THE UNCERTAINTY PRINCIPLE FOR FOURIER TRANSFORMS ON THE REAL LINE

MITCH HILL

Abstract. This paper will explore the heuristic principle that a function on the line and its Fourier transform cannot both be concentrated on small sets. We begin with the basic properties of the Fourier transform and show that a function and its Fourier transform cannot both have compact support. From there we prove the Fourier inversion theorem and use this to prove the classical uncertainty principle which shows that the spread of a function and its Fourier transform are inversely proportional. Finally, we extend our compactness result from earlier and show that a function and its Fourier transform cannot both be supported on finite sets.

Contents

1. Introduction

For certain well-behaved functions from the real line to the complex plane, one can define a related function which is known as the *Fourier transform*. The Fourier transform of a function $f\colon \mathbb{R} \to \mathbb{C}$ is formally defined as

$$\mathcal{F}[f](\xi) = \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)\, e^{-2\pi i x\xi}\, dx\,, \quad \xi \in \mathbb{R}\,.$$

Heuristically, the Fourier transform of a function has (most of) the same properties as the original function, so that no information is lost. The utility of the Fourier transform lies in the fact that some problems are much easier to tackle in the dual domain in ξ than in the original domain in x. However, we shall see that the concentration properties of a function do not carry over to its Fourier transform. In particular, the Fourier transform "smears out" functions.

An important and famous result by Heisenberg and Bernstein, often called the Uncertainty Principle, states that the "spread" of a function and its Fourier transform are inversely

Date: August 30, 2013.

proportional – that is, if the majority of the mass of the original function is clustered tightly in one area, the mass of the Fourier transform of that function must be spread more widely over the line. We state the quantitative version of this result as our first theorem. (In the theorem, $\mathcal{S}(\mathbb{R})$ refers to Schwartz functions on the real line, a restriction which will become clear later.)

Theorem 1.1. *(The Uncertainty Principle) For any $f \in \mathcal{S}(\mathbb{R})$ and any $x_0, \xi_0 \in \mathbb{R}$, we have the following inequality:*

$$\|f(x)\|_2^2 \leq 4\pi \|(x - x_0)f(x)\|_2 \, \|(\xi - \xi_0)\hat{f}(\xi)\|_2 \, . \tag{1.2}$$

Once the uncertainty principle has been established, one can ask more questions about the Fourier transform of functions with different kinds of support. If a function has finite support, so that the function is non-zero only on a set of finite measure, one might suspect from the uncertainty principle that the support of the Fourier transform of such a function must be larger, and therefore infinite, since a function of finite support has its mass concentrated in a small, finite area. This intuition turns out to be correct, and studying this problem leads us to the second major result of the paper.

Theorem 1.3. *(Amrein-Berthier) Let $f \in \mathcal{S}(\mathbb{R})$ and E, $F \subset \mathbb{R}$ be sets of finite measure. Then*

$$\|f\|_{L^2(\mathbb{R})} \leq C(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)}) \tag{1.4}$$

for some constant C that depends only on E and F.

Note that this results implies that if both f and $\hat{f}$ have finite support, then we must have $f = 0$ a.e., which means that the Fourier transform of a function of finite support must have infinite support.

2. The Schwartz Class

Before we define the Fourier transform and give its basic properties, we will define a class of functions for which the behavior of the Fourier transform is particularly nice. This set of functions is known as the Schwartz class, which can be thought of as smooth functions that vanish rapidly towards infinity. The formal definition is given below.

Definition 2.1. We say a function $f : \mathbb{R} \to \mathbb{C}$ is in the *Schwartz class*, denoted by $\mathcal{S}(\mathbb{R})$, if $f \in C^\infty(\mathbb{R})$ and, for all $m, n \in \mathbb{N}$, there exists a constant $C_{m,n} > 0$ such that

$$\rho_{m,n}(f) := \sup_{x \in \mathbb{R}} |x^m f^{(n)}(x)| = C_{m,n} < \infty \, . \tag{2.2}$$

The values $\rho_{m,n}(f)$ are called the *Schwartz semi-norms* of f.

Remark 2.3. It is clear that $\mathcal{S}(\mathbb{R})$ is closed under addition of functions, multiplication of a function by a scalar, and differentiation of a function. The product rule can be used to show that $\mathcal{S}(\mathbb{R})$ is also closed under multiplication of functions.

Examples 2.4. The following are examples of some functions which are in the Schwartz class and some functions which are not in the Schwartz class.

(1) Let f be the Gaussian function $f(x) = Ce^{-\beta x^2}$, with $\beta > 0$. Then $f \in \mathcal{S}(\mathbb{R})$, because e^{-x^2} is infinitely differentiable on $\mathbb{R}$ and decays faster than any power of x. We will see later that the Gaussians form a special subset of $\mathcal{S}(\mathbb{R})$, because they are the only functions for which equality holds in the uncertainty principle inequality.

(2) Let $f(x) = e^{-|x|}$. Then $f \notin \mathcal{S}(\mathbb{R})$, because f is not differentiable at 0.
(3) Let $f(x) = (1 + x^{2k})^{-1}$ for some $k \in \mathbb{Z}^+$. Then $f \notin \mathcal{S}(\mathbb{R})$, since f does not decay faster than the power x^{2k+1}.
(4) Let $f \in \mathcal{S}(\mathbb{R})$ and let P be a polynomial. Then $Pf \in \mathcal{S}(\mathbb{R})$, since f already decays faster than any power of x.
(5) Let $C_0^\infty(\mathbb{R})$ denote the set of C^∞ functions on $\mathbb{R}$ with compact support. Then $C_0^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$, because if $f \in C_0^\infty(\mathbb{R})$, then

$$\sup_{x\in\mathbb{R}} |x^m f^{(n)}(x)| = \sup_{-R\le x\le R} |x^m f^{(n)}(x)| \le R^m \left(\sup_{-R\le x\le R} |f^{(n)}(x)| \right) < \infty ,$$

so we have $f \in \mathcal{S}(\mathbb{R})$.

Although $\mathcal{S}(\mathbb{R})$ is not a normed linear space, the semi-norms $\rho_{m,\,n}$ can be used to make it into a locally convex complete metric space, otherwise known as a Fréchet space. We will not need this characterization of $\mathcal{S}(\mathbb{R})$ in this paper, but still we will define convergence in $\mathcal{S}(\mathbb{R})$ using these semi-norms and show that this notion of convergence is stronger than convergence in any L^p space.

Definition 2.5. Let $\{f_k\}_{k=1,\,2,\,\dots}$ and f be functions in $\mathcal{S}(\mathbb{R})$. We say that $\{f_k\}$ converges to f in $\mathcal{S}(\mathbb{R})$ if, for all $m,\, n \in \mathbb{N}$, we have

$$\rho_{m,\,n}(f_k - f) = \sup_{x\in\mathbb{R}} \left| x^m \left(f_k^{(n)}(x) - f^{(n)}(x) \right) \right| \to 0 \tag{2.6}$$

as $k \to \infty$.

Proposition 2.7. *Let $\{f_k\}_{k=1,\,2,\,\dots}$ and f be functions in $\mathcal{S}(\mathbb{R})$. If $f_k \to f$ in $\mathcal{S}(\mathbb{R})$, then $f_k \to f$ in $L^p(\mathbb{R})$ for all $1 \le p \le \infty$. Moreover, there exists a constant $C_p > 0$ such that, for all $f \in \mathcal{S}(\mathbb{R})$, we have*

$$\|f^{(n)}\|_p \le C_p \left(\|f^{(n)}\|_\infty + \rho_{\lfloor 2/p \rfloor + 1,\, n}(f) \right) < \infty . \tag{2.8}$$

Remark 2.9. Observe that (2.8) implies $\mathcal{S}(\mathbb{R}) \subset L^p(\mathbb{R})$ for all $1 \le p \le \infty$, which can be seen by letting $n = 0$.

Proof. Take some $g \in \mathcal{S}(\mathbb{R})$. Then

$$\begin{aligned} \|g\|_p &\le \left(\int_{|x|<1} \|g\|_\infty^p \, dx + \int_{|x|\ge 1} x^2 \, |g(x)|^p \, x^{-2} \, dx \right)^{1/p} \\ &\le \left(2\|g\|_\infty^p + \sup_{|x|\ge 1} \left\{ |x|^2 \, |g(x)|^p \right\} \int_{|x|\ge 1} x^{-2} \, dx \right)^{1/p} \\ &\le C_p \left(\|g\|_\infty + \sup_{x\in\mathbb{R}} |x|^{\lfloor 2/p \rfloor + 1} |g(x)| \right) . \end{aligned}$$

(2.8) immediately follows if we let $f^{(n)} = g$, and the convergence result follows by letting $g_k = f_k - f$ and applying the definition of convergence for Schwartz functions. ☐

3. The Fourier Transform and Basic Properties

Now that we are familiar with the Schwartz class, we give the definition of the Fourier transform and develop some of its basic properties. We conclude the section by giving the first of our "uncertainty" results, which states that the only function of compact support with a Fourier transform of compact support is the zero function.

Definition 3.1. Let $f \in L^1(\mathbb{R})$. Then the Fourier transform of f, denoted by $\mathcal{F}[f]$ or $\hat{f}$, is defined as follows:

$$\mathcal{F}[f](\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} f(x)\, e^{-2\pi i x\xi}\, dx, \quad \xi \in \mathbb{R}\,. \tag{3.2}$$

Observe that this function is well-defined, since

$$|\hat{f}(\xi)| \le \int_{\mathbb{R}} \left| f(x)\, e^{-2\pi i x\xi} \right| dx = \|f\|_1, \tag{3.3}$$

which is finite if $f \in L^1(\mathbb{R})$. One can also show that $\hat{f}$ is continuous, which is done as follows. Take some $\varepsilon > 0$. By the Dominated Convergence Theorem, the sequence $\{f|_{\mathrm{B}(0,N)}\}_{N\in\mathbb{N}}$ restricting f to the ball $\mathrm{B}(0,N)$ converges to f in $L^1(\mathbb{R})$. Therefore, there exists an $R > 0$ such that $\int_{|x|>R} |f(x)|\, dx < \frac{\varepsilon}{4}$. Now we have

$$\begin{aligned} |\hat{f}(\xi+h) - \hat{f}(\xi)| &= \left| \int_{\mathbb{R}} f(x)\, e^{-2\pi i\xi x}\, (e^{-2\pi i x h} - 1)\, dx \right| \\ &\le \int_{|x|\le R} |f(x)|\, |e^{-2\pi i x h} - 1|\, dx + \int_{|x|>R} |f(x)|\, |e^{-2\pi i x h} - 1|\, dx\,. \end{aligned}$$

In general, we have the inequalities $|e^{it} - 1| \le |t|$ and $|e^{it} - 1| \le 2$ for all $t \in \mathbb{R}$. Therefore, if $|h| \le \delta := \frac{\varepsilon}{4\pi R\|f\|_1}$ and $|x| \le R$, then $|e^{-2\pi i x h} - 1| \le \frac{\varepsilon}{2\|f\|_1}$ and we have

$$|\hat{f}(\xi+h) - \hat{f}(\xi)| \le \frac{\varepsilon}{2\|f\|_1} \int_{|x|\le R} |f(x)|\, dx + 2\int_{|x|>R} |f(x)|\, dx \;\le\; \varepsilon\,.$$

Since δ doesn't depend on the value of ξ, this shows that $\hat{f}$ is uniformly continuous.

Note that since $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$ by the previous section, the above definition of the Fourier transform extends to the Schwartz class. It is important to note that $\hat{f}$ is not necessarily in $L^1(\mathbb{R})$ even if f is in $L^1(\mathbb{R})$, so in general $\mathcal{F}[\hat{f}]$ might not be well-defined. In later sections of the paper, we will want to invert the Fourier transform to reclaim our original function, but this inversion is not always possible for a function that is simply L^1. In this section we will prove that the Fourier transform of a Schwartz function is also in the Schwartz class, which means we can find the Fourier transform of the Fourier transform of our original function, so that the inversion process we will describe later is well-defined.

Next we will explicitly calculate the Fourier transform of a Gaussian function.

Example 3.4. Let f be the Gaussian $f(x) = e^{-\pi x^2}$. Plugging f into (3.2) gives

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-\pi x^2} e^{-2\pi i x\xi}\, dx = \int_{\mathbb{R}} e^{-\pi(x+i\xi)^2} e^{\pi(i\xi)^2}\, dx = e^{-\pi\xi^2} \int_{-\infty}^{\infty} e^{-\pi(x+i\xi)^2}\, dx$$

where the second inequality follows from completing the square. The function $g(z) = e^{-\pi z^2}$ is holomorphic on $\mathbb{C}$, so integrating along the boundary of the rectangle

$$R_M = \{z \in \mathbb{C} : -M \le \mathrm{Re}(z) \le M \;\text{ and }\; 0 \le \mathrm{Im}(z) \le \xi\}$$

gives $\int_{\partial R_M} g(z)\,dz = 0$. (If $\xi < 0$, then take the boundaries of R_M so that $\xi \le \mathrm{Im}(z) \le 0$.) Therefore we have

$$\left(\int_{-M}^{M} e^{-\pi x^2}\,dx - \int_{-M}^{M} e^{-\pi(x+i\xi)^2}\,dx\right) + \left(\int_{0}^{\xi} e^{-\pi(-M+iy)^2}\,dy - \int_{\xi}^{0} e^{-\pi(M+iy)^2}\,dz\right) = 0\,.$$

As $M \to \infty$, the terms in parentheses on the right both go to zero, which means that

$$\int_{-\infty}^{\infty} e^{-\pi(x+i\xi)^2}\,dx = \int_{-\infty}^{\infty} e^{-\pi x^2}\,dx = 1$$

for any $\xi \in \mathbb{R}$. Therefore we have $\hat{f}(\xi) = e^{-\pi\xi^2} = f(\xi)$, so there is a non-zero function f whose Fourier transform is itself.

The following proposition gives some of the important basic properties of the Fourier transform.

Proposition 3.5. *Let $f \in \mathcal{S}(\mathbb{R})$, $\alpha \in \mathbb{R}$, and $n, m \in \mathbb{N}$. Let $\tilde{f}(x) = f(-x)$ and $f_a(x) = f(x-a)$. Then:*

(1) $\|\hat{f}\|_\infty \le \|f\|_1$
(2) $\widehat{f+g} = \hat{f} + \hat{g}$
(3) $\widehat{\alpha f} = \alpha \hat{f}$
(4) $\hat{\tilde{f}} = \tilde{\hat{f}}$
(5) $\hat{\bar{f}} = \bar{\tilde{\hat{f}}}$
(6) $\widehat{f_a}(\xi) = e^{-2\pi i a\xi}\,\hat{f}(\xi)$
(7) $(e^{2\pi i x a}\,f(x))\hat{\ }(\xi) = (\hat{f})_a(\xi)$
(8) $(f(ax))\hat{\ }(\xi) = \frac{1}{a}\,\hat{f}(\xi/a)$
(9) $\lim_{|\xi|\to\infty} |\hat{f}(\xi)| = 0$
(10) $\widehat{f^{(n)}}(\xi) = (2\pi i\xi)^n \hat{f}(\xi)$
(11) $(\hat{f})^{(n)}(\xi) = ((-2\pi i x)^n\,f(x))\hat{\ }(\xi)$
(12) $\hat{f} \in \mathcal{S}(\mathbb{R})$

Remark 3.6. Properties (1) – (9) hold for $f \in L^1(\mathbb{R})$ as well.

Proof. (1) follows from the fact that (3.3) holds for any $\xi \in \mathbb{R}$. (2) and (3) follow from the linearity of integration. For (4), we see that

$$\int_{\mathbb{R}} f(-x)\,e^{-2\pi i x\xi}\,dx = \int_{\mathbb{R}} f(u)\,e^{2\pi i u\xi}\,du = \hat{f}(-\xi)\,.$$

For (5), we have

$$\int_{\mathbb{R}} \overline{f(x)}\,e^{-2\pi i x\xi}\,dx = \int_{\mathbb{R}} \overline{f(x)\,e^{2\pi i x\xi}}\,dx = \overline{\hat{f}(-\xi)}\,.$$

Part (6) follows from writing

$$\int_{\mathbb{R}} f(x-a)\,e^{-2\pi i x\xi}\,dx = e^{-2\pi i a\xi}\int_{\mathbb{R}} f(x-a)\,e^{-2\pi i(x-a)\xi}\,dx = e^{-2\pi i a\xi}\int_{\mathbb{R}} f(u)\,e^{-2\pi i u\xi}\,dx$$

and (7) follows similarly from

$$\int_{\mathbb{R}} e^{2\pi i x a}\,f(x)\,e^{-2\pi i x\xi}\,dx = \int_{\mathbb{R}} f(x)\,e^{-2\pi i x(\xi-a)}\,dx = \hat{f}(\xi-a)\,.$$

To obtain (8), we can write

$$\int_{\mathbb{R}} f(ax)\, e^{-2\pi i x\xi}\, dx = \frac{1}{a}\int_{\mathbb{R}} f(u)\, e^{-2\pi i u(\xi/a)}\, du = \frac{1}{a}\,\hat{f}(\xi/a)\,.$$

To prove (9), first observe that

$$\lim_{|\xi|\to\infty}\int_a^b e^{-2\pi i x\xi}dx = \lim_{|\xi|\to\infty}\frac{e^{-2\pi i b\xi}-e^{-2\pi i a\xi}}{-2\pi i\xi} = 0$$

Therefore (9) holds for simple functions $\psi = \sum_{k=1}^{n} c_k \chi_{I_k}$ where I_k are finite disjoint intervals. Take $\varepsilon > 0$. Since the family of simple functions is dense in L^1, there exists a simple function ψ such that $\|\psi - f\|_1 < \varepsilon/2$, and there exists an $R > 0$ such that $|\int_{\mathbb{R}} \psi(x)\, e^{-2\pi i x\xi}\, dx| < \varepsilon/2$ if $|\xi| > R$. Now if $|\xi| > R$, then

$$|\hat{f}(\xi)| = \left|\int_{\mathbb{R}} f(x)\, e^{-2\pi i x\xi}\, dx\right| \le \int_{\mathbb{R}} |f(x)-\psi(x)|\, dx + \left|\int_{\mathbb{R}} \psi(x)\, e^{-2\pi i x\xi}\, dx\right| < \varepsilon\,.$$

For (10), we will prove the case $n = 1$, and the rest follows from the same proof using induction. Using integration by parts, we can write

$$\int_{\mathbb{R}} f'(x)\, e^{-2\pi i x\xi}\, dx = f(x)e^{-2\pi i x\xi}\Big|_{-\infty}^{\infty} + (2\pi i\xi)\int_{\mathbb{R}} f(x)\, e^{-2\pi i x\xi}\, dx$$

where the evaluation at infinity is zero since the function is Schwartz. We will also prove (11) for the $n = 1$ case. We write

$$(\hat{f})'(\xi) = \frac{d}{d\xi}\int_{\mathbb{R}} f(x)\, e^{-2\pi i x\xi}\, dx = \int_{\mathbb{R}} \frac{d}{d\xi} f(x)\, e^{-2\pi i x\xi}\, dx = \int_{\mathbb{R}} (-2\pi i x) f(x)\, e^{-2\pi i x\xi}\, dx$$

where bringing the differential operator inside of the integral is justified by the fast convergence of f. For the final part of the proposition, we have the bound

$$\begin{aligned}
\|x^m (\hat{f})^{(n)}(x)\|_\infty &= \|x^m\left((-2\pi i x)^n f(x)\right)\hat{}\,\|_\infty \\
&= (2\pi)^n \left\| \frac{(2\pi)^m}{(2\pi)^m} x^m \left(x^n f(x)\right)\hat{}\, \right\|_\infty \\
&= (2\pi)^{n-m} \left\| \left[\frac{d^m}{dx^m}\left(x^n f(x)\right)\right]\hat{}\, \right\|_\infty \\
&\le (2\pi)^{n-m} \left\| \frac{d^m}{dx^m}\left(x^n f(x)\right) \right\|_1 \qquad \text{by (1)}
\end{aligned}$$

where the bottom term is finite since $\frac{d^m}{dx^m}\left(x^n f(x)\right) \in \mathcal{S}(\mathbb{R})$. □

Example 3.7. Let f be the Schwartz function $f(x) = e^{2\pi i x t} e^{-\pi(\beta x)^2}$. Then

$$\hat{f}(\xi) = \frac{1}{\beta}\, e^{-\pi\left(\frac{\xi - t}{\beta}\right)^2},$$

which follows from Example 3.4 and parts (7) and (8) of the proposition. This function will be very useful when proving the Fourier inversion formula.

In the above proposition, we have shown that the Fourier transform of a Schwartz function is itself Schwartz, and that differentiating the Fourier transform of a function results in a multiplication of the Fourier transform by $2\pi i\xi$. The fact that differentiation of the Fourier transform results in multiplication, which follows from a simple integration by parts, will be the key to proving the classical uncertainty principle. We need Fourier inversion before we can prove the uncertainty principle, but we are now in a position to prove the first of our uncertainty results, which states that f and $\hat{f}$ cannot both be compactly supported. We begin with some results about the extension of $\hat{f}$ to a holomorphic function on $\mathbb{C}$.

Lemma 3.8. *(Uniqueness of the Fourier Transform) If $f, g \in L^1(\mathbb{R})$ and $\hat{f}(x) = \hat{g}(x)$ for all $x \in \mathbb{R}$, then $f \equiv 0$ a.e. on $\mathbb{R}$.*

Proof. First recall that $\hat{f}$ is continuous if $f \in L^1(\mathbb{R})$, so in proving this theorem we suppose $\hat{f} = \hat{g}$ for *all* $x \in \mathbb{R}$ rather than just a.e. By the linearity of the Fourier transform, the theorem is equivalent to proving that $f = 0$ a.e. if $\hat{f}(x) = 0$ for all $x \in \mathbb{R}$.

We will prove the theorem by constructing a function H such that $H(0) = \int_{-\infty}^{0} f(x)\,dx = 0$. This is enough to show that $f = 0$ a.e., because we can define a new function $g(x) = f(x-a)$ with $\hat{g}(x) = e^{-2\pi i x a}\,\hat{f}(x) = 0$ for all $x \in \mathbb{R}$. Then we apply the same process to g to obtain $\int_{-\infty}^{0} g(x)\,dx = \int_{-\infty}^{-a} f(x)\,dx = 0$ for any $a \in \mathbb{R}$, which shows that $\int_A f(x)\,dx = 0$ for any measurable set A, so $f = 0$ a.e.

To construct H, we first break $\hat{f}$ apart into two different pieces, $\hat{f} = F_- + F_+$, where

$$F_-(\xi) = \int_{-\infty}^{0} f(t)\,e^{-2\pi i \xi t}\,dt \text{ and } F_+(\xi) = \int_{0}^{\infty} f(t)\,e^{-2\pi i \xi t}\,dt\,. \tag{3.9}$$

Now take $z = x + iy$ from the upper half of the plane (so $\text{Im}(z) = y \geq 0$) and observe that

$$|F_-(z)| = \left| \int_{-\infty}^{0} f(t) e^{-2\pi i x t}\, e^{2\pi y t}\,dt \right| \leq \int_{-\infty}^{0} |f(t)|\,dt \leq \|f\|_1\,.$$

Therefore F_- is uniformly bounded on $\{z \in \mathbb{C} \colon \text{Im}(z) \geq 0\}$, and applying the same proof as before shows that the function is uniformly continuous on this set. Therefore, we can differentiate under the integral sign, and since the interior is holomorphic as a function of ξ, we know F_- is holomorphic on the upper half-plane. A similar argument shows that F_+ is holomorphic on the lower half-plane.

We will now use F_- and F_+ to construct an entire holomorphic function H. By assumption, $F_-(x) + F_+(x) = \hat{f}(x) = 0$ if x is real, which means $F_+ = -F_-$ on the real line. Define H by

$$H(z) := \begin{cases} F_-(z) & \text{if } \text{Im}(z) \geq 0 \\ F_+(z) & \text{if } \text{Im}(z) \leq 0 \end{cases}.$$

H is uniformly continuous and uniformly bounded by $\|f\|_1$, and holomorphic when restricted to the upper half-plane or to the lower-half plane. By Morera's theorem, if $\int_\gamma H(z)\,dz = 0$ for any simple closed C^1 curve γ, then H is entire holomorphic, because we can construct an antiderivative to H by integration along curves independent of the curves themselves. If γ is contained entirely in the upper half or lower half of the plane, then $\int_\gamma H(z)\,dz = 0$ by Cauchy's theorem since H is holomorphic in these regions. For a curve that crosses the real axis, split the curve into its lower plane and upper plane parts and complete the new curves along the real line. The integral around the new curves will still be zero, and integrating along the real line in opposite directions will cancel out. Therefore $\int_\gamma H(z)\,dz = 0$ for any simple closed C^1 gamma, so H is entire holomorphic.

H is now an entire bounded holomorphic function, so by Liouville's Theorem H is constant. An analog of the proof of (7) from the previous proposition shows that $\lim_{|z|\to\infty} H(z) = 0$, so we must have $H(z) = 0$ for all $z \in \mathbb{C}$. Thus $H(0) = \int_{-\infty}^{0} f(x)\,dx = 0$ and we are done. □

Remark 3.10. Another version of the above lemma can be obtained using Fourier inversion, but the version proved here is stronger because it shows that if the Fourier transforms of two functions are equal, then the original functions are equal a.e. *even if the Fourier transform of the functions is not invertible.* While the Fourier transform of a Schwartz function is always invertible, this is not the case for L^1.

Lemma 3.11. *(Paley-Wiener) Let $f \in C_0^\infty(\mathbb{R})$, and suppose $f(x) = 0$ if $|x| > R$. Then $\hat{f}$ can be extended to a holomorphic function on all of $\mathbb{C}$, and we have the decay estimate*

$$|\hat{f}(z)| \le C_n(1+|z|)^{-n} e^{2\pi|Im(z)|R} \tag{3.12}$$

for any $n \in \mathbb{Z}^+$.

Proof. First observe that $\hat{f}(z)$ is absolutely convergent for all $z \in \mathbb{C}$, since

$$|\hat{f}(z)| = \left|\int_{\mathbb{R}} f(x)\, e^{-2\pi i z x}\, dx\right| = \left|\int_{-R}^{R} f(x)\, e^{-2\pi i \mathrm{Re}(z) x} e^{2\pi \mathrm{Im}(z) x}\, dx\right| \le e^{2\pi|\mathrm{Im}(z)|R}\, \|f\|_1 .$$

Therefore we can differentiate under the integral sign, and $\hat{f}$ is holomorphic on $\mathbb{C}$ since the interior of the integral is holomorphic on $\mathbb{C}$. To prove the bound, the fact that $\hat{f}$ is holomorphic allows us to restrict our attention to $|z| > 1$, since $\hat{f}$ is uniformly bounded on the closed unit disk. Using integration by parts yields

$$\hat{f}(z) = \int_{-R}^{R} \frac{1}{(-2\pi i z)^n} \left(\frac{d^n}{dx^n}\, e^{-2\pi i z x}\right) f(x) dx = \frac{1}{(-2\pi i z)^n} \int_{-R}^{R} e^{-2\pi i z x}\, \frac{d^n f}{dx^n}(x)\, dx\, .$$

Since f is Schwartz, there is some $K_n > 0$ such that $|f^{(n)}(x)| < K_n$ for all $x \in \mathbb{R}$, so we obtain

$$|\hat{f}(z)| \le \left|\frac{1}{(-2\pi i z)^n} \int_{-R}^{R} K_n\, e^{-2\pi i z x}\, dx\right| \le \frac{2RK_n}{(2\pi)^n}\, z^{-n}\, e^{2\pi|\mathrm{Im}(z)|\, R}$$

which concludes the proof. □

Theorem 3.13. *Let $f \in C_0^\infty(\mathbb{R})$. If $\hat{f}$ has compact support, then $f \equiv 0$.*

Proof. We will use a proof by contradiction. Suppose f is a non-zero infinitely differentiable function of compact support and that the support of $\hat{f}$ is also compact. Since f and $\hat{f}$ have compact support, there is some $R > 0$ such that $f(x) = \hat{f}(x) = 0$ if $|x| > R$. By the previous lemma, $\hat{f}$ can be extended to an entire holomorphic function. For a non-zero holomorphic function g, if $g(z_0) = 0$ then there exists a $\delta > 0$ such that $g(z) \neq 0$ for any $z \in \mathrm{B}(z_0, \delta)$. Since $\hat{f}(2R) = 0$ and for any $\delta > 0$, we have $2R + \delta/2 \in \mathrm{B}(2R,\, \delta)$ and $\hat{f}(2R + \delta/2) = 0$, we must have $\hat{f} = 0$ on all of $\mathbb{C}$, and therefore also on all of $\mathbb{R}$. By the Uniqueness of the Fourier Transform, this implies that $f = 0$ a.e. on $\mathbb{R}$, and since $f \in C_0^\infty$, $f = 0$ everywhere on $\mathbb{R}$. □

4. Fourier Inversion

One of the most useful properties of the Fourier transform is the fact that under certain circumstances, the Fourier transform of a function can be inverted to recover the original function. This means that when working with the Fourier transform of a function, we don't have to worry about losing information contained in the original function, since this information can be recovered using the inversion process. An easier corollary of Fourier inversion is the Plancherel identity, which states that the L^2 norm of a function and of its

Fourier transform are equal, and this identity will appear very frequently when proving the uncertainty principle and other uncertainty results.

Some quick informal scratch work sheds light on why the process of Fourier inversion works in certain cases, and also gives an explanation of why the Fourier transform of a function on $\mathbb{R}$ is defined as a function on $\mathbb{R}$, whereas the Fourier series for a function on a bounded interval is only a collection of coefficients defined on $\mathbb{Z}$.

Consider some function $f : \mathbb{R} \to \mathbb{C}$ with period 1. The n^{th} Fourier coefficient of f for some integer n is given by

$$\hat{f}(n) = \int_{-1/2}^{1/2} f(t)\, e^{-2\pi int}\, dt\,,$$

and if f is a well-behaved function then we are able to reconstruct f from its Fourier coefficients and write

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

Now, consider some function $f : \mathbb{R} \to \mathbb{C}$ with arbitrary period $P > 0$. Then $g(x) = f(Px)$ is a function of period 1, so from a change of variables we could choose to define the Fourier coefficients of f by

$$\hat{f}(n) = \frac{1}{P} \int_{-P/2}^{P/2} f(t)\, e^{-\frac{2\pi int}{P}}\, dt\,.$$

Reconstructing f from these coefficients in the same way as before, we get

$$f(t) = \sum_{n=-\infty}^{\infty} \hat{f}(n)\, e^{\frac{2\pi int}{P}} = \frac{1}{P} \sum_{n=-\infty}^{\infty} \left(\int_{-P/2}^{P/2} f(x)\, e^{-\frac{2\pi inx}{P}}\, dx \right) e^{\frac{2\pi int}{P}}\,.$$

Taking the limit as $P \to \infty$ and ignoring problems of convergence, the sum becomes an integral and we might hope to write

$$f(t) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x) e^{-2\pi ix\xi}\, dx \right) e^{2\pi it\xi}\, d\xi \tag{4.1}$$

for a function of period ∞, which is just a function on the line. This is exactly what we will prove in the upcoming Fourier inversion theorem, and the above equation turns out to always be true when f is Schwartz. Motivated by this analysis, we give a definition for the inverse of the Fourier transform.

Definition 4.2. Given $f \in \mathcal{S}(\mathbb{R})$, we define

$$\check{f}(x) = \int_{\mathbb{R}} f(t)\, e^{2\pi itx}\, dt = \hat{f}(-x) \tag{4.3}$$

for all $x \in \mathbb{R}$. The operation $f \to \check{f}$ is called the *inverse Fourier transform.*

Before proving the Fourier inversion theorem, we first need some results about *approximate identities.* In physics, one often uses the Dirac-delta distribution $\delta(x)$, which satisfies

$$\int_{\mathbb{R}} f(t)\delta(x-t)\, dt = f(x) \tag{4.4}$$

for all f in some set of functions. In general, the integral $\int_{\mathbb{R}} f(t)g(x-t)\, dt$ is known as the *convolution* of f and g, which itself is a function from $\mathbb{R}$ to $\mathbb{C}$, and we might wonder if there is some function g such that the convolution of f and g returns the orignal function

f. Although such a function does not exist in the conventional sense of functions, often we *can* find a family of functions $g(x, \varepsilon)$ with the property that

$$\lim_{\varepsilon \to 0} \int_{\mathbb{R}} f(x)\, g(x-t,\, \varepsilon)\, dt = f(x)\,. \tag{4.5}$$

Such functions are known as approximate identities.

Definition 4.6. An *approximate identity* (as $\varepsilon \to 0$) is a family of functions $k_\varepsilon \in L^1(\mathbb{R})$ with the following properties:

(1) There exists a constant $c > 0$ such that $\|k_\varepsilon\|_1 \leq c$ for all $\varepsilon > 0$.
(2) $\int_{\mathbb{R}} k_\varepsilon(x)\, dx = 1$ for all $\varepsilon > 0$.
(3) For any $\delta > 0$, we have $\int_{|x| \geq \delta} |k_\varepsilon(x)|\, dx \to 0$ as $\varepsilon \to 0$.

From this definition, we get an idea of how an approximate identity works. As $\varepsilon \to 0$, the mass of k_ε becomes concentrated closer and closer to the origin. If k_ε replaces g in (4.3) and x is fixed, then the product inside of the integral vanishes for all t except those close to x when ε is small, and since integrating k_ε over $\mathbb{R}$ always gives 1, the value of the integral is close to $f(x)$ for small ε. Before we prove that (4.3) does indeed hold for families of functions that satisfy the three properties listed above, we give some examples of approximate identities on $\mathbb{R}$.

Example 4.7. The following are examples of approximate identities.

- Let k be a function such that $\int_{\mathbb{R}} k(x)\, dx = 1$. Then $k_\varepsilon = \varepsilon^{-1}\, k(\varepsilon^{-1} x)$ is an approximate identity. A simple substitution of variables shows that k_ε satisfies properties (1) and (2) since k satisfies both of these properties. Property (3) is satisfied since
$$\lim_{\varepsilon \to 0} \int_{|x| \geq \delta/\varepsilon} |k(x)|\, dx = 0$$
for all $\delta > 0$. This simple trick shows that approximate identites are quite common, and that many different types of functions can be altered slightly to serve as an approximate identity. This method of converting a function into an approximate identity bears some similarity to (8) from Proposition 3.5, and we will exploit this when proving Fourier inversion.
- Let $P(x) = (\pi(1+x^2))^{-1}$, and let $P_\varepsilon(x) = \varepsilon^{-1} P(\varepsilon^{-1} x)$. Then by the above argument P_ε is an approximate identity, since
$$\int_{\mathbb{R}} \frac{1}{\pi}\, \frac{1}{1+x^2}\, dx = \frac{1}{\pi} \arctan(x) \bigg|_{-\infty}^{\infty} = 1\,.$$
P_ε is commonly known as the Poisson kernel.
- Let $g(x) = e^{-\pi x^2}$. Then $g_\varepsilon(x) = \varepsilon^{-1}\, g(\varepsilon^{-1} x)$ is an approximate identity, since $\int_{\mathbb{R}} g(x)\, dx = 1$. To see this, let $A = \int_{\mathbb{R}} e^{-\pi x^2}\, dx$. Then
$$\begin{aligned} A^2 &= \left(\int_{\mathbb{R}} e^{-\pi x^2}\, dx\right)\left(\int_{\mathbb{R}} e^{-\pi y^2}\, dy\right) = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{-\pi(x^2+y^2)}\, dy\, dx \\ &= \int_0^{2\pi} \int_0^{\infty} r\, e^{-\pi r^2}\, dr\, d\theta = 1 \end{aligned}$$
where the third equality follows from a change to polar coordinates, and the last integral is easily evaluted using substitution. We have already used the fact that $\int_{\mathbb{R}} g(x)\, dx = 1$ earlier in Example 3.4. We will use the approximate identity g_ε in our proof of the Fourier inversion theorem.

Theorem 4.8. *Let k_ε be an approximate identity, and let $f \in \mathcal{S}(\mathbb{R})$. Then*

$$\lim_{\varepsilon \to 0} \left\| \int_{\mathbb{R}} f(t)\, k_\varepsilon(x-t)\, dt - f(x) \right\|_\infty = 0\,, \tag{4.9}$$

so the convolution of f and k_ε converges to f for all $x \in \mathbb{R}$ as $\varepsilon \to 0$.

Proof. First we observe that we can write

$$\int_{\mathbb{R}} f(t)\, k_\varepsilon(x-t)\, dt = \int_{\mathbb{R}} f(x-t)\, k_\varepsilon(t)\, dt$$

by using a simple substitution, so the convolution operation is commutative and we can instead prove the theorem for $\int_{\mathbb{R}} f(x-t)\, k_\varepsilon(t)\, dt$.

Next observe that if $f \in \mathcal{S}(\mathbb{R})$, then f is uniformly continuous on $\mathbb{R}$, which is shown as follows: since $\sup_{x\in\mathbb{R}} |xf(x)| < \infty$, we must have $\lim_{|x|\to\infty} f(x) = 0$, and we know f is continuous by definition. Therefore there exists an $R > 0$ such that if $|x| > R$, then $|f(x)| < \varepsilon/2$. Any continuous function is uniformly continuous on a compact set, so there exists some $\delta_0 > 0$ such that if $x,\, y \in [-R-1, R+1]$ and $|x-y| < \delta$, then $|f(x)-f(y)| < \varepsilon$. Therefore, for any $x,\, y \in \mathbb{R}$, if $|x-y| < \min(\delta, 1)$, we have $|f(x) - f(y)| < \varepsilon$, so f is uniformly continuous.

Now choose some arbitrary $\lambda > 0$. Since f is uniformly continuous, there is some $\delta > 0$ such that if $|h| < \delta$, then

$$|f(x-h) - f(x)| < \frac{\lambda}{2c}$$

for all $x \in \mathbb{R}$, where c is the constant from part (1) of Definition 4.4. Now by part (2) of Definition 4.4, we know $1 = \int_{\mathbb{R}} k_\varepsilon(x)\, dx = \int_{\mathbb{R}} k_\varepsilon(x-t)\, dx$, so we can write

$$\begin{aligned}
\int_{\mathbb{R}} f(x-t)\, k_\varepsilon(t)\, dt - f(x) &= \int_{\mathbb{R}} f(x-t)\, k_\varepsilon(t)\, dt - f(x) \int_{\mathbb{R}} k_\varepsilon(t)\, dt \\
&= \int_{\mathbb{R}} (f(x-t) - f(x))\, k_\varepsilon(t)\, dt \\
&= \int_{|t|<\delta} (f(x-t) - f(x))\, k_\varepsilon(t)\, dt + \int_{|t|\ge\delta} (f(x-t) - f(x))\, k_\varepsilon(t)\, dt\,.
\end{aligned}$$

Taking the L^∞ norm of the first term in the bottom sum gives

$$\begin{aligned}
\sup_{x\in\mathbb{R}} \left| \int_{|t|<\delta} (f(x-t) - f(x))\, k_\varepsilon(t)\, dt \right| &\le \int_{|t|<\delta} \sup_{x\in\mathbb{R}} |f(x-t) - f(x)|\, |k_\varepsilon(t)|\, dt \\
&< \int_{|t|<\delta} \frac{\lambda}{2c} |k_\varepsilon(t)|\, dt \le \frac{\lambda}{2}
\end{aligned}$$

and taking the L^∞ norm of the second term gives

$$\begin{aligned}
\sup_{x\in\mathbb{R}} \left| \int_{|t|\ge\delta} (f(x-t) - f(x))\, k_\varepsilon(t)\, dt \right| &\le \int_{|t|\ge\delta} \sup_{x\in\mathbb{R}} |f(x-t) - f(x)|\, |k_\varepsilon(t)|\, dt \\
&\le 2\, \|f\|_\infty \int_{|t|\ge\delta} |k_\varepsilon(t)|\, dt\,.
\end{aligned}$$

Now, by property (3) of Definition 4.4, there exists an $\varepsilon_0 > 0$ such that if $\varepsilon < \varepsilon_0$, then

$$\int_{|t|\ge\delta} |k_\varepsilon(t)|\, dt < \frac{\lambda}{4\|f\|_\infty}\,.$$

Therefore $\| \int_{\mathbb{R}} f(x-t)\, k_\varepsilon(t)\, dt - f(x)\|_\infty < \lambda$, so the theorem is proven. ☐

With the above theorem in hand, we are ready to prove Fourier inversion.

Theorem 4.10. *Let f, g, $h \in \mathcal{S}(\mathbb{R})$. Then the following hold:*

(1)
$$\int_{\mathbb{R}} f(x)\hat{g}(x)\, dx = \int_{\mathbb{R}} \hat{f}(x) g(x)\, dx$$

(2) (Fourier Inversion)
$$(\hat{f})\check{} = f = (\check{f})\,\hat{}$$

(3)
$$\int_{\mathbb{R}} f(x)\overline{h(x)}\, dx = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{h}(\xi)}\, d\xi$$

(4) (Plancherel's Identity)
$$\|f\|_2 = \|\hat{f}\|_2 = \|\check{f}\|_2$$

(5)
$$\int_{\mathbb{R}} f(x)g(x)\, dx = \int_{\mathbb{R}} \hat{f}(x)\check{g}(x)\, dx$$

Proof. For (1), we have
$$\int_{\mathbb{R}} f(x) \left(\int_{\mathbb{R}} g(y)\, e^{-2\pi i y x}\, dy \right) dx = \int_{\mathbb{R}} g(y) \left(\int_{\mathbb{R}} f(x)\, e^{-2\pi i y x}\, dx \right) dy\,,$$
where the change in the order of integration is justified by Fubini's Theorem, which is valid because of the absolute convergence of the integrals. To prove (2), we use (1) with
$$g(x) = e^{2\pi i x t}\, e^{-\pi(\varepsilon x)^2}\,.$$
Fortunately, we have already done enough work to find $\hat{g}$ without an explicit computation. Combining the results of Example 3.4 and Proposition 3.5 parts (7) and (8) yields
$$\hat{g}(x) = \frac{1}{\varepsilon}\, e^{-\pi\left(\frac{x-t}{\varepsilon}\right)^2}\,,$$
which is an approximate identity by Example 4.7 and the fact that $\int_{\mathbb{R}} e^{-\pi(x-t)^2}\, dx = \int_{\mathbb{R}} e^{-\pi x^2}\, dx = 1$ for any real number t. Applying (1) now yields
$$\int_{\mathbb{R}} f(x)\varepsilon^{-1} e^{-\pi\left(\frac{x-t}{\varepsilon}\right)^2}\, dx = \int_{\mathbb{R}} \hat{f}(x)\, e^{2\pi i x t}\, e^{-\pi(\varepsilon x)^2}\, dx\,.$$
We now consider the above equation as $\varepsilon \to 0$. By Theorem 4.8, the left-hand side converges to $f(t)$. Since $e^{-\pi(\varepsilon x)^2} \leq 1$, the right-hand side converges to $(\hat{f})\check{}(t)$ by the Lebesgue Dominated Convergence Theorem. Therefore $f = (\hat{f})\check{}$ on $\mathbb{R}$, and the other equality in (2) follows in a similar way.

To prove (3), observe that
$$(\bar{\hat{g}})\widehat{\ } = (\hat{g})^{\bar{\check{}}} = \bar{g}$$
where the first equality follows from Proposition 3.5 (5). Therefore (3) follows from (1) with $g = \bar{\hat{h}}$. Plancherel's identity follows from (3) with $h = f$, and (5) follows from (1) and Fourier inversion. ☐

We remark that Plancherel's identity allows us to extend $\mathcal{F}$ uniquely to an isometry on $L^2(\mathbb{R}$. The following corollary sums up why the Schwartz class is a natural environment for studying the Fourier transform.

Corollary 4.11. *The Fourier transform is a homeomorphism from $\mathcal{S}(\mathbb{R})$ onto itself.*

Proof. The fact that $\mathcal{F}$ is bijective follows from Proposition 3.5 (12) and Fourier inversion. To prove continuity, we show that if $f_k \to f$ in $\mathcal{S}(\mathbb{R})$, then $\hat{f}_k \to \hat{f}$ in $\mathcal{S}(\mathbb{R})$. For any $g \in \mathcal{S}(\mathbb{R})$, we have

$$\begin{aligned} \left\| x^m \hat{g}^{(n)}(x) \right\|_\infty &= (2\pi)^{n-m} \left\| \mathcal{F}\left[\frac{d^m}{dx^m}(x^n g(x)) \right] \right\|_\infty \\ &\leq (2\pi)^{n-m} \left\| \frac{d^m}{dx^m}(x^n g(x)) \right\|_1 \end{aligned}$$

just as in the proof of Proposition 3.5 (12). Now if we let $g = f_k - f$, by using the product rule and the fact that convergence in $\mathcal{S}(\mathbb{R})$ implies convergence in $L^1(\mathbb{R})$, it is simple to show that $\rho_{m,n}(\hat{f}_k - \hat{f}) \to 0$ as $k \to \infty$ for arbitrary $m, n \in \mathbb{Z}^+$. Therefore $\hat{f}_k \to f$ in $\mathcal{S}(\mathbb{R})$ and $\mathcal{F}$ is continuous. □

5. The Uncertainty Principle

Now that we have established all the necessary preliminary results, we obtain the uncertainty principle from what is essentially a simple integration by parts. After the proof of the uncertainty principle, we give some physical interpretations of our new formula.

Theorem 5.1. *For any $f \in \mathcal{S}(\mathbb{R})$ and any $x_0, \xi_0 \in \mathbb{R}$, we have the following inequality:*

$$\|f(x)\|_2^2 \leq 4\pi \|(x - x_0)f(x)\|_2 \, \|(\xi - \xi_0)\hat{f}(\xi)\|_2 \,. \tag{5.2}$$

Moreover, we have equality in the above expression if and only if f is a modulated and shifted Gaussian, which means it has the form $f(x) = c_0\, e^{ic_1(x-\xi_0)}\, e^{-c_2(x-x_0)^2}$ with $c_2 > 0$.

Proof. We will first prove a weaker version of the inequality before we prove the theorem in its full generality. We write

$$\begin{aligned} \int_{\mathbb{R}} |f(x)|^2\, dx &= \int_{\mathbb{R}} f(x)\overline{f(x)}\, dx \\ &= x|f(x)|^2 \,\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}} x\,(f'(x)\overline{f(x)} + f(x)\overline{f'(x)})\, dx \\ &= -2\int_{\mathbb{R}} x\, \mathrm{Re}('(x)\overline{f'(x)})\, dx \end{aligned}$$

where the second equality follows from integration by parts and the last equality holds because $\lim_{|x|\to\infty} |xf(x)^2| = 0$ since f is Schwartz. Now, we can write

$$\|f\|_2^2 = \left| -2\int_{\mathbb{R}} x\, \mathrm{Re}(f(x)\overline{f'(x)})\, dx \right| \leq 2\, \|xf(x)\|_2\, \|f'(x)\|_2 \tag{5.3}$$

which follows from the Cauchy-Schwartz inequality. Recalling that differentiation of the original function translates to multiplication by $2\pi i\xi$ in the domain of the Fourier transform, we have

$$\|\overline{f'(x)}\|_2 = \|f'(x)\|_2 = \|\widehat{(f')}(\xi)\|_2 = \|(2\pi i\xi)\hat{f}(\xi)\|_2 = 2\pi\|\xi\, \hat{f}(\xi)\|_2$$

where the second equality follows from Plancherel's Identity and the third equality follows from Proposition 3.5 (10). Combining this with our previous results gives

$$\|f\|_2^2 \le 4\pi \|x\, f(x)\|_2 \, \|\xi\, \hat{f}(\xi)\|_2 \,. \tag{5.4}$$

To prove the theorem in its full generality, we define a new function

$$g(x) = e^{-2\pi i x \xi_0}\, f(x + x_0)\,, \quad \text{so that} \quad \hat{g}(\xi) = e^{2\pi i x_0 \xi}\, \hat{f}(\xi + \xi_0)\,.$$

For this function g, we have

$$\|g\|_2^2 = \int_{\mathbb{R}} \left| e^{-2\pi i x \xi_0}\, f(x + x_0)\right|^2 \, dx = \int_{\mathbb{R}} |f(t)|^2 \, dt = \|f\|_2^2\,,$$

$$\begin{aligned} \|u\, g(u)\|_2 &= \left(\int_{\mathbb{R}} \left| u\, e^{-2\pi i u \xi_0}\, f(u + x_0)\right|^2 \, du \right)^{1/2} \\ &= \left(\int_{\mathbb{R}} |(x - x_0)\, f(x)|^2 \, dx \right)^{1/2} = \|(x - x_0) f(x)\|_2\,, \end{aligned}$$

$$\begin{aligned} \|t\, \hat{g}(t)\|_2 &= \left(\int_{\mathbb{R}} \left| t\, e^{2\pi i x_0 t}\, \hat{f}(t + \xi_0)\right|^2 \, dt \right)^{1/2} \\ &= \left(\int_{\mathbb{R}} \left| (\xi - \xi_0)\, \hat{f}(\xi)\right|^2 \, d\xi \right)^{1/2} = \|(\xi - \xi_0)\, \hat{f}(\xi)\|_2\,. \end{aligned}$$

Therefore, applying (5.4) to g proves (5.2).

If we want equality to hold in (5.2), then we must have equality when we apply Cauchy-Schwartz in (5.3). Equality holds when applying Cauchy-Schwartz to the inner product of u and v if and only if $v = \lambda u$ for some scalar λ. Therefore, for equality to hold in (5.2) we must have $f'(x) = \lambda\, x\, f(x)$, a differential equation which has the solution $f(x) = C\, e^{\lambda/2\, x^2}$. □

If we consider f and $\hat{f}$ as probability distributions, then the expressions

$$\inf_{x_0 \in \mathbb{R}} \|(x - x_0) f(x)\|_2 \quad \text{and} \quad \inf_{\xi_0 \in \mathbb{R}} \|(\xi - \xi_0)\, \hat{f}(\xi)\|_2$$

correspond to the standard deviation of f and the standard deviation of $\hat{f}$ respectively. A probabilistic interpretation of the uncertainty principle says that the standard deviation of f and $\hat{f}$ exhibit an inversely proportional relationship, so if f is tightly concentrated in a small area then $\hat{f}$ has a much wider spread, and vice-versa. The smaller the standard deviation of a probability distribution is, the more precisely one can predict the outcome of the random event, so if two probability distributions are Fourier transforms of one another, we can accurately predict at best one event. Moreover, paired normal distributions give us the most overall predictive power of any distribution, since equality in the theorem only holds for normal distributions.

There are two important physical interpretations of Theorem 5.1, which arise from two sets of paired domains. Taking the Fourier transform of a probability distribution in the position domains gives a probability distribution in the momentum domain, so Theorem 5.1 states that we cannot precisely know both the position and momentum of a particle to arbitrary certainty. This is the well-known Uncertainty Principle from quantum mechanics, which arises from little more than integration by parts and the Cauchy-Schwartz inequality but has huge ramifications for the way we look at the world we live in. Time and frequency are another set of paired domains, so Theorem 5.1 also states that sampling a sound over smaller and smaller time intervals results in a loss in accuracy of the frequency sampled, which has important implications for audio technology.

6. The Amrein-Berthier Theorem

We conclude our study of the uncertainty principle with the Amrein-Berthier theorem. As an easy corollary, we obtain the result that a non-zero function f supported on a set of finite measure cannot have a Fourier transform which is supported on a set of finite measure. We have already shown that this is the case if f is compactly supported, but proving the result for sets of finite measure (which might not be bounded) is more difficult and will require methods significantly different from those already used.

Suppose E, $F \subset \mathbb{R}$ have finite measure. We are interested in determining whether there is a non-zero $f \in L^2(\mathbb{R})$ such that

$$\operatorname{supp}(f) \subset E \quad \text{and} \quad \operatorname{supp}(f) \subset F\,. \tag{6.1}$$

We begin by defining a linear operator $T[f] = \chi_E(\chi_F \hat{f})\check{\ }$, where T is considered as an operator from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$. If there is a function f that satisfies (6.1), then it is clear that $T(f) = f$, which would imply that $\|T\|_{\text{op}} \geq 1$. On the other hand, we have

$$\|T[f]\|_2 = \|\chi_E(\chi_F\hat{f})\check{\ }\|_2 \leq \|(\chi_F\hat{f})\check{\ }\|_2 = \|\chi_F\hat{f}\|_2 \leq \|\hat{f}\|_2 = \|f\|_2$$

where the equalities all follow from Plancherel's Identity, so $\|T\|_{\text{op}} \leq 1$. Therefore, if we can show that $\|T\|_{\text{op}} < 1$, then there is no non-zero f that satisfies (6.1). We can rewrite $T[f]$ in the following way:

$$\begin{aligned} T[f](x) = (\chi_E(\chi_F\hat{f})\check{\ })(x) &= \chi_E(x)\int_{\mathbb{R}} e^{2\pi i t x}\,\chi_F(t)\left(\int_{\mathbb{R}} f(y)\,e^{-2\pi i y t}\,dy\right)dt \\ &= \chi_E(x)\int_{\mathbb{R}} f(y)\int_{\mathbb{R}}\chi_F(t)\,e^{2\pi i t(x-y)}\,dt\,dy \end{aligned}$$

so we have

$$T[f](x) = \int_{\mathbb{R}} \chi_E(x)\,\check{\chi}_F(x-y) f(y)\,dy\,. \tag{6.2}$$

Therefore T is an integral transform with kernel $K(x, y) = \chi_E(x)\,\check{\chi}_F(x-y)$, and

$$\|T\|_{\text{op}} \leq \left(\int_{\mathbb{R}}\int_{\mathbb{R}} |K(x, y)|^2 dx\,dy\right)^{1/2} = (|E|\cdot|F|)^{1/2} := \sigma < \infty\,.$$

Therefore $\|T\|_{\text{op}} \leq \min(\sigma, 1)$, and if $\sigma < 1$ it immediately follows that the only f that satisfies (6.1) is the zero function. With a little more work, we can show that the previous statement is true for arbitrary σ. Before doing so, we temporarily drop the requirement that E and F have finite measure and show that $\|T\|_{\text{op}} < 1$ is equivalent to several other statements, some of which are more immediately applicable.

Lemma 6.3. *Let E, F be measurable subsets of $\mathbb{R}$, let $f \in L^2(\mathbb{R})$, and let C_1, C_2 be positive constants. Then the following are equivalent:*

(1) $\|f\|_{L^2(\mathbb{R})} \leq C_1(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)})$

(2) There exists some $\varepsilon > 0$ such that $\|f\|^2_{L^2(E)} + \|\hat{f}\|^2_{L^2(F)} \leq (2-\varepsilon)\|f\|_2^2$

(3) If $\operatorname{supp}(\hat{f}) \subset F$, *then* $\|f\|_2 \leq C_2\|f\|_{L^2(E^c)}$

(4) *If* $\operatorname{supp}(f) \subset E$*, then* $\|\hat{f}\|_2 \le C_2 \|\hat{f}\|_{L^2(E^c)}$
(5) *There exists* $0 < \rho < 1$ *such that* $\|\chi_E(\chi_F \hat{f})\check{}\|_2 \le \rho \|f\|_2$

Remark 6.4. $\|T\|_{\text{op}} < 1$ is the same as statement (5).

Proof. Plancherel's Identity identity is used frequently throughout the proof, and we will no longer make note when it is used.

(1) ⇒ (2): We can write

$$\begin{aligned} \|f\|^2_{L^2(E)} + \|\hat{f}\|^2_{L^2(F)} &= 2\|f\|_2^2 - \|f\|^2_{L^2(E^c)} - \|\hat{f}\|^2_{L^2(F^c)} \\ &\le (2 - (2C_1)^{-1})\|f\|_2^2 \,. \end{aligned}$$

(2) ⇒ (3): If $\operatorname{supp}(\hat{f}) \subset F$, then $\|\hat{f}\|^2_{L^2(F)} = \|\hat{f}\|^2_{L^2(\mathbb{R})}$, so $0 \le \|f\|^2_{L^2(E)} \le (1-\varepsilon)\|f\|_2^2$. Now

$$\varepsilon \|f\|_2^2 \le \|f\|^2_{L^2(E)} \le \|f\|_2^2 \,, \qquad \text{or} \qquad \|f\|_2^2 \le \varepsilon^{-1} \|f\|^2_{L^2(E)} \,.$$

(3) ⇒ (1): Observe that $\operatorname{supp}(\chi_F \hat{f}) \subset F$ by definition, so

$$\begin{aligned} \|f\|_2 &\le \|(\chi_F \hat{f})\check{}\|_2 + \|(\chi_{F^c} \hat{f})\check{}\|_2 \\ &\le C_2 \|\chi_{E^c}(\chi_F \hat{f})\check{}\|_2 + \|(\chi_{F^c}\hat{f})\check{}\|_2 \\ &\le C_2 \|\chi_{E^c} f\|_2 + C_2 \|\chi_{E^c}(\chi_{F^c}\hat{f})\check{}\|_2 + \|(\chi_{F^c}\hat{f})\check{}\|_2 \\ &\le C_2 \|\chi_{E^c} f\|_2 + (C_2 + 1)\|(\chi_{F^c}\hat{f})\check{}\|_2 \,. \end{aligned}$$

Simply replacing f by $\hat{f}$ gives (2) ⇒ (4) and (4) ⇒ (1).

(3) ⇒ (5): As before, $\operatorname{supp}(\chi_F \hat{f}) \subset F$, so

$$\begin{aligned} \|(\chi_F \hat{f})\check{}\|_2^2 &= \|\chi_E(\chi_F \hat{f})\check{}\|_2^2 + \|\chi_{E^c}(\chi_F \hat{f})\check{}\|_2^2 \\ &\ge \|\chi_E(\chi_F \hat{f})\check{}\|_2^2 + C_2^{-2}\|(\chi_F \hat{f})\check{}\|_2^2 \,, \end{aligned}$$

which means $\|\chi_E(\chi_F \hat{f})\check{}\|_2 \le (1 - C_2^{-2})^{1/2} \|(\chi_F \hat{f})\check{}\|_2 \le (1 - C_2^{-2})^{1/2}\|f\|_2$.

(5) ⇒ (3): If $\operatorname{supp}(\hat{f}) \subset F$, then $(\chi_F \hat{f})\check{} = f$, so we have

$$\begin{aligned} \|f\|_2 &\le \|\chi_E f\|_2 + \|\chi_{E^c} f\|_2 \\ &\le \rho\|f\|_2 + \|\chi_{E^c} f\|_2 \,, \end{aligned}$$

which implies $\|f\|_2 \le (1-\rho)^{-1} \|\chi_{E^c} f\|_2$.

□

We are ready to prove our final theorem.

Theorem 6.5. *(Amrein-Berthier) Let* $f \in L^2(\mathbb{R})$ *and* $E, F \subset \mathbb{R}$ *be sets of finite measure. Then*

$$\|f\|_{L^2(\mathbb{R})} \le C(\|f\|_{L^2(E^c)} + \|\hat{f}\|_{L^2(F^c)}) \tag{6.6}$$

for some constant C *that depends only on* E *and* F*.*

Proof. Recall our linear operator $T[f] = \chi_E(\chi_F \hat{f})\check{}$, and that $\|T\|_{\text{op}} \le \sigma := (|E| \cdot |F|)^{1/2}$. Let

$$A_\lambda := \{f \in L^2(\mathbb{R}) : T[f] = \lambda f\} \tag{6.7}$$

be the left λ-eigenspace of T. First we claim that $\dim(A_\lambda) \le \lambda^{-2}\sigma^2$. To prove this claim, suppose that $\{f_k\}_{k=1}^m$ is an orthonormal sequence sequence in A_λ. Now, with $K(x, y)$ as the kernel of the integral transform T, Bessel's Inequality allows us to write

$$m\lambda^2 = \sum_{k=1}^{m} \left| \int_{\mathbb{R}} \int_{\mathbb{R}} K(x, y) f_k(x) \overline{f_k(y)}\, dx\, dy \right|^2 \le \int_{\mathbb{R}} \int_{\mathbb{R}} |K(x, y)|^2 dx\, dy = \sigma^2 ,$$

because $\int_{\mathbb{R}} K(x, y) f_k(x)\, dx = \lambda f_k(y)$ since $\{f_k\} \subset A_\lambda$ and because

$$\int_{\mathbb{R}} f_j(y) \overline{f_k(y)}\, dy = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \ne k \end{cases}$$

since $\{f_k\}$ is an orthonormal sequence. In particular, $\dim(A_\lambda) < \infty$ for any λ since σ is finite by assumption.

Now we will use proof by contradiction to prove the theorem. Suppose $\|T\|_{\text{op}} = 1$, which contradicts the conclusion of the theorem since the previous lemma shows that the conclusion of the theorem is equivalent to $\|T\|_{\text{op}} < 1$. A bounded integral transform is always a compact operator, so if $\|T\|_{\text{op}} = 1$ then there is some $f \in L^2(\mathbb{R})$ that satisfies (6.1). We now show that $\dim(A_1) = \infty$ to obtain our contradiction. Inductively, we define

$$S_0 := \operatorname{supp}(f)\,, \quad S_1 := S_0 \cup (S_0 - y_0)\,, \quad S_{k+1} := S_k \cup (S_0 - y_k) \quad \text{for } k \ge 1,$$

where the translations $\{y_k\}_{k\in\mathbb{N}}$ are chosen so that $|S_k| < |S_{k+1}| < |S_k| + 2^{-k}$. Using the f that satisfies (6.1), we define $f_k(x) = f(x + y_k)$, so that $\{f_k\}$ is an infinite linearly independent set with

$$\operatorname{supp}(f_k) \subset S_\infty := \bigcup_{j=0}^{\infty} S_j$$

for all $k \ge 0$. Observe that $\operatorname{supp}(\hat{f}_k) \subset F$ for all $k \ge 0$ since translation in the original domain translates to multiplication by a constant in the Fourier domain. Since $|S_\infty| < \infty$, we obtain the desired contradiction with $E = S_\infty$.

□

Acknowledgments. I would like to thank my mentor Casey Rodriguez for his time, knowledge and guidance, and Peter May for organizing the REU program that gave me the opportunity to write this paper.

References

[1] Grafakos, Loukas. Classical and Modern Fourier Analysis. Pearson Education Inc., 2004.
[2] Lang, Serge. Complex Analysis, Fourth Edition. Springer, 2003.
[3] Muscalu, Camil, and Wilhelm Schlag. Classical and Multilinear Harmonic Analysis. Cambridge University Press, 2008.
[4] Vretblad, Anders. Fourier Analysis and Its Applications. Springer-Verlag, New York, 2003.