THE FOURIER UNCERTAINTY PRINCIPLES

KABIR DUBEY

Abstract. Over the past century, Fourier uncertainty principles have been a highly active area of study. Despite their far-reaching implications, the inequalities have simple origins in mathematics. In this paper, we arrange some foundational and powerful perspectives.

Contents

1. Introduction

Uncertainty principles are not formally defined. In general, uncertainty principles refer to a meta-theorem in Fourier analysis that states that a nonzero function and its Fourier transform cannot be localized to arbitrary precision [1]. We will define and prove the fundamental properties of Fourier analysis in sections two and three, commenting on real-world applications along the way.

The localization of a function refers to the behavior of the "size" of a function near and away from zero. Any mathematical statement which multiplies a function that describes localization with its Fourier transform is arguably an uncertainty principle. In section four, we will formalize this overarching theme.

The most popular use of Fourier uncertainty principles is as a description of the natural tradeoff between the stability and measurability of a system, particularly quantum mechanical systems. After introducing the general case in section four, we prove the Heisenberg Uncertainty Principle, as a consequence, in section five. We close the paper by confirming our result with a highly flexible technique based on a "primary uncertainty principle" that yields most, if not all, uncertainty principles in the literature.

We only consider functions from the real line to the complex plane. Every integral is a Lesbesgue integral over the entire real line, unless mentioned otherwise.

Date: Summer 2021.

2. Properties of L^p

The function space L^p organizes functions which have a Fourier transform. We briefly recall some basic definitions and results.

Definition 2.1. Let $I \subset \mathbb{R}$ be an interval and $1 \leq p < \infty$. The $\mathbf{L^p(I)}$**-norm** of a function f is given by,

$$\|f\|_p = \left(\int_I |f(x)|^p \, dx \right)^{1/p}.$$

Functions are in $L^p(I)$ if their $L^p(I)$-norm is finite.

Theorem 2.2 (L^p Completeness)**.** *Let $I \subset \mathbb{R}$ be an interval and $1 \leq p < \infty$. Then, the function space $L^p(I)$ is complete.*

Proof. See Theorem 15.4 in [2]. □

Theorem 2.3 (Hölder's Inequality)**.** *If $1 < p, q < \infty$, $p^{-1} + q^{-1} = 1$, and f and g are measurable, then,*

$$\int |fg| \, d\mu \leq \|f\|_p \|g\|_q.$$

This also holds if $p = \infty$ and $q = 1$ or if $p = 1$ and $q = \infty$. The inequality when $p = q = 2$ is called the Cauchy-Schwartz Inequality.

Proof. See Proposition 15.1 in [2]. □

Theorem 2.4 (Minkowski's Integral Inequality)**.** *Suppose $1 < p < \infty$, $f : X \times Y \to \mathbb{R}$, and $f \in L^p(\mathbb{R})$. Then,*

$$\left(\int_X \left| \int_Y f(x,y) \, dy \right|^p dx \right)^{\frac{1}{p}} \leq \int_Y \left(\int_X |f(x,y)|^p \, dx \right)^{\frac{1}{p}} dy.$$

Proof. This follows from Hölder's Inequality and Fubini's theorem. See [5]. □

Theorem 2.5 (Dominated Convergence Theorem)**.** *Suppose that f_n are measurable real-valued functions and $f_n(x)$ converges to $f(x)$ for each x. Suppose there exists a non-negative integrable function g such that $|f_n(x)| \leq g(x)$ for all x. Then,*

$$\lim_{n \to \infty} \int f_n \, d\mu = \int f \, d\mu.$$

Proof. See Theorem 7.9 in [2]. □

Proposition 2.6. *Functions in the intersection $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ are dense in $L^2(\mathbb{R})$.*

Proof. Let f be a function in $L^2(\mathbb{R})$. Then, define the sequence of functions,

$$f_n(x) := \begin{cases} f(x), & |x| < n \\ 0, & |x| \geq n. \end{cases}$$

We know f_n is in $L^2(\mathbb{R})$ from the fact that f and zero are in $L^2(\mathbb{R})$. Furthermore, f_n is also in $L^1(\mathbb{R})$ because, by the Cauchy-Schwartz inequality, we have,

$$\|f_n\|_1 = \int_{|x| \leq n} |f(x) \cdot 1| \, dx \leq \left(\int_{|x| \leq n} |f(x)|^2 \, dx \right)^{\frac{1}{2}} \left(\int_{|x| \leq n} 1^2 \, dx \right)^{\frac{1}{2}} < \infty.$$

We want to show that the sequence f_n converges to f in the $L^2(\mathbb{R})$-norm as n approaches infinity. By definition of f_n, the difference $f - f_n$ converges to zero for each x. Additionally, we know that the difference $|f - f_n|^2$ is at most $|f|^2$. Thus, the Dominated Convergence Theorem gives,

$$\lim_{n\to\infty} \int |f - f_n|^2 = 0,$$

which implies $\lim_{n\to\infty} \|f - f_n\|_2 = 0$, as desired. □

Theorem 2.7 (Hardy's Inequality). *Suppose $1 < p < \infty$, $f : (0, \infty) \to \mathbb{R}$, and $f \in L^p(\mathbb{R})$. Define $g(x) := \frac{1}{x}\int_0^x f(y)\,dy$. Then,*

$$\left(\int_0^\infty |g(x)|^p\,dx\right)^{\frac{1}{p}} \leq \frac{p}{p-1}\left(\int_0^\infty |f(x)|^p\,dx\right)^{\frac{1}{p}}.$$

Proof. Use change of variables to rewrite,

$$g(x) = \frac{1}{x}\int_0^x f(z)\,dz = \frac{1}{x}\int_{0/x=0}^{x/x=1} f(xy)x\,dy = \int_0^1 f(xy)\,dy.$$

Bound the left hand side with Minkowski's integral inequality,

$$\left(\int_0^\infty \left|\int_0^1 f(xy)\,dy\right|^p dx\right)^{\frac{1}{p}} \leq \int_0^1 \left(\int_0^\infty |f(xy)|^p\,dx\right)^{\frac{1}{p}} dy.$$

Use the same change of variables to spilt the integral. Then, evaluate the definite integral to finish the proof,

$$\int_0^1 y^{-\frac{1}{p}}\,dy\left(\int_0^\infty |f(x)|^p\,dx\right)^{\frac{1}{p}} = \frac{p}{p-1}\left(\int_0^\infty |f(x)|^p\,dx\right)^{\frac{1}{p}}.$$

□

Proposition 2.8 (Modified Hardy's Inequality). *Hardy's inequality holds for integrals bounded over the entire real line,*

$$\left(\int_{-\infty}^\infty |g(x)|^p\,dx\right)^{\frac{1}{p}} \leq \frac{p}{p-1}\left(\int_{-\infty}^\infty |f(x)|^p\,dx\right)^{\frac{1}{p}}.$$

Proof. For clarity, we will begin by splitting the integrals along zero,

$$\left(\int_{-\infty}^0 |g(x)|^p\,dx + \int_0^\infty |g(x)|^p\,dx\right)^{\frac{1}{p}} \leq \frac{p}{p-1}\left(\int_{-\infty}^0 |f(x)|^p\,dx + \int_0^\infty |f(x)|^p\,dx\right)^{\frac{1}{p}}.$$

We already know that the inequality is true over zero to positive infinity, so it suffices to show that,

$$\left(\int_{-\infty}^0 |g(x)|^p\,dx\right)^{\frac{1}{p}} \leq \frac{p}{p-1}\left(\int_{-\infty}^0 |f(x)|^p\,dx\right)^{\frac{1}{p}}.$$

This result follows from the exact same argument given for the original theorem. □

3. Properties of Fourier Transforms

Definition 3.1 (Fourier Transform in L^1)**.** Given the function $f \in L^1(\mathbb{R})$, the Fourier transform $\hat{f}$ is defined as,

$$\hat{f}(\xi) = \int f(x) e^{-i\xi x}\, dx,$$

for any $\xi \in \mathbb{R}$.

Property 3.2 (Derivative-to-Multiplication Property)**.** Let f be a differentiable function. If f and its first derivative f' are in $L^2(\mathbb{R})$, then the Fourier transform of f' is given by, $\widehat{(f')}(\xi) = i\xi\hat{f}(x)$.

Proof. Integrate the definition of the Fourier transform by parts,

$$\widehat{(f')}(\xi) = \int f'(x) e^{-i\xi x}\, dx = f(x)e^{-i\xi x}|_{\mathbb{R}} + i\xi \int f(x) e^{-i\xi x}\, dx.$$

The second term $i\xi \int f(x)e^{-i\xi x}\, dx$ equals $i\xi\hat{f}(x)$. So all that must be shown is that the first term evaluates to zero because f vanishes at infinity. Using the Fundamental Theorem of Calculus, we can write,

$$f(x) = \left([f(0)]^2 + 2\int_0^x f(t)f'(t)\, dt \right)^{\frac{1}{2}}.$$

We want to show that the radicand converges to zero at infinity. The first term converges because it is a constant. The second term converges by comparison to the $L^2(\mathbb{R})$-norms of f and f'; by the Cauchy-Schwartz Inequality, we have,

$$\begin{aligned} \lim_{x\to\pm\infty} 2\int_0^x |f(t)f'(t)|\, dt &= 2\int_0^{\pm\infty} |f(t)f'(t)|\, dt \\ &\leq 2\left(\int_0^{\pm\infty} |f(t)|^2\, dt\right)^{\frac{1}{2}} \left(\int_0^{\pm\infty} |f'(t)|^2\, dt\right)^{\frac{1}{2}} < \infty. \end{aligned}$$

Furthermore, f must vanish at infinity. Otherwise, at large positive and negative x, the distance between $f(x)$ and zero would be bounded below by a strictly positive number. This would make it impossible for the $L^2(\mathbb{R})$-norm of f to be finite. □

The previous result can be generalized to functions that are absolutely continuous, see Proposition 16.3 in [2]. We argued for functions in $L^2(\mathbb{R})$ because that is a key hypothesis in section five, but the fact that the Fourier transform relates derivatives to multiplication is of independent importance as a technique for solving differential equations.

Theorem 3.3 (Fourier Inversion in L^1)**.** *Suppose f and $\hat{f}$ are both in $L^1(\mathbb{R})$. Then,*

$$f(x) = \frac{1}{(2\pi)^n} \int \hat{f}(\xi) e^{i\xi x}\, d\xi$$

almost everywhere.

Proof. In explicit terms of the Fourier transform, we want to show,

$$f(x) = \frac{1}{(2\pi)^n} \int \left(\int f(y) e^{-i\xi y}\, dy \right) e^{i\xi x}\, d\xi, \tag{3.4}$$

but we are unable to change the order of integration because the integrand, $f(y)e^{-i\xi(x-y)}$, is not integrable in $\mathbb{R}^2$. To resolve this, we will make use of the fact that the Gaussian function is the Fourier transform of itself (see equation 2.32 of [3]). Specifically, the Fourier transform of the Gaussian function, $G(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, is given by,

$$\hat{G}(\xi) = e^{-\frac{\xi^2}{2}}.$$

Multiply the Gaussian, $e^{-\frac{s^2\xi^2}{2}}$, to the integrand of equation 3.4 such that as the dilation parameter, $s > 0$, approaches zero, the term itself approaches one, so the integrand converges to $\hat{f}(\xi)e^{i\xi x}$ for each x. We have the following approximation,

$$I_s(x) := \frac{1}{(2\pi)^n} \iint f(y) e^{i\xi(x-y)} e^{-\frac{s^2\xi^2}{2}} \, dy d\xi.$$

Since $|\hat{f}(\xi)e^{-\frac{s^2\xi^2}{2}} e^{i\xi x}| \leq |\hat{f}(\xi)|$ for all x, we have,

$$\lim_{s\to 0} I_s(x) = \frac{1}{(2\pi)^n} \left(\lim_{s\to 0} \int \hat{f}(\xi) e^{-\frac{s^2\xi^2}{2}} e^{i\xi x} \, dx \right) = \frac{1}{(2\pi)^n} \int \hat{f}(\xi) e^{i\xi x} \, d\xi, \tag{3.5}$$

by the Dominated Convergence Theorem. Next, notice that we can use Fubini's theorem to change the order of integration of I_s because it has a finite $L^1(\mathbb{R})$-norm. Define $g_s := \frac{1}{s} G\left(\frac{x}{s}\right)$. From change of variables and the definition of G, we get,

$$I_s(x) = \int f(y) \int \frac{1}{(2\pi)^n} e^{-\frac{s^2\xi^2}{2}} e^{i\xi(x-y)} \, d\xi dy = \int f(y) g_s(x-y) \, dy.$$

As s approaches zero, the approximate identity g_s concentrates around the origin and has an integral that always equals one (see Proposition 16.6 in [2]). Thus, we can verify that,

$$\lim_{s\to 0} \int |I_s(x) - f(x)| \, dx = \lim_{s\to 0} \iint g_s(x-y)|f(y) - f(x)| \, dy dx = 0,$$

which combines with line 3.5 to finish the proof.

□

Definition 3.6 (Convolution)**.** The convolution of two measurable functions f and g is defined by,

$$(f * g)(x) = \int f(x-y) g(y) \, dy,$$

provided the integral exists.

Property 3.7 (Convolution Theorem)**.** Let $f, g \in L^1(\mathbb{R})$. Then, $h := f * g$ is also in $L^1(\mathbb{R})$, and

$$\hat{h}(\xi) = \hat{f}(\xi)\hat{g}(\xi).$$

Proof. By Fubini's theorem and change of variables,

$$\int |h(x)| \, dx \leq \iint |f(y)||g(x-y)| \, dy dx = \left(\int |f(y)| \, dy \right) \left(\int |g(x)| \, dx \right) < \infty.$$

Thus, h is in $L^1(\mathbb{R})$. Next, use Fubini's theorem and change of variables to calculate,

$$\begin{aligned}\hat{h}(\xi) &= \int \left(\int f(y)g(x-y)\, dy \right) e^{-i\xi x}\, dx \\ &= \iint f(y)g(x-y)e^{-i\xi x}\, dydx \\ &= \iint e^{-i\xi(x+y)} f(y)g(x)\, dydx = \hat{f}(\xi)\hat{g}(\xi),\end{aligned}$$

as desired. □

We can use the properties introduced thus far to explain how Fourier transforms lie at the heart of signal processing. Let the function $f(t)$ represent a signal, where t is time. An operator L is called a **time-invariant operator** if an input signal delayed by time τ is mapped by L to an output signal that is also delayed by τ. That is, given $L(f_i) = f_f$, we have,

$$L(f_i(t-\tau)) = f_f(t-\tau).$$

In a continuous time system, all time-invariant operators can be described with convolution [3]. Furthermore, the convolution theorem implies that these time-invariant operators act as multiplication in Fourier space. Let $\vee$ denote Fourier inversion and $\times$ denote multiplication. When the relevant operators apply, the following diagram commutes,

$$\begin{array}{ccc} f & \xrightarrow{\ \wedge\ } & \hat{f}(\xi) \\ {\scriptstyle L}\big\downarrow & & \big\downarrow{\scriptstyle \times\hat{g}(\xi)} \\ (f * g)(t) & \xleftarrow[\ \vee\ ]{} & \hat{f}(\xi)\hat{g}(\xi). \end{array}$$

A signal processer uses the Fourier transform to decompose signals into complex exponentials, multiplies them with a transfer function $\hat{g}$, and then inverts the product to get the time-invariant operation $L(f)$. Multiplication is more simple to compute than the convolution itself (the Fourier transform *diagonalizes* the convolution operator), so Fourier transforms are well suited for computers and models.

Notations 3.8. Let f^* denote the complex conjugate of $f \in \mathbb{C}$. If f is a function, let $\bar{f}(x) = [f(-x)]^* = f^*(-x)$.

Property 3.9 (Complex Conjugate Property)**.** If the function $f(x)$ has a Fourier transform equal to $\hat{f}(\xi)$, then its complex conjugate $f^*(x)$ has a Fourier transform equal to $\hat{f}^*(-\xi)$.

Proof. This is just a calculation,

$$\hat{f}^*(x) = \int f^*(x)e^{-i\xi x}\, dx = \left[\int f(x)e^{i\xi x} \right]^* = \hat{f}^*(-\xi).$$

□

Theorem 3.10 (Parseval's, Plancherel's Theorem)**.** *If f and g are in $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then Parseval's theorem states,*

$$\int f(x)g^*(x)\, dx = \frac{1}{(2\pi)^n} \int \hat{f}(\xi)\hat{g}^*(\xi)\, d\xi.$$

For $g = f$, Plancherel's theorem follows,

$$\|f\|_2 = \frac{1}{(2\pi)^{n/2}}\|\hat{f}\|_2.$$

Proof. Let $h := f * \bar{g}$. The convolution theorem and complex conjugate property indicate that $\hat{h}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$. The remaining step is to take $h(0)$ and change the variables such that the inverse Fourier transform of $h(0)$ finishes the proof,

$$\int f(x)g^*(x)\,dx = h(0) = \frac{1}{(2\pi)^n}\int \hat{h}(\xi)\,d\xi = \frac{1}{(2\pi)^n}\int \hat{f}(\xi)\hat{g}^*(\xi)\,d\xi.$$

□

4. The General Uncertainty Principle

In this section, we take a moment to introduce the Fourier uncertainty principle in general terms in order to de-mystify the entire class of theorems which follow. In essence, we want to study the epistemic tradeoff between a nonzero function and its Fourier transform. This can be illustrated by manipulating the localization of a function in a similar way to how we treated the Gaussian function in our argument for the Fourier Inversion theorem in section three.

Property 4.1 (Dilation Property)**.** Let f be a function in $L^2(\mathbb{R})$. Consider the function,

$$f_s(x) := \frac{1}{\sqrt{s}} f\left(\frac{x}{s}\right),$$

where the parameter $0 < s < \infty$ dilates the localization of the function. Note that for every s, the $L^2(\mathbb{R})$-norms of f_s and f are equal. So, for any variations in s, the consequent localization of f_s and f is the same. Calculate the Fourier transform using change of variables $u = \frac{x}{s}$,

$$\begin{aligned}\hat{f}_s(\xi) = \frac{1}{\sqrt{s}}\int f\left(\frac{x}{s}\right)e^{-i\xi x}\,dx &= \frac{1}{\sqrt{x/u}}\int f(u)e^{-i(\xi s)u}s\,du \\ &= \sqrt{s}\int f(u)e^{-i(\xi s)u}\,du = \sqrt{s}\hat{f}(s\xi).\end{aligned}$$

Notice that s has opposite proportional effects on the domain of f_s and $\hat{f}_s$. This property holds for any function with a defined Fourier transform and an appropriately-constructed dilation. A visual using a Gaussian function can be found in [6].

Qualitatively, this means a narrow function has a wide Fourier transform, and a wide function has a narrow Fourier transform. In either domain, a wider function means there is literally a wide distribution of data, so there *always* exists uncertainty in one domain.

The tradeoff is intuitive in the case of a time and frequency domain; if one observes a signal for a long duration, they can be confident about only a limited range of corresponding frequencies. For most physical systems, it is plausible that the "true" frequency will reliably lie within this interval. However, as we will show the next section, there is no natural reliability for quantum mechanical systems. Indeed, the extraordinary parts of the uncertainty principle arise from the ideas that the principle describes, rather than the abstraction itself.

5. The Heisenberg Uncertainty Principle

The study of uncertainty principles began with Werner Heisenberg's argument that it is impossible to simultaneously determine a free particle's position and momentum to arbitrary precision. In quantum mechanics, the wave function of position is the Fourier transform of the wave function of momentum. In this section, we will prove Heisenberg's uncertainty principle, but initially formulate the principle in a way that clearly expounds from the properties we established in section two and three. Then, we will reformulate our result in terms of the aforementioned wave functions and discuss its implications.

To genuinely analyze the uncertainty principle in the framework of quantum mechanics, we must center our attention on functions that are in **Hilbert spaces**. Hilbert spaces are complete normed linear spaces that have an inner product [2]. Physicists represent quantum measurements in terms of linear operators on a Hilbert space. The exact nature of the Hilbert space depends on the observables one wants to describe. This paper concerns position and momentum, which are characterized by $L^2(\mathbb{R})$, but *not* $L^1(\mathbb{R})$ functions. Thus, we must extend all the properties of Fourier transforms in the previous section from $L^1(\mathbb{R})$ to $L^2(\mathbb{R})$ via the following density argument.

Remark 5.1 (Fourier Analysis in L^2)**.** Recall that functions in the intersection $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ are dense in $L^2(\mathbb{R})$ (proposition 2.5). This means, given a function $f \in L^2(\mathbb{R})$, we can define a sequence of functions $f_n, f_m \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ that converge pointwise to f. In other words, f_n and f_m are Cauchy sequences, so the difference of their respective $L^2(\mathbb{R})$-norms, $\|f_n - f_m\|_2$, gets arbitrarily small for large enough n and m. Moreover, since each function in f_n is in $L^1(\mathbb{R})$, their corresponding Fourier transform is well-defined. Denote this sequence $\hat{f}_n$. The key is that $\hat{f}_n$ forms a Cauchy sequence because,

$$\|\hat{f}_n - \hat{f}_m\|_2 = (2\pi)^{-n/2}\|f_n - f_m\|_2,$$

holds by Plancherel's theorem; the difference of the norms of the Fourier transforms approaches zero as n and m increase. The final realization is that this Cauchy property implies that $\hat{f}_n$ must converge to an element of $L^2(\mathbb{R})$ because $L^2(\mathbb{R})$ is complete (theorem 2.2). Thus, there must exist a Fourier transform $\hat{f} \in L^2(\mathbb{R})$ such that the the difference $\|\hat{f} - \hat{f}_n\|$ converges to zero, as desired.

Now that we've established that the we can do Fourier analysis with functions in the Hilbert space $L^2(\mathbb{R})$, we can move on and prove the Heisenberg uncertainty principle. For simplicity, we begin with a lemma.

Lemma 5.2. *If f is differentiable, then,*

$$\left(\int |f|^2 \, dx\right)^2 \leq 4\left(\int |xf(x)|^2 \, dx\right)\left(\int |f'|^2 \, dx\right).$$

Proof. Using the Fundamental Theorem of Calculus, we can rewrite the integrand on the left hand side as follows,

$$\left(\int |f|^2 \, dx\right)^2 = \left(\int \left|xf(x)\frac{1}{x}\int_0^x f'(y)\, dy\right|^2 \, dx\right)^2.$$

Then, we can split the integrand into parts and apply the Cauchy-Schwartz Inequality to get,

$$\int |xf(x)|^2 \left|\frac{1}{x}\int_0^x f'(y)\,dy\right|^2 dx \le \left(\int |xf(x)|^2\,dx\right)^{\frac{1}{2}} \left(\int \left|\frac{1}{x}\int_0^x f'(y)\,dy\right|^2 dx\right)^{\frac{1}{2}}.$$

Square both sides to recover the left hand side,

$$\left(\int |f|^2\,dx\right)^2 \le \left(\int |xf(x)|^2\,dx\right)\left(\int \left|\frac{1}{x}\int_0^x f'(y)\,dy\right|^2 dx\right).$$

All that remains is to bound the second term on the right. This can be done using the modified Hardy's inequality, but that only applies if the integrand is non-negative. So we must write,

$$\left(\int \left|\frac{1}{x}\int_0^x f'(y)\,dy\right|^2 dx\right) \le \left(\int \left|\frac{1}{x}\int_0^x |f'(y)|\,dy\right|^2 dx\right).$$

Now, let $g(x) := \frac{1}{x}\int_0^x |f'(y)|\,dy$. Apply the modified Hardy's inequality when $p = 2$ to get $\left(\int |g|^2\,dx\right)^{\frac{1}{2}} \le 2\left(\int |f'|^2\,dx\right)^{\frac{1}{2}}$. Then, square both sides to get the upper bound $4\int |f'|^2\,dx$, as desired. □

Theorem 5.3 (The Heisenberg Uncertainty Principle [2])**.** *Let $a, b \in \mathbb{R}$ and f be a differentiable function. If f and f' are in $L^2(\mathbb{R})$ then,*

$$\left(\int (x-a)^2|f(x)|^2\,dx\right)\left(\int (u-b)^2|\hat{f}(u)|^2\,du\right) \ge \frac{\pi}{2}\left(\int |f(x)|^2\,dx\right)^2.$$

Proof. Start with the inequality $\frac{4}{(2\pi)^n} \le \frac{2}{\pi}$. Multiply both sides by the integral $\int u^2|\hat{f}(u)|^2\,du$. Then, manipulate the left hand side as follows,

$$\begin{aligned}\frac{4}{(2\pi)^n}\int u^2|\hat{f}(u)|^2\,du &= \frac{4}{(2\pi)^n}\int |iu\hat{f}(u)|^2\,du \\ &= \frac{4}{(2\pi)^n}\int |(\hat{f'})(u)|^2\,du = 4\int |f'(x)|^2\,dx,\end{aligned}$$

where the second equality follows from the deriviative-to-multiplication property and the third from Plancherel's theorem. Next, multiply both sides by the integral $\int |xf(x)|^2\,dx$ to get,

$$4\left(\int |xf(x)|^2\,dx\right)\left(\int |f'(x)|^2\,dx\right) \le \frac{2}{\pi}\left(\int |xf(x)|^2\,dx\right)\left(\int u^2|\hat{f}(u)|^2\,du\right).$$

By lemma 5.2, we can bound the right hand side by the integral $(\int |f|^2\,dx)^2$. Multiplying both sides of the resultant inequality by $\frac{\pi}{2}$ yields the desired result when $a = b = 0$.

Now we will argue that the inequality holds for nonzero a, b. We want to show that the function $g(x) := e^{-ibx}f(x+a)$ counteracts any changes caused by nonzero a, b without loss of generality. First, consider nonzero a. From $|e^{-ibx}| = 1$ we get,

$$\int x^2|g(x)|^2\,dx = \int x^2|f(x+a)|^2\,dx,$$

and from change of variables we get,

$$\int x^2|f(x+a)|^2\,dx = \int (x-a)^2|f(x)|^2\,dx.$$

Likewise, consider nonzero b. Write $\hat{g}$ in terms of $\hat{f}$ using change of variables with $y = x + a$,

$$\begin{aligned}\hat{g}(u) = \int e^{-iux}g(x)\,dx &= \int e^{-ix(u+b)}f(x+a)\,dx \\ &= \int e^{-i(y-a)(u+b)}f(y)\,dy = e^{i(u+b)a}\hat{f}(u+b).\end{aligned}$$

For the same reasons as before, we have,

$$\int u^2|\hat{g}(u)|^2\,du = \int u^2|\hat{f}(u+b)|^2\,du = \int (u-b)^2|\hat{f}(u)|^2\,du,$$

as desired. □

Remark 5.4. We claim equality for the Heisenberg Uncertainty Principle holds if and only if f is a translation and modulation of a Gaussian function. Forcing equality in the Cauchy-Schwartz Inequality implies there must exist a $p \in \mathbb{R}$ such that the differential equation $f'(x) = -2pxf(x)$ holds. By separation of variables, we find that there exists an $q \in \mathbb{R}$ such that $f(x) = qe^{-bx^2}$ is a solution. When $a \neq 0$ and $b \neq 0$ we get the Gaussian solution,

$$f(x) = qe^{ibx-p(x-a)^2}. \tag{5.5}$$

The converse can be verified by using this Gaussian to calculate the lower bound.

The main difference between the Heisenberg Uncertainty Principle and the dilation property is that the former multiplies the **variance** of the function and its Fourier transform so that there is a clearly delineated lower bound. In practice, the Heisenberg Uncertainty principle captures the experimental fact that measurements of quantum mechanical systems with low position uncertainty correlate to a large range of momentum measurements. This partly follows from the probabilistic interpretation of quantum mechanical systems, which use variance to represent the uncertainty around the average position and momentum measurements. The derivations of the position and momentum wave functions are outside the scope of this mathematics paper, so we will borrow the result from [3] and omit the physics.

Let $f \in L^2(\mathbb{R})$ be the wave function of a free particle. The probability density that this particle is at a position x is given by $\frac{|f(x)|^2}{\|f\|_2}$. The probability density that its momentum is equal to p is given by $\frac{|\hat{f}(p)|^2}{2\pi\|f\|_2}$. From probability theory, we can write the variance σ^2 by multiplying displacement around the domain in the integrand,

$$\sigma_x^2 = \frac{1}{\|f\|_2}\int (x-x_0)^2|f(x)|^2\,dx \tag{5.6}$$

$$\sigma_p^2 = \frac{1}{2\pi\|f\|_2}\int (p-p_0)^2|\hat{f}(p)|^2\,dp. \tag{5.7}$$

Expressing theorem 5.3 in terms of variance gives,

$$\sigma_x^2\sigma_p^2 \geq \frac{1}{4}\|f\|_2^{-1}\|\hat{f}\|_2^{-4}. \tag{5.8}$$

Note that the generalization in the last part of our argument for theorem 5.3 implies that translations and modulations of f do not affect the variance of f. This further implies that dividing f by $\|f\|_2$ does not change variance, so we may assume $\|f\|_2 = 1$. For nonzero f, Plancherel's theorem gives,

$$\sigma_x^2 \sigma_p^2 \geq \frac{1}{4}, \tag{5.9}$$

which is a more concise, applicable, and standard formulation of the Heisenberg Uncertainty principle.

6. The Primary Uncertainty Principle

We will close by applying recently-discovered techniques. In [4], readers are introduced to a never-before-seen proof template of Fourier uncertainty principles. The primary theorem stems from the fact that Fourier transforms, or more generally, "k-Hadamard" operators, are bounded as an operator from $L^1 \to L^\infty$, and also unitary. The authors use this to prove an entire survey of inequalities; we will show their argument for the Heisenberg uncertainty principle.

The following theorem captures the operator properties which yield uncertainty principles in infinite dimensions. We weakened the original statement to avoid having to introduce new mathematical concepts; the reader is encouraged to read [4] if they are interested in the general statement.

Theorem 6.1 (Primary Uncertainty Principle, Infinitary version, Weakened)**.** *Let $f \in L^1(\mathbb{R})$ such that $\hat{f} \in L^1(\mathbb{R})$. Then,*

$$\|f\|_1 \|\hat{f}\|_1 \geq \|f\|_\infty \|\hat{f}\|_\infty.$$

Proof. For any $\xi \in \mathbb{R}$, we have,

$$|\hat{f}(\xi)| = \left| \int f(x) e^{-i\xi x} \, dx \right| \leq \int |f(x)| \, dx = \|f\|_1,$$

which implies that $\|\hat{f}\|_\infty \leq \|f\|_1$. Similarly, for any $x \in \mathbb{R}$, we have,

$$|f(x)| = \left| \frac{1}{(2\pi)^n} \int \hat{f}(\xi) e^{i\xi x} \, d\xi \right| \leq \int |\hat{f}(\xi)| \, d\xi = \|\hat{f}\|_1,$$

which implies that $\|f\|_\infty \leq \|\hat{f}\|_1$. Since norms are non-negative, we can multiply both of these inequalities to get $\|f\|_\infty \|\hat{f}\|_\infty \leq \|f\|_1 \|\hat{f}\|_1$, as desired. $\square$

We can use this to observe another uncertainty principle, which relates the L^p-norm ratios of functions to their Fourier transforms.

Theorem 6.2 (Norm Uncertainty Principle, Weakened)**.** *For any $1 \leq p \leq \infty$ and any non-zero $f : \mathbb{R} \to \mathbb{C}$, we have,*

$$\frac{\|f\|_1}{\|f\|_p} \cdot \frac{\|\hat{f}\|_1}{\|\hat{f}\|_p} \geq 1.$$

Proof. For any function $g : \mathbb{R} \to \mathbb{C}$, we calculate,

$$\|g\|_p^p = \int |g(x)|^p \, dx \leq \|g\|_\infty^{p-1} \int |g| \, dx = \|g\|_\infty^{p-1} \|g\|_1.$$

Multiplying $\|g\|_1^{p-1}$, we get $\|g\|_1^{p-1}\|g\|_p^p \leq \left(\frac{\|g\|_1}{\|g\|_\infty}\right)^{\frac{p-1}{p}}$. Plug in f and $\hat{f}$ in for g and then multiply the resulting inequalities. Then, the proof is finished by writing,

$$\frac{\|f\|_1}{\|f\|_p} \cdot \frac{\|\hat{f}\|_1}{\|\hat{f}\|_p} \geq \left(\frac{\|f\|_1}{\|f\|_\infty} \cdot \frac{\|\hat{f}\|_1}{\|\hat{f}\|_\infty}\right)^{\frac{p-1}{p}} \geq 1,$$

where the second inequality follows from the primary uncertainty principle. □

Finally, we will prove the Heisenberg Uncertainty Principle. Recall the definition of variance from display lines 5.6 and 5.7. We will, once again, assume $\|f\|_2 = 1$.

Theorem 6.3 (The Heisenberg Uncertainty Principle [4])**.** *There exists a constant $C > 0$ such that for any nonzero function $f \in L^1(\mathbb{R})$ such that $\hat{f} \in L^1(\mathbb{R})$, we have,*

$$\sigma_x^2 \sigma_p^2 \geq C\|f\|_2^2 \|\hat{f}\|_2^2.$$

We want to use the norm uncertainty principle when $p = 2$. This means that we should look for a lower bound on the variance of an arbitrary function g that is in terms of the ratio $\frac{\|g\|_1}{\|g\|_2}$.

Lemma 6.4. *There exists some constant $c > 0$ such that, for any nonzero function g, we have,*

$$\frac{\sigma^2(g)}{\|g\|_2^2} \geq c\left(\frac{\|g\|_1}{\|g\|_2}\right)^4.$$

Proof. Let $T = \frac{1}{8}\left(\frac{\|g\|_1}{\|g\|_2}\right)^2$. Note that the $L^2([-T,T])$-norm of g is at least the $L^2(\mathbb{R})$-norm of g because $[-T,T]$ is a proper subset of $\mathbb{R}$. Multiply both sides of this inequality by $\frac{\|g\|_1}{2\|g\|_2}$ to get,

$$\left(\int_{-T}^{T} |g|^2\right)^{\frac{1}{2}} \left(\frac{\|g\|_1}{2\|g\|_2}\right) \leq \|g\|_2 \left(\frac{\|g\|_1}{2\|g\|_2}\right) = \frac{1}{2}\|g\|_1.$$

Note that for the second term on the left-hand side, we can write,

$$\left(\frac{\|g\|_1}{2\|g\|_2}\right) = \left(\frac{1}{4}\left(\frac{\|g\|_1}{\|g\|_2}\right)^2\right)^{\frac{1}{2}} = \left(\int_{-T}^{T} 1\right)^{\frac{1}{2}}.$$

Thus, by Cauchy Schwartz, we have $\int_{-T}^{T} |g| \leq \left(\int_{-T}^{T} |g|^2\right)^{\frac{1}{2}} \left(\int_{-T}^{T} 1\right)^{\frac{1}{2}}$, which implies,

$$\int_{-T}^{T} |g| \leq \frac{1}{2}\|g\|_1.$$

Note that this further implies that $\int_{|x|>T} |g| \geq \frac{1}{2}\|g\|_1$. Apply Cauchy Schwartz as follows,

$$\begin{aligned}\frac{1}{2}\|g\|_1 \leq \int_{|x|>T} \frac{x}{x}|g(x)|\, dx &\leq \left(\int_{|x|>T} x^{-2}\, dx\right)\left(\int x^2 |g(x)|^2\, dx\right)^{\frac{1}{2}} \\ &= \sqrt{2/T} \cdot \sigma(g).\end{aligned}$$

Then, isolate $\frac{\sigma(g)^2}{\|g\|_2^2}$ to get

$$\frac{T}{8}\frac{\|g\|_1^2}{\|g\|_2^2} \leq \frac{\sigma(g)^2}{\|g\|_2^2}.$$

All that remains is the bound the left hand side of this result below by $c\left(\frac{\|g\|_1}{\|g\|_2}\right)^4$. Take any $0 < c \leq \frac{1}{64}$. Then, multiply both sides of the second inequality by $\left(\frac{\|g\|_1}{\|g\|_2}\right)^4$ to get $c\left(\frac{\|g\|_1}{\|g\|_2}\right)^4 \leq \frac{T}{8}\left(\frac{\|g\|_1}{\|g\|_2}\right)^2$, as desired. □

Proof of Theorem 6.3. Write the result from lemma 6.4 as follows,

$$\frac{\|g\|_1}{\|g\|_2} \leq \left(\frac{1}{c}\frac{\sigma^2(g)}{\|g\|_2^2}\right)^4.$$

Plug in f and $\hat{f}$ for g. Then, multiply both inequalities and define C as the product of the constants terms from both functions. □

Since our argument does not depend on any analytic techniques (e.g. integration by parts), nor properties of Fourier transforms, the Heisenberg Uncertainty principle can be generalized to operators beyond Fourier transforms, such as the n-th moment of the position and momentum operators in quantum mechanics. Notably, the constant term C is *not* the most optimal constant in the inequality. This implies that the tightness of the constant in uncertainty principles does not depend on same the properties which the uncertainty principle depend on, but rather something deeper about the operator itself.

Acknowledgments

This paper would not have been possible without the guidance I received from my mentor, Phillip Lo. I am also grateful to Professor Daniil Rudenko, who prepared the apprenticeship program, and Professor Peter May, who organized the entire REU experience.

References

[1] Gerald B. Folland, Alladi Sitaram. *The Uncertainty Principle: A Mathematical Survey.* The Journal of Fourier Analysis and Applications. Vol 3, Number 3, 1997.
[2] Richard F. Bass. *Real Analysis for Graduate Students.* Version 4.2.
[3] Stéphane Mallat. *A wavelet tour of signal processing.* Third Edition. 2008.
[4] Avi Wigderson, Yuval Wigderson. *The uncertainty principle: Variations on a theme.* Bull. Amer. Math. Soc. 58 (2021), 225-261. https://www.ams.org/journals/bull/2021-58-02/S0273-0979-2021-01715-0/
[5] https://calclab.math.tamu.edu/~sivan/math663_06a/appendix2.pdf
[6] https://www.desmos.com/calculator/jmksqol7ca