MATH 4330/5330, Fourier Analysis
Section 10
The L^2 Fourier Transform on the Line

The Fourier transform on the real line appears to be restricted to absolutely integrable (L^1) functions. However, just as in the case of the Fourier transform on the circle, there is a remarkable L^2 theory. (Remember the marvelous Parseval Equality!) This time the L^2 theory requires some more detailed analysis. There is no help for it. We must deal with *Dirac δ functions*. The following is a fundamental construction.

EXERCISE 10.1. Let $k(x)$ be a nonnegative function on the real line $\mathbb{R}$ for which $\int_{\mathbb{R}} k(x)\,dx = 1$. For each $t > 0$, define

$$k_t(x) = \frac{1}{t}k(\frac{x}{t}).$$

(a) Prove that $\int_{\mathbb{R}} k_t(x)\,dx = 1$ for all $t > 0$.
(b) Show that, for each $\delta > 0$, we have

$$\lim_{t\to 0}\int_{-\delta}^{\delta} k_t(x)\,dx = 1.$$

HINT: Write down the integral, and change variables. Then take the limit.
(c) Show that part (b) implies that for each $\delta > 0$

$$\lim_{t\to 0}[\int_{-\infty}^{-\delta} k_t(x)\,dx + \int_{\delta}^{\infty} k_t(x)\,dx] = 0.$$

(d) Can you sketch the graph of k_t and deduce what must be happening to it as t approaches 0?

REMARK. The construction in Exercise 10.1 produces an example of a *Dirac δ function*. These particular examples are nonnegative functions (parameterized families of functions). The Dirichlet kernel is an example of a Dirac δ function that is **not** nonnegative, but rather oscillatory. Likewise, the kernel K_B of Section 10 is a Dirac δ function that is oscillatory

The Dirac δ functions constructed in Exercise 10.1 have an important connection with the convolution operation. They form what are called *approximate identities*.

THEOREM 10.1. *Let $k(x)$ be a nonnegative function on $\mathbb{R}$ for which $\int_{\mathbb{R}} k(x)\,dx =$ 1. For each $t > 0$, set*

$$k_t(x) = \frac{1}{t}k(\frac{x}{t}).$$

Suppose f is a bounded function on $\mathbb{R}$, and assume that f is continuous at a point x. Then

$$f(x) = \lim_{t\to 0} f * k_t(x) = \lim_{t\to 0}\int_{\mathbb{R}} f(x-y)k_t(y)\,dy.$$

PROOF. Let ϵ be a positive number. We wish to show that $\lim_{t\to 0}|f(x) - f * k_t(x)| \le \epsilon$. (Why will this suffice to prove the theorem?) First, choose a $\delta > 0$ so that $|f(x) - f(x-y)| < \epsilon$ if $|y| < \delta$. (How is this possible?) Next, since f is a bounded function, let M be a number for which $|f(t)| \le M$ for all real numbers t. We then have

$$\begin{aligned}
|f(x) - f * k_t(x)| &= |f(x) \int_{\mathbb{R}} k_t(y)\,dy - \int_{\mathbb{R}} f(x-y)k_t(y)\,dy| \\
&= |\int_{\mathbb{R}} (f(x) - f(x-y))k_t(y)\,dy| \\
&\le \int_{\mathbb{R}} |f(x) - f(x-y)|k_t(y)\,dy \\
&= \int_{-\infty}^{-\delta} |f(x) - f(x-y)|k_t(y)\,dy \\
&\quad + \int_{-\delta}^{\delta} |f(x) - f(x-y)|k_t(y)\,dy \\
&\quad + \int_{\delta}^{\infty} |f(x) - f(x-y)|k_t(y)\,dt \\
&\le \int_{-\infty}^{-\delta} (|f(x)| + |f(x-y)|)k_t(y)\,dy \\
&\quad + \int_{-\delta}^{\delta} \epsilon k_t(y)\,dy \\
&\quad + \int_{\delta}^{\infty} (|f(x)| + |f(x-y)|)k_t(y)\,dy \\
&\le \int_{-\infty}^{-\delta} 2Mk_t(y)\,dy + \int_{-\delta}^{\delta} \epsilon k_t(y)\,dy + \int_{\delta}^{\infty} 2Mk_t(y)\,dy \\
&= \int_{-\infty}^{-\delta} 2Mk_t(y)\,dy + \int_{\delta}^{\infty} 2Mk_t(y)\,dy + \int_{-\delta}^{\delta} \epsilon k_t(y)\,dy \\
&= 2M(\int_{-\infty}^{-\delta} k_t(y)\,dy + \int_{\delta}^{\infty} k_t(y)\,dy) + \int_{-\delta}^{\delta} \epsilon k_t(y)\,dy \\
&\le 2M(1 - \int_{-\delta}^{\delta} k_t(y)\,dy) + \epsilon \int_{\mathbb{R}} k_t(y)\,dy \\
&= 2M(1 - \int_{-\delta}^{\delta} k_t(y)\,dy) + \epsilon.
\end{aligned}$$

So, from part (b) of Exercise 10.1, we see that

$$\lim_{t\to 0} |f(x) - f * k_t(x)| \le \epsilon.$$

Since this is true for an arbitrary ϵ, it follows that

$$\lim_{t\to 0} |f(x) - f * k_t(x)| = 0,$$

as desired.

THEOREM 10.2. *Let k and k_t be as in Exercise 10.1. Then*

(1) *For each real number ω, we have*

$$\lim_{t \to 0} \widehat{k_t}(\omega) = 1.$$

(2) *If f is any absolutely integrable function, then*

$$\lim_{t \to 0} \widehat{f * k_t}(\omega) = \widehat{f}(\omega).$$

PROOF. We calculate

$$\begin{aligned}
\widehat{k_t}(\omega) &= \int_{\mathbb{R}} k_t(x) e^{-2\pi i x \omega}\, dx \\
&= \int_{\mathbb{R}} \frac{1}{t} k(\frac{x}{t}) e^{-2\pi i x \omega}\, dx \\
&= \int_{\mathbb{R}} k(x) e^{-2\pi i t x \omega}\, dx \\
&= \int_{\mathbb{R}} k(x) e^{-2\pi i x (t\omega)}\, dx \\
&= \widehat{k}(t\omega).
\end{aligned}$$

Therefore, since $\widehat{k}$ is continuous,

$$\lim_{t \to 0} \widehat{k_t}(\omega) = \lim_{t \to 0} \widehat{k}(t\omega) = \widehat{k}(0) = 1,$$

which proves part (1).

Now, part (2) is easy:

$$\lim_{t \to 0} \widehat{f * k_t}(\omega) = \lim_{t \to 0} \widehat{f}(\omega) \widehat{k_t}(\omega) = \widehat{f}(\omega).$$

EXERCISE 10.2. Can you explain why these Dirac δ functions are called approximate identities? For instance, suppose f is a bounded function that is continuous everywhere. What can you say about the parameterized family of functions $f * k_t$? Is there a single function g such that $f = f * g, i.e., an actual identity for convolution$?

EXERCISE 10.3. (a) Verify that the function defined by

$$k(x) = \frac{1}{\pi} \frac{1}{1 + x^2}$$

satisfies the hypotheses of the preceding theorem, and write down an explicit formula for the functions k_t. (This Dirac δ function is called the *Poisson kernel.*

(b) Let k be a nonnegative function on $\mathbb{R}$ for which $\int_{\mathbb{R}} k(x)\, dx = 1$. For each positive t, define $k_t(x) = (1/\sqrt{t}) k(x/\sqrt{t})$. Show that the conclusions of Exercise 10.1 and Theorem 10.1 hold for this kernel. In fact, verify that all the same conclusions hold if, instead of $\sqrt{t}$, we use any positive power of t or, for that matter, any scalar multiple of a positive power of t, e.g., $\sqrt{4\pi t}$.

(c) Let $g(x) = e^{-\pi x^2}$. For each $t > 0$ define $g_t(x) = (1/\sqrt{4\pi t})g(x/\sqrt{4\pi t})$. Use part (b) to prove that the conclusions of the preceding theorem hold for this kernel. Work out exactly what the functions g_t are. Do you recognize these functions? (This Dirac δ function is called the *heat kernel*.)

Here is a generalization of Theorem 10.1.

EXERCISE 10.4. Suppose the f in Theorem 10.1 is bounded and uniformly continuous. Show that $f(x) = \lim_{t\to 0} f * k_t(x)$, and that the convergence is uniform. That is, show that, given an $\epsilon > 0$, there exists a $\alpha > 0$ such that $|f(x) - f * k_t(x)| \leq \epsilon$ for all x and for all t for which $t < \alpha$.
HINT: By the uniform continuity of f, choose a δ such that $|f(x) - f(x-y)| < \epsilon/2$ for all x and all $|y| < \delta$. Then, using part (b) of Exercise 10.1, choose α such that $1 - \int_{-\delta}^{\delta} k_t(x)\,dx < \epsilon/(4M)$ if $t < \alpha$. Now, make the same computation as in the proof of Theorem 10.1.

Here's another important analytical result about convolution.

THEOREM 10.3. *Let f and g be uniformly continuous, absolutely integrable functions. Then the convolution $f * g$ is uniformly continuous and absolutely integrable.*

PROOF. Note that

$$\begin{aligned} |f * g(x) - f * g(x')| &= |\int_{\mathbb{R}} (f(x-y) - f(x'-y))g(y)\,dy| \\ &\leq \int_{\mathbb{R}} |f(x-y) - f(x'-y)||g(y)|\,dy, \end{aligned}$$

so that, by the uniform continuity of f, we would have that

$$|f * g(x) - f * g(x')| \leq \int_{\mathbb{R}} \epsilon |g(y)|\,dy$$

providing $|x - x'|$ is less than some δ. The uniform continuity of $f * g$ follows from this.

As for the absolute integrability, note that

$$\begin{aligned} \int_{\mathbb{R}} |f * g(x)|\,dx &= \int_{\mathbb{R}} |\int_{\mathbb{R}} f(x-y)g(y)\,dy|\,dx \\ &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x-y)g(y)|\,dydx \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x-y)g(y)|\,dxdy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x)g(y)|\,dxdy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x)|\,dx|g(y)|\,dy \\ &< \infty. \end{aligned}$$

We come next to the analog in this $L^1(\mathbb{R})$ context of the Parseval Equality. Here, it is called the *Plancherel Theorem.* In preparation for its proof, we have the following technical exercises.

EXERCISE 10.5. Let f be an absolutely integrable function. Define a function f^* by $f^*(x) = \overline{f(-x)}$.

(a) Show that f^* is absolutely integrable.

(b) Prove that $\widehat{f^*}(\omega) = \overline{\widehat{f}(\omega)}$.

(c) Prove that $\widehat{f * f^*}(\omega) = |\widehat{f}(\omega)|^2$.

EXERCISE 10.6. Let k_t be the Poisson Kernel; i.e., the Dirac δ function determined by the function

$$k(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

See part (a) of Exercise 10.3.

(a) Prove that $k_t(-x) = k_t(x)$.

(b) Compute $\widehat{k}$. You should get

$$\widehat{k}(\omega) = e^{-|2\pi\omega|}.$$

(You probably can't compute this transform directly. Consult Exercise 8.8.)

(c) Prove that $\widehat{\widehat{k_t}} = k_t$.

THEOREM 10.4. (Plancherel Theorem) Let f be a continuous, absolutely integrable and square-integrable function on the real line. Then $\widehat{f}$ is square-integrable, and

$$\int_{\mathbb{R}} |f(x)|^2\,dx = \int_{\mathbb{R}} |\widehat{f}(\omega)|^2\,d\omega.$$

PROOF. We will use in this proof the Poisson kernel k_t of the preceding exercise. We will also use the function f^* of Exercise 10.5, as well as the earlier theorems in this section and the hat trick from Section 8. Watch out for the math! We have

$$\begin{aligned}
\int_{\mathbb{R}} |f(x)|^2\,dx &= \int_{\mathbb{R}} f(x)\overline{f^*}(0-x)\,dx \\
&= f * F^*(0) \\
&= \lim_{t\to 0} f * f^* * k_t(0) \\
&= \lim_{t\to 0} \int_{\mathbb{R}} f * f^*(x) k_t(0-x)\,dx \\
&= \lim_{t\to 0} \int_{\mathbb{R}} f * f^*(x) k_t(x)\,dx \\
&= \lim_{t\to 0} \int_{\mathbb{R}} f * f^*(x) \widehat{\widehat{k_t}}(x)\,dx \\
&= \lim_{t\to 0} \int_{\mathbb{R}} \widehat{f * f^*}(\omega) \widehat{k_t}(\omega)\,d\omega \\
&= \lim_{t\to 0} \int_{\mathbb{R}} |\widehat{f}(\omega)|^2 \widehat{k_t}(\omega)\,d\omega \\
&= \int_{\mathbb{R}} |\widehat{f}(\omega)|^2\,d\omega,
\end{aligned}$$

proving the theorem.

REMARK. In fact, the preceding theorem holds for any square-integrable function. That is, the L^2 norm of f equals the L^2 norm of its Fourier transform. Another perfect theorem!!

The Poisson Summation Formula

Let f be a function on the real line, and define a function F by

$$F(x) = \sum_{k=-\infty}^{\infty} f(x+k).$$

In the next two theorems, we will assume that f has sufficient extra properties so that F is differentiable. Then F is a periodic function, and we call F the *periodization* of f.

EXERCISE 10.7. Suppose f is differentiable everywhere and assume further that f has bounded support. that is, $f(x) = 0$ unless x belongs to a bounded interval $[-b, b]$. Show that in this case F is differentiable.

THEOREM 10.5. (Poisson Summation Formula) Suppose f and F are as above,. Then, for every x,

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

whence, in particular,

$$\sum_{k=-\infty}^{\infty} f(k) = \sum_{n=-\infty}^{\infty} \widehat{f}(n).$$

(This last equation is what is usually called the *Poisson summation formula.*)

PROOF. First, we calculate the Fourier coefficients $\widehat{F}(n)$ for the periodic function F.

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

So, using Theorem 6.1, we have

$$\begin{aligned}\sum_{k=-\infty}^{\infty} f(x+k) &= F(x) \\ &= \sum_{n=-\infty}^{\infty} \widehat{F}(n)e^{2\pi inx} \\ &= \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{2\pi inx}.\end{aligned}$$

Of course, the Poisson Summation Formula now follows by substituting $x = 0$ in the preceding.

EXERCISE 10.8. Let f be as in the preceding theorem, and let α be a positive number. Define $f_\alpha(x) = f(\alpha x)$, and let $F_\alpha(x) = \sum_{k=-\infty}^{\infty} f_\alpha(x+k)$. Write β for the reciprocal $1/\alpha$ of α. Use the theorem above, applied to f_α, to conclude that for every x

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

Then, deduce that

$$\sum_{k=-\infty}^{\infty} f(k\alpha) = \beta \sum_{n=-\infty}^{\infty} \widehat{f}(n\beta).$$

THEOREM 10.6. (Sampling Theorem) Suppose f has bounded support. That is, suppose that $f(x) = 0$ if x is outside a finite interval $[-b, b]$. If β is any positive number smaller than $1/2b$, then the function f is totally reconstructible from the "samples" $\{\widehat{f}(n\beta)\}$ of the Fourier transform $\widehat{f}$.

PROOF. Let β be a positive number smaller than $1/2b$, and write α for the reciprocal $1/\beta$. From the preceding exercise, we know that

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

Note that, for any x in the interval $[-b, b]$, the only value of k for which $f(x+k\alpha) \neq 0$ is $k = 0$. So, for any such x, we have

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

Hence, f is totally reconstructible from the values of $\widehat{f}$ at the points $n\beta$.

The Heat Equation on the Line

Now, suppose $u(t, x)$ is a solution of the heat equation on the line and that $u(0, x) = f(x)$ is its initial condition. Assume that f is square-integrable. (Of course, as in Section 1, we assume that $u(t,x)$ is square-integrable in the variable x for each $t > 0$.)

For each $t \geq 0$, write f_t for the function of x given by $f_t(x) = u(t,x)$. Note that $f_0(x) = u(0,x) = f(x)$. Then, for every $t > 0$, f_t is differentiable and square-integrable, so that Fourier's inversion formula works for it. Hence, we may write

$$\begin{aligned} u(t,x) &= f_t(x) \\ &= \int_{\mathbb{R}} \widehat{f_t}(\omega) e^{2\pi i\omega x}\, d\omega \\ &= \int_{\mathbb{R}} g_\omega(t) e^{2\pi i\omega x}\, d\omega, \end{aligned}$$

where $g_\omega(t)$ is just $\widehat{f_t}(\omega)$. In particular, $g_\omega(0) = \widehat{f_0}(\omega) = \widehat{f}(\omega)$.

So, writing

$$u(t,x) = \int_{\mathbb{R}} g_\omega(t) e^{2\pi i\omega x}\, d\omega,$$

what is the partial derivative of u with respect to t?

$$\begin{aligned} \frac{\partial u}{\partial t} u(t,x) &= \frac{d}{dt} \int_{\mathbb{R}} g_\omega(t) e^{2\pi i\omega x}\, d\omega \\ &= \int_{\mathbb{R}} {g_\omega}'(t) e^{2\pi i\omega x}. \end{aligned}$$

And, what is the second partial derivative of u with respect to x?

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} u(t,x) &= \frac{d^2}{dx^2} \int_{\mathbb{R}} g_\omega(t) e^{2\pi i\omega x}\, d\omega \\ &= \int_{\mathbb{R}} g_\omega(t)(-4\pi^2\omega^2) e^{2\pi i\omega x}\, d\omega. \end{aligned}$$

Now u satisfies the heat equation

$$\frac{\partial u}{\partial t} u(t,x) = \frac{\partial^2 u}{\partial x^2} u(t,x).$$

Hence, from our calculations above, for each $t > 0$, we have an equality between two Fourier transforms, i.e., of the function ${g_\omega}'(t)$ and of the function $-4\pi^2\omega^2 g_\omega(t)$. Hence, since the Fourier transform is 1-1, these two functions of ω must be equal for every t. Turning things around, we must have that for every ω the two functions $g'_\omega(t)$ and $-4\pi^2\omega^2 g_\omega(t)$ of t must be equal.

Now we can solve this simple differential equation for the functions g_ω, using Theorem 2.1 for instance. Indeed, we get that

$$g_\omega(t) = g_\omega(0) e^{-4\pi^2\omega^2 t} = \widehat{f}(\omega) e^{-4\pi^2\omega^2 t}.$$

For each t, let g_t be the function whose Fourier transform is given by $\widehat{g_t}(\omega) =$

$e^{-4\pi^2\omega^2 t}$. Then a formula for the solution to the heat equation is given by

$$\begin{aligned} u(t,x) &= \int_{\mathbb{R}} g_\omega(t) e^{2\pi i\omega x}\, dx \\ &= \int_{\mathbb{R}} \widehat{f}(\omega) e^{-4\pi^2\omega^2 t} e^{2\pi i\omega x} \\ &= \int_{\mathbb{R}} \widehat{f}(\omega)\widehat{g_t}(\omega) e^{2\pi i\omega x}\, d\omega \\ &= \int_{\mathbb{R}} \widehat{f * g_t}(\omega) e^{2\pi i\omega x}\, d\omega \\ &= f * g_t(x) \\ &= \int_{\mathbb{R}} f(x - y g_t(y)\, dy. \end{aligned}$$

Now, what is this function g_t? It's just the inverse Fourier transform of the function $e^{-4\pi^2\omega^2 t}$.

EXERCISE 10.9. (a) Show that g_t is given by

$$g_t(x) = \frac{1}{\sqrt{4\pi t}} g(\frac{x}{\sqrt{4\pi t}}) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}.$$

(This Dirac δ function is called the *heat kernel.*)

(b) Show that the solution to the heat equation on the line, having initial condition $f(x)$, is given by

$$u(t,x) = \int_{\mathbb{R}} f(x-y) \frac{1}{\sqrt{4\pi t}} e^{-\frac{y^2}{4t}}\, dy..$$

EXERCISE 10.10. Deduce that, for any square-integrable initial condition $f(x)$, there is a solution $u(t,x)$ of the heat equation. Moreover, show there is only one such solution.

Mathematical Formulation of the Heisenberg Uncertainty Principle

We begin with a famous inequality that holds for square-integrable functions, either on the circle or on the real line.

THEOREM 10.7. (Cauchy-Schwarz Inequality) Let f and g be two square-integrable functions. Then

$$\left|\int f(x)\overline{g(x)}\, dx\right|^2 \leq \int |f(x)|^2\, dx \times \int |g(x)|^2\, dx.$$

PROOF. Because the inequality we want to prove is certainly correct if g is the 0 function, we assume that g is not the 0 function. This means that $\int |g(x)|^2\, dx > 0$.

We will want to divide by this number in the argument below. Note next that the following calculation is correct for any complex number λ :

$$
\begin{aligned}
0 &\leq \int |f(x) + \lambda g(x)|^2 \, dx \\
&= \int (f(x) + \lambda g(x))\overline{(f(x) + \lambda g(x))} \, dx \\
&= \int f(x)\overline{f(x)} \, dx + \int f(x)\overline{\lambda g(x)} \, dx \\
&\quad + \int \lambda g(x)\overline{f(x)} \, dx + \int \lambda g(x)\overline{\lambda g(x)} \, dx \\
&= \int |f(x)|^2 \, dx + \overline{\lambda} \int f(x)\overline{g(x)} \, dx \\
&\quad + \lambda \int g(x)\overline{f(x)} \, dx + \lambda\overline{\lambda} \int g(x)\overline{g(x)} \, dx \\
&= \int |f(x)|^2 \, dx + \overline{\lambda} \int f(x)\overline{g(x)} \, dx \\
&\quad + \lambda \int g(x)\overline{f(x)} \, dx + |\lambda|^2 \int |g(x)|^2 \, dx \\
&= \int |f(x)|^2 \, dx + \overline{\lambda} \int f(x)\overline{g(x)} \, dx \\
&\quad + \lambda \overline{\int f(x)\overline{g(x)} \, dx} + |\lambda|^2 \int |g(x)|^2 \, dx.
\end{aligned}
$$

Since this inequality is true for any complex number λ, it must hold for

$$\lambda = -\frac{\int f(x)\overline{g(x)} \, dx}{\int |g(x)|^2 \, dx}.$$

Substituting this value in for λ gives

$$
\begin{aligned}
0 &\leq \int |f(x)|^2 \, dx - \overline{\frac{\int f(x)\overline{g(x)} \, dx}{\int |g(x)|^2 \, dx}} \int f(x)\overline{g(x)} \, dx \\
&\quad - \frac{\int f(x)\overline{g(x)} \, dx}{\int |g(x)|^2 \, dx} \overline{\int f(x)\overline{g(x)} \, dx} + \frac{|\int f(x)\overline{g(x)} \, dx|^2}{(\int |g(x)|^2 \, dx)^2} \int |g(x)|^2 \, dx \\
&= \int |f(x)|^2 \, dx - \frac{|\int f(x)\overline{g(x)} \, dx|^2}{\int |g(x)|^2 \, dx},
\end{aligned}
$$

from which the desired inequality is an immediate consequence.

Here is a little stuff from probability theory.

If p is a nonnegative function on $\mathbb{R}$ for which $\int_{\mathbb{R}} p(x) \, dx = 1$, we call p a *probability density function.* This density function p is associated to a "random variable" X, in the sense that we say that the probability that the value of the random variable X is between points a and b is equal to $\int_a^b p(x) \, dx$.

The number (if it is finite) $\mu = E(X) = \int_{\mathbb{R}} xp(x) \, dx$ is called the *expected value* or *mean* of the random variable X. Finally, the number (if it's finite) defined by $\sigma^2 = \mathrm{var}(X) = \int_{\mathbb{R}} (x - \mu)^2 p(x) \, dx$ is called the *variance* of the random variable X.

The mean of a random variable is what we think of as the average value, with respect to the probability density function. The variance is a measure of how concentrated the probability is near the mean. If the variance is small, then the density function $p(x)$ must be concentrated near μ; i.e., everywhere the quantity $(x-\mu)^2$ is large, the quantity $p(x)$ must be small. The smaller the variance, the more concentrated the random variable is near its mean value.

The assertion of the next theorem is that the variances of two particular random variables cannot both be small. Their product is always greater than a fixed constant.

Let f be an element of $L^2(\mathbb{R})$ for which $\|f\|^2 = \int_{\mathbb{R}} |f(x)|^2\,dx = 1$. Think of $|f(x)|^2$ as a density function associated to a random variable X. By the Plancherel Theorem, we know that the function $|\widehat{f}(\omega)|^2$ is the density function for another random variable that we denote by P. Suppose that the expected value $\mu = E(X) = \int_{\mathbb{R}} x|f(x)|^2\,dx$ exists, and suppose also that the expected value $\nu = E(P) = \int_{\mathbb{R}} \omega|\widehat{f}(\omega)|^2\,d\omega$ exists.

THEOREM 10.8. *Let the notation be as in the preceding paragraph. Then*

$$var(X) \times var(P) \geq \frac{1}{16\pi^2}.$$

That is,

$$\int_{\mathbb{R}} (x-\mu)^2|f(x)|^2\,dx \times \int_{\mathbb{R}} (\omega-\nu)^2|\widehat{f}(\omega)|^2\,d\omega \geq \frac{1}{16\pi^2}.$$

REMARK. What this theorem has to do with the physical understanding of the Heisenberg uncertainty principle is not immediately obvious. Ask your Quantum Mechanics teacher. In any case, this is kind of a peculiar relation between a function and its Fourier transform.

PROOF. We make a few simplifications before we do any computations. First of all, if $g(x)$ is defined to be $f(x+\mu)$, and x_1 is a random variable associated to the density function $|g(x)|^2$, then

$$\begin{aligned}
E(x_1) &= \int_{\mathbb{R}} x|g(x)|^2\,dx \\
&= \int_{\mathbb{R}} x|f(x+\mu)|^2\,dx \\
&= \int_{\mathbb{R}} (x-\mu)|f(x)|^2\,dx \\
&= \int_{\mathbb{R}} x|f(x)|^2\,dx - \mu\int_{\mathbb{R}} |f(x)|^2\,dx \\
&= \mu - \mu \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{var}(x_1) &= \int_{\mathbb{R}} (x-0)^2 |g(x)|^2 \, dx \\
&= \int_{\mathbb{R}} x^2 |f(x+\mu)|^2 \, dx \\
&= \int_{\mathbb{R}} (x-\mu)^2 |f(x)|^2 \, dx \\
&= \operatorname{var}(X).
\end{aligned}$$

Hence, we need only prove that

$$\operatorname{var}(x_1) \times \operatorname{var}(P) \geq \frac{1}{16\pi^2}.$$

Moreover, since $|\widehat{g}(\omega)| = |\widehat{f}(\omega)|$, the two density functions $|\widehat{g}(\omega)|^2$ and $|\widehat{f}(\omega)|^2$ are identical, which means that the random variable P is associated to the density function $|\widehat{g}(\omega)|^2$.

Secondly, if $h(x)$ is defined by $h(x) = e^{2\pi i \nu x} g(x)$, then $|h(x)| = |g(x)|$, so that the random variable X_1 is associated to the density function $|h(x)|^2$. Also, $\widehat{h}(\omega) = \widehat{g}(\omega + \nu)$, so that just as above, the random variable P_1 associated to the density function $\widehat{h}(\omega)|^2$ satisfies $E(P_1) = 0$, and $\operatorname{var}(P_1) = \operatorname{var}(P)$. Hence, we need only prove that

$$\operatorname{var}(X_1) \times \operatorname{var}(P_1) \geq \frac{1}{16\pi^2}.$$

Now we just compute, using what we know about Fourier transforms and, at a crucial point, using the Cauchy-Schwarz Inequality. Also, for notational purposes, we will write $H(x)$ for the function $xh(x)$. See if you can follow this!

$$\begin{aligned}
\operatorname{var}(X_1) \times \operatorname{var}(P_1) &= \int_{\mathbb{R}} x^2 |h(x)|^2 \, dx \times \int_{\mathbb{R}} \omega^2 |widehath(\omega)|^2 \, d\omega \\
&= \int_{\mathbb{R}} x^2 |h(x)|^2 \, dx \frac{1}{4\pi^2} \int_{\mathbb{R}} |\widehat{h'}(\omega)|^2 \, d\omega \\
&= \frac{1}{4\pi^2} \|H\|^2 \|h'\|^2 \\
&\geq \frac{1}{4\pi^2} |\langle H \mid h' \rangle|^2 \\
&= \frac{1}{16\pi^2} (|\langle H \mid h' \rangle| + |\langle h' \mid H \rangle|)^2 \\
&= \frac{1}{16\pi^2} | \int_{\mathbb{R}} H(x)\overline{h'(x)} + h'(x)\overline{H(x)} \, dx|^2 \\
&= \frac{1}{16\pi^2} | \int_{\mathbb{R}} xh(x)\overline{h'(x)} + xh'(x)\overline{h(x)} \, dx|^2 \\
&= \frac{1}{16\pi^2} | \int_{\mathbb{R}} x(h(x)\overline{h'(x)} + h'(x)\overline{h(x)}) \, dx|^2 \\
&= \frac{1}{16\pi^2} (xh(x)\overline{h(x)}[_{-\infty}^{\infty} - \int_{\mathbb{R}} h(x)\overline{h(x)} \, dx|^2 \\
&= \frac{1}{16\pi^2} | \int_{\mathbb{R}} |h(x)|^2 \, dx|^2 \\
&= \frac{1}{16\pi^2},
\end{aligned}$$

completing the proof.

EXERCISE 10.11. Show that the Heisenberg uncertainty inequality above is actually an equality if $f(x) = e^{-\pi x^2}$.