int_{I_k} |g(x)|\,dx \le \int_{I_k} |f(x)|\,dx$ for all k and that $|g(x)| = |f(x)|$ for $x \notin \bigcup_k I_k$; therefore $\int_{\mathbb{R}} |g(x)\,dx \le \int_{\mathbb{R}} |f(x)|\,dx$.

Also, properties (ii) and (iii) of the I_k's show that $|g(x)| \le 2\xi$ a.e. These two inequalities, together with the bound $\|H_1 g\|_2 \le A\|g\|_2$, give

$$\textstyle\int_{\mathbb{R}} |H_1 g(x)|^2\,dx \le A^2 \int_{\mathbb{R}^1} |g(x)|^2\,dx \le 2\xi A^2 \int_{\mathbb{R}} |g(x)|\,dx \le 2\xi A^2 \int_{\mathbb{R}} |f(x)|\,dx.$$

Combining this result with Chebyshev's inequality $m\big(\big\{x \bigm| |F(x)| > \beta\big\}\big) \le \beta^{-2}\int_{\mathbb{R}} |F(x)|^2\,dx$ for the function $F = H_1 g$ and the number $\beta = \xi/2$, we obtain

$$m\big(\big\{x \bigm| |H_1 g(x)| > \xi/2\big\}\big) \le \frac{4}{\xi^2}\, 2\xi A^2 \int_{\mathbb{R}} |f(x)|\,dx = \frac{8A^2\|f\|_1}{\xi}.$$

This proves the bulleted item about g.

For b, let b_k be the product of b with the indicator function of I_k. Then we have $b = \sum_k b_k$ with the sum convergent in L^1. Since H_1 is convolution by the L^2 function h_1, $H_1 b = \sum_k H_1 b_k$ with the sum convergent in L^2. Lumping terms via Theorem 5.59 if necessary, we may assume that the convergence takes place a.e. Therefore $|H_1 b(x)| \le \sum_k |H_1 b_k(x)|$ a.e. Using monotone convergence and $(**)$, we conclude that

$$\begin{aligned} \|H_1 b\|_{L^1\left(\mathbb{R}-\bigcup_k I_k^*\right)} &\le \sum_k \|H_1 b_k\|_{L^1\left(\mathbb{R}-\bigcup_j I_j^*\right)} \\ &\le \sum_k \|H_1 b_k\|_{L^1(\mathbb{R}-I_k^*)} \le 6\|b_k\|_1 = 6\|b\|_1 \le 6\|f\|_1. \end{aligned}$$

Thus $\quad m\big(\big\{x \in \mathbb{R} - \bigcup_k I_k^* \bigm| |H_1 b(x)| > \xi/2\big\}\big) \le 6\|f\|_1\big/(\xi/2) = 12\|f\|_1/\xi.$

Since $m\big(\big\{x \in \bigcup_k I_k^*\big\}\big) \le 5\|f\|_1\big/\xi$ by (i), we obtain $m\big(\big\{x \bigm| |H_1 b(x)| > \xi/2\}\big) \le 17\|f\|_1\big/\xi$, and the bulleted item about b follows. $\square$

8. Problems

1. For each of the following alternative definitions of the Fourier transform in $\mathbb{R}^N$, find a constant α such that the Fourier inversion formula is as indicated, and find a constant β such that when convolution is defined by

$$f * g(x) = \beta \textstyle\int_{\mathbb{R}^N} f(x-t)g(t)\,dt,$$

then the Fourier transform of the convolution is the product of the Fourier transforms.

(a) Fourier transform $\widehat{f}(y) = \int_{\mathbb{R}^N} f(x)e^{-ix\cdot y}\,dy$ and inverse Fourier transform $f(x) = \alpha \int_{\mathbb{R}^N} \widehat{f}(y)e^{ix\cdot y}\,dy$.

(b) Fourier transform $\widehat{f}(y) = (2\pi)^{-N} \int_{\mathbb{R}^N} f(x)e^{-ix\cdot y}\,dy$ and inverse Fourier transform $f(x) = \alpha \int_{\mathbb{R}^N} \widehat{f}(y)e^{ix\cdot y}\,dy$.

(c) Fourier transform $\widehat{f}(y) = (2\pi)^{-N/2} \int_{\mathbb{R}^N} f(x)e^{-ix\cdot y}\,dy$ and inverse Fourier transform $f(x) = \alpha \int_{\mathbb{R}^N} \widehat{f}(y)e^{ix\cdot y}\,dy$.

2. Let $(u, v)_2 = \int_{\mathbb{R}^N} u(x)\overline{v(x)}\,dx$ if u and v are in $L^2(\mathbb{R}^N)$, and let $\mathcal{F}$ denote the Fourier transform on $L^2(\mathbb{R}^N)$. Prove for every pair of functions f and g in L^2 that $(f, g)_2 = (\mathcal{F}(f), \mathcal{F}(g))_2$.

3. Prove that the Poisson kernel P and the conjugate Poisson kernel Q for $\mathbb{R}^2_+$ satisfy the identity $Q_\varepsilon * Q_{\varepsilon'} = P_{\varepsilon+\varepsilon'}$.

4. This problem is an analog for the Fourier transform of Problem 20c of Chapter VI concerning Fourier series and weak-star convergence. Weak-star convergence was defined in Section V.9.

(a) If f is in $L^\infty(\mathbb{R}^N)$ and P_t is the Poisson kernel, prove that $P_t * f$ converges to f weak-star against $L^1(\mathbb{R}^N)$ as t decreases to 0. In other words, prove that $\lim_{t\downarrow 0} \int_{\mathbb{R}^N} (P_t * f)(x)g(x)\,dx = \int_{\mathbb{R}^N} f(x)g(x)\,dx$ for every g in L^1.

(b) Theorem 8.19b shows that $\|P_t * f\|_\infty \le \|f\|_\infty$ if f is in $L^\infty(\mathbb{R}^N)$. Prove that $\lim_{t\downarrow 0} \|P_t * f\|_\infty = \|f\|_\infty$.

5. Let $\mathcal{M}^+(\mathbb{R}^N)$ be the space of finite Borel measures on $\mathbb{R}^N$. This problem introduces convolution and the Poisson integral formula for $\mathcal{M}^+(\mathbb{R}^N)$. Each finite Borel measure on $\mathbb{R}^N$ defines, by means of integration, a bounded linear functional on the normed linear space $C_{\text{com}}(\mathbb{R}^N)$ equipped with the supremum norm, and thus it is meaningful to speak of weak-star convergence of such measures against $C_{\text{com}}(\mathbb{R}^N)$.

(a) The convolution of a finite Borel measure μ on $\mathbb{R}^N$ with an integrable function f is defined by $(f * \mu)(x) = \int_{\mathbb{R}^N} f(x - y)\,d\mu(y)$. Define the **convolution** $\mu = \mu_1 * \mu_2$ of two members of $\mathcal{M}^+(\mathbb{R}^N)$ by $\mu(E) = \int_{\mathbb{R}^N} \mu_1(E - x)\,d\mu_2(x)$ for all Borel sets E. Check that the result is a Borel measure and that the definition for $f\,dx * \mu$, i.e., for the situation in which μ_1 and μ_2 are specialized so that $\mu_1 = f\,dx$ and $\mu_2 = \mu$, is consistent with the definition in the special case.

(b) With convolution of finite Borel measures on $\mathbb{R}^N$ defined as in (a), prove that $\int_{\mathbb{R}^N} g\,d(\mu_1 * \mu_2) = \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} g(x + y)\,d\mu_1(x)\,d\mu_2(y)$ for every bounded Borel function g on $\mathbb{R}^N$.

(c) Verify that $\|P_t * \mu\|_1 \le \mu(\mathbb{R}^N)$ if μ is in $\mathcal{M}^+(\mathbb{R}^N)$. Prove the limit formula $\lim_{t\downarrow 0} \|P_t * \mu\|_1 = \mu(\mathbb{R}^N)$.

(d) If μ is in $\mathcal{M}^+(\mathbb{R}^N)$, prove that the measures $(P_t * \mu)(x)\,dx$ converge to μ weak-star against $C_{\text{com}}(\mathbb{R}^N)$ as t decreases to 0. In other words, prove that $\lim_{t\downarrow 0} \int_{\mathbb{R}^N} (P_t * \mu)(x)g(x)\,dx = \int_{\mathbb{R}^N} g(x)\,d\mu(x)$ for every g in $C_{\text{com}}(\mathbb{R}^N)$.

Problems 6–12 examine the Fourier transform of a measure in $\mathcal{M}^+(\mathbb{R}^N)$, ultimately proving "Bochner's theorem" characterizing the "positive definite functions" on $\mathbb{R}^N$. They take for granted the **Helly–Bray Theorem**, i.e., the statement that if $\{\mu_n\}$ is a sequence in $\mathcal{M}^+(\mathbb{R}^N)$ with $\{\mu_n(\mathbb{R}^N)\}$ bounded, then there is a subsequence $\{\mu_{n_k}\}$ convergent to some member μ of $\mathcal{M}^+(\mathbb{R}^N)$ weak-star against $C_{\mathrm{com}}(\mathbb{R}^N)$. The Helly–Bray Theorem will be proved in something like this form in Chapter XI.

6. If μ is in $\mathcal{M}^+(\mathbb{R}^N)$, the **Fourier transform** of μ is defined to be the function $\widehat{\mu}(y) = \int_{\mathbb{R}^N} e^{-2\pi i x\cdot y}\, d\mu(x)$.
   (a) Prove that $\widehat{\mu}$ is bounded and continuous.
   (b) Prove that the Fourier transform of the delta measure at 0 does not vanish at infinity.
   (c) Prove that $\widehat{\mu_1 * \mu_2} = \widehat{\mu_1}\widehat{\mu_2}$ when convolution of finite measures is defined as in Problem 5.
   (d) By forming $\varphi_\varepsilon * \mu$, prove that $\widehat{\mu}$ can equal 0 for some μ in $\mathcal{M}^+(\mathbb{R}^N)$ only if $\mu = 0$.
7. A continuous function $F : \mathbb{R}^N \to \mathbb{C}$ is called **positive definite** if for each finite set of points $x_1, \dots, x_k$ in $\mathbb{R}^N$ and corresponding system of complex numbers $\xi_1, \dots, \xi_k$, the inequality $\sum_{i,j} F(x_i - x_j)\xi_i\overline{\xi_j} \geq 0$ holds. Prove that the continuous function F is positive definite if and only if the inequality $\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} F(x-y)g(x)\overline{g(y)}\,dx\,dy \geq 0$ holds for each member g of $C_{\mathrm{com}}(\mathbb{R}^N)$.
8. Prove that the Fourier transform of any member μ of $\mathcal{M}^+(\mathbb{R}^N)$ is a positive definite function.
9. Using sets of one and then two elements x_i in the definition of positive definite, prove that a positive definite function F must have $F(0) \geq 0$ and $|F(x)| \leq F(0)$ for all x.
10. Suppose that F is positive definite, that $\varphi \geq 0$ is in $L^1(\mathbb{R}^N)$, and that $\Phi(x) = \int_{\mathbb{R}^N} e^{2\pi i x\cdot y}\varphi(y)\,dy$. Prove that $F(x)\Phi(x)$ is positive definite.
11. Suppose that F is positive definite. Let $\varepsilon > 0$, and let φ be as in Problem 10, so that $\varphi(x) = \varepsilon^{-N}e^{-\pi\varepsilon^{-2}|x|^2}$ and $\Phi(x) = e^{-\pi\varepsilon^2|x|^2}$.
    (a) The function $F_0(x) = F(x)\Phi(x)$ is positive definite by Problem 10. Prove that it is integrable.
    (b) Using Problem 2 and the alternative definition of positive definite in Problem 7, prove that $\int_{\mathbb{R}^N} \widehat{F_0}(y)|\widehat{g}(y)|^2\,dy \geq 0$ for every function g in $C_{\mathrm{com}}(\mathbb{R}^N)$.
    (c) Deduce from (b) that the function $f_0 = \widehat{F_0}$ is ≥ 0.
    (d) Conclude from (c) that f_0 is integrable with $\int_{\mathbb{R}^N} f_0\,dy = F(0)$, hence that $f_0(y)\,dy$ is a finite Borel measure.
12. **(Bochner's Theorem)** By combining the results of the previous problems with the Helly–Bray Theorem, prove that each positive definite function on $\mathbb{R}^N$ is the Fourier transform of a finite Borel measure.

Problems 13–18 concern a version of the Fourier transform for finite abelian groups, along with the Poisson Summation Formula in that setting. They show for a cyclic group of order $m = pq$ that the use of the idea behind the Poisson Summation Formula makes it possible to compute a Fourier transform in about $pq(p+q)$ steps rather than the expected $m^2 = p^2q^2$ steps. This savings may be iterated in the case of a cyclic group of order 2^n so that the Fourier transform is computed in about $n2^n$ steps rather than the expected 2^{2n} steps. An organized algorithm to implement this method of computation is known as the **fast Fourier transform**.

13. Let G be a finite abelian group. A **multiplicative character** χ of G is a homomorphism of G into the circle group $\{e^{i\theta}\}$. If f and g are two complex-valued functions on G, their L^2 inner product is defined to be $\sum_{t\in G} f(t)\overline{g(t)}$.
    (a) Prove that the set of multiplicative characters of G forms an abelian group under pointwise multiplication, the identity element being the constant function 1 and the inverse of χ being $\overline{\chi}$. This group $\widehat{G}$ is called the **dual group** of G.
    (b) Prove that distinct multiplicative characters are orthogonal and hence that the members of $\widehat{G}$ form a linearly independent set.
    (c) Let J_m be the cyclic group $\{0, 1, 2, \dots, m-1\}$ of integers modulo m under addition, and let $\zeta_m = e^{2\pi i/m}$. For k in J_m define a multiplicative character χ_n of J_m by $\chi_n(k) = (\zeta_m^n)^k$. Prove that the resulting m multiplicative characters exhaust $\widehat{J_m}$ and that $\chi_n\chi_{n'} = \chi_{n+n'}$. Therefore $\widehat{J_m}$ is isomorphic to J_m. For Problems 16–18 below, it will be convenient to identify χ_n with $\chi_n(1) = \zeta_m^n$.
    (d) If G is a direct sum of cyclic groups of orders $m_1, \dots, m_r$, use (c) to exhibit $\prod_{j=1}^r m_j$ distinct members of $\widehat{G}$. Using (b) and the theorem that every finite abelian group is the direct sum of cyclic groups, conclude for any finite abelian group G that these members of $\widehat{G}$ exhaust $\widehat{G}$ and form a basis of $L^2(G)$.

14. Let G be a finite abelian group, and let $\widehat{G}$ be its dual group. The **Fourier transform** of a function f in $L^2(G)$ is the function $\widehat{f}$ on $\widehat{G}$ given by $\widehat{f}(\chi) = \sum_{t\in G} f(t)\overline{\chi(t)}$. Prove that the Fourier transform mapping carries $L^2(G)$ one-one onto $L^2(\widehat{G})$ and that the correct analog of the Fourier inversion formula is $f(t) = |G|^{-1}\sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(t)$, where $|G|$ is the order of G.

15. Let G be a finite abelian group, let H be a subgroup, and let G/H be the quotient group. If t is in G, write $\dot{t}$ for the coset of t in G/H. Let f be in $L^2(G)$ and define $F(\dot{t}) = \sum_{h\in H} f(t+h)$ as a function on G/H. Suppose that χ is a member of $\widehat{G}$ that is identically 1 on H, so that χ descends to a member $\dot{\chi}$ of $\widehat{G/H}$. By imitating steps in the proof of Theorem 8.15, prove that $\widehat{f}(\chi) = \widehat{F}(\dot{\chi})$.

16. Now suppose that $G = J_m$ with $m = pq$; here p and q need not be relatively prime. Let $H = \{0, q, 2q, \dots, (p-1)q\}$ be the subgroup of G isomorphic to J_p, so that $G/H = \{0, 1, 2, \dots, q-1\}$ is isomorphic to J_q. Prove that the characters χ of G identified as in Problem 13c with $\zeta_m^0, \zeta_m^p, \zeta_m^{2p}, \dots, \zeta_m^{(q-1)p}$ are the ones that are identically 1 on H and therefore descend to characters of G/H. Verify that the descended characters $\dot\chi$ are the ones identified with $\zeta_q^0, \zeta_q^1, \zeta_q^2, \dots, \zeta_q^{q-1}$. Consequently the formula $\widehat{f}(\chi) = \widehat{F}(\dot\chi)$ of the previous problem provides a way of computing $\widehat{f}$ at $\zeta_m^0, \zeta_m^p, \zeta_m^{2p}, \dots, \zeta_m^{(q-1)p}$ from the values of $\widehat{F}$. Show that if $\widehat{F}$ is computed from the definition of Fourier transform in Problem 14, then the number of steps involved in its computation is about q^2, apart from a constant factor. Show therefore that the total number of steps in computing $\widehat{f}$ at these special values of χ is therefore on the order of $q^2 + pq$.

17. In the previous problem show for each k with $0 \le k \le p-1$ that the value of $\widehat{f}$ at $\zeta_m^k, \zeta_m^{p+k}, \zeta_m^{2p+k}, \dots, \zeta_m^{(q-1)p+k}$ can be handled in the same way with a different F by replacing f by a suitable variant of f. Doing so for each k requires p times the number of steps detected in the previous problem, and therefore all of $\widehat{f}$ can be computed in about $p(q^2 + pq) = pq(p+q)$ steps.

18. Show how iteration of this process to compute the Fourier transform of each F, together with further iteration of this process, allows one to compute a Fourier transform for $J_{m_1 m_2 \cdots m_r}$ in about $m_1 m_2 \cdots m_r (m_1 + m_2 + \cdots + m_r)$ steps.

Problems 19–27 combine the Poisson Summation Formula of this chapter with elementary complex analysis as in Appendix B to establish the analytic continuation and functional equation of the Riemann zeta function. As in Section 5 the Riemann zeta function is defined initially for complex s with $\operatorname{Re} s > 1$ by

$$\zeta(s) = \sum_{n=1}^{\infty} \tfrac{1}{n^s} = \prod_{p \text{ prime}} \big(1 - \tfrac{1}{p^s}\big)^{-1}.$$

Problems 19–20 establish that $\zeta(s)$ extends to be a meromorphic function defined for $\operatorname{Re} \zeta > 0$ in such a way that its only pole is simple and occurs at $s = 1$. Afterward Problems 21–27 establish that ζ extends analytically to all of $\mathbb{C} - \{1\}$ in such a way that

$$\Lambda(1-s) = \Lambda(s), \qquad \text{where } \Lambda(s) = \zeta(s)\Gamma\big(\tfrac12 s\big)\pi^{-\frac12 s}.$$

It is to be understood in this statement that Γ has been extended to a meromorphic function on $\mathbb{C}$ whose only poles are at the nonpositive integers. This result was established in Problems 26–27 at the end of Chapter VI.

19. By manipulating the defining formula for $\zeta(s)$, show for $\operatorname{Re} s > 1$ that

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=1}^{\infty} \int_n^{n+1} (n^{-s} - t^{-s})\, dt.$$

20. Prove that the series on the right side in Problem 19 converges uniformly on each compact subset of s with $\operatorname{Re} s > 0$, and deduce that $\zeta(s) - \frac{1}{s-1}$ extends to be analytic for $\operatorname{Re} s > 0$.

21. Let $\tau = \rho + i\sigma$ be a complex variable, and define

$$\theta(\tau) = \sum_{n=-\infty}^{\infty} e^{in^2\pi\tau} = 1 + 2\sum_{n=1}^{\infty} e^{in^2\pi\tau}.$$

Show that $\theta(\tau)$ is an analytic function for $\sigma = \operatorname{Im}\tau > 0$ and that $\theta(\tau+2) = \theta(\tau)$.

22. Derive the formula $\theta(-1/(i\sigma)) = \sigma^{1/2}\theta(i\sigma)$ from the special case of the Poisson Summation Formula given in Corollary 8.16. Here the branch of the square root is understood to be the principal branch. Explain how it follows that

$$\theta(-1/\tau) = (\tau/i)^{1/2}\theta(\tau) \qquad \text{for } \operatorname{Im}\tau > 0.$$

23. Show by a change of variables that

$$\int_0^\infty e^{-n^2\pi\sigma}\sigma^{\frac{1}{2}s-1}\,d\sigma = n^{-s}\Gamma(\tfrac{1}{2}s)\pi^{-\frac{1}{2}s} \qquad \text{for } \operatorname{Re} s > 1.$$

24. Sum the identity in the previous problem over n, and justify the conclusion that

$$\Lambda(s) = \zeta(s)\Gamma(\tfrac{1}{2}s)\pi^{-\frac{1}{2}s} = \int_0^\infty \tfrac{1}{2}[\theta(i\sigma) - 1]\sigma^{\frac{1}{2}s-1}\,d\sigma \qquad \text{for } \operatorname{Re} s > 1.$$

25. By estimating the series for $\frac{1}{2}[\theta(i\sigma) - 1]$, show that $\int_1^\infty \frac{1}{2}[\theta(i\sigma) - 1]\,\sigma^{\frac{1}{2}s-1}\,d\sigma$ converges for all complex s and defines an entire function $h(s)$.

26. Applying the result of Problem 22 and then making the change of variables $\sigma \mapsto 1/\sigma$ prove that

$$\int_0^1 \tfrac{1}{2}\theta(i\sigma)\sigma^{\frac{1}{2}s-1}\,d\sigma = \int_1^\infty \tfrac{1}{2}[\theta(i\sigma) - 1]\,\sigma^{\frac{1}{2}(1-s)-1}\,d\sigma - \frac{1}{1-s}$$

for $\operatorname{Re} s > 1$.

27. Combining the results of Problems 24, 25, and 26, obtain the conclusion that

$$\Lambda(s) = h(s) + h(1-s) - \frac{1}{1-s} - \frac{1}{s} \qquad \text{for } \operatorname{Re} s > 1.$$

with $h(s)$ entire. Conclude that $\Lambda(s)$ extends to a meromorphic function in $\mathbb{C}$ whose only possible poles are at 0 and 1 and that $\Lambda(1-s) = \Lambda(s)$. Taking for granted that $\Gamma(s)$ is nowhere 0, a fact that would follow by first proving the identity $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$, deduce that $\zeta(s)$ extends to a meromorphic function on $\mathbb{C}$ whose only pole is at $s = 1$.