A Primer on the Fourier Transform

Evan Randles

In this note, I will introduce the Fourier transform on $\mathbb{R}$ and derive a number of its amazing properties. I have sprinkled in a few exercises for your benefit as you actively read this document; you needn't turn them in. If you have any questions or comments, please let me know.

1 The Schwartz Class $\mathcal{S}$

The Fourier transform can be seen as a linear transformation between function (vector) spaces. In this section, we introduce a particularly nice function space, called the Schwartz class of functions, on which we will define the Fourier transform. The following section is dedicated to introducing the Fourier transform and proving that, in fact, it is an automorphism[1] of the Schwartz class.

We shall denote by $C^\infty(\mathbb{R})$ the set of smooth (infinitely differentiable) complex-valued functions[2]. In other words, each $f \in C^\infty(\mathbb{R})$ is of the form

$$f(x) = u(x) + iv(x)$$

defined for $x \in \mathbb{R}$ where $u(x)$ and $v(x)$ are in infinitely differentiable real-valued functions called the real and imaginary parts of f, respectively; these are also written $\mathrm{Re}(f) = u$ and $\mathrm{Im}(f) = v$. The complex conjugate of f is the function $\overline{f}$ defined by

$$\overline{f}(x) = \overline{f(x)} = u(x) - iv(x)$$

for $x \in \mathbb{R}$ and the absolute value[3] is the real-valued (non-negative-valued) function $|f|$ given by

$$|f|(x) = |f(x)| = \sqrt{f(x)\overline{f}(x)} = \sqrt{(u(x))^2 + (v(x))^2}$$

for $x \in \mathbb{R}$. Note that the final equality holds because

$$\begin{aligned}
f(x)\overline{f(x)} &= (u(x) + iv(x))(u(x) - iv(x)) \\
&= ((u(x))^2 - iv(x)u(x) + iv(x)u(x) - (i)^2(v(x))^2 \\
&= (u(x))^2 + 0 - (-1)(v(x))^2 \\
&= (u(x))^2 + (v(x))^2
\end{aligned}$$

for $x \in \mathbb{R}$. For each $n \in \mathbb{N} =: \{0, 1, 2, \dots\}$ and $f = u + iv \in C^\infty(\mathbb{R})$, we will write $f^{(n)}$ for the nth-derivative of f; this is defined by

$$f^{(n)}(x) = \frac{d^n f}{dx^n}(x) = \frac{d^n u}{dx^n}(x) + i\frac{d^n v}{dx^n}(x)$$

[1]Given a vector space V, an automorphism of V is an invertible linear transformation, i.e., an isomorphism, from V onto itself.

[2]To be more explicit about the codomain of such functions, we could write $C^\infty(\mathbb{R};\mathbb{C})$ instead of $C^\infty(\mathbb{R})$, but we won't.

[3]This is also commonly called the *complex modulus* of f.

for $x \in \mathbb{R}$.

The following example of an element in $C^\infty(\mathbb{R})$ is paramount: Given any number $\xi \in \mathbb{R}$, consider the function

$$e^{ix\xi} = \cos(x\xi) + i\sin(x\xi)$$

defined for $x \in \mathbb{R}$. We observe that $x \mapsto e^{ix\xi}$ is a member of $C^\infty(\mathbb{R})$ because $x \mapsto \cos(x\xi)$ and $x \mapsto \sin(x\xi)$ are both infinitely differentiable. We see that

$$\overline{e^{ix\xi}} = \cos(x\xi) - i\sin(x\xi) = \cos(-x\xi) + i\sin(-x\xi) = e^{-ix\xi}$$

for all $x \in \mathbb{R}$ where we have used the fact that cosine is an even function and sine is an odd function. Also,

$$|e^{ix\xi}| = \sqrt{\cos^2(x\xi) + \sin^2(x\xi)} = \sqrt{1} = 1$$

for all $x \in \mathbb{R}$, in view of the Pythagorean identity $\cos^2(\theta) + \sin^2(\theta) = 1$. Observe that

$$\begin{aligned}
\frac{d}{dx}e^{ix\xi} &= \frac{d}{dx}\cos(x\xi) + i\frac{d}{dx}\sin(x\xi) \\
&= -\xi\sin(x\xi) + i\xi\cos(x\xi) \\
&= i\xi\cos(x\xi) + (i\cdot i)\xi\sin(x\xi) \\
&= i\xi(\cos(x\xi) + i\sin(x\xi)) \\
&= i\xi e^{ix\xi}
\end{aligned}$$

for $x \in \mathbb{R}$. In fact, you can easily verify that $\frac{d^n}{dx^n}e^{ix\xi} = (i\xi)^n e^{ix\xi}$ for each $n \in \mathbb{N}$.

Exercise 1. *Let $\alpha = a + ib \in \mathbb{C}$ and define*

$$e^{\alpha x} = e^{ax}\cos(bx) + ie^{ax}\sin(bx)$$

for $x \in \mathbb{R}$. Then:

1. *$e^{\alpha x} \in C^\infty(\mathbb{R})$.*
2. *$|e^{\alpha x}| = e^{ax}$ for $x \in \mathbb{R}$.*
3. *$\frac{d^n}{dx^n}e^{\alpha x} = (\alpha)^n e^{\alpha x}$ for $x \in \mathbb{R}$.*
4. *If $\alpha_1 = a_1 + ib_1$ and $\alpha_2 = a_2 + ib_2$, then $e^{\alpha_1 x}e^{\alpha_2 x} = e^{(\alpha_1+\alpha_2)x}$ for $x \in \mathbb{R}$.*

The following proposition may be familiar to you from MA311.

Proposition 1. *The set of smooth functions $C^\infty(\mathbb{R})$, equipped with the usual function addition and scalar multiplication, is a vector space over $\mathbb{C}$.*

Exercise 2. *Prove the proposition.*

Now that we are familiar with $C^\infty(\mathbb{R})$, we introduce an important subspace with which we will work (almost) exclusively henceforth. Named after the French mathematician, Laurent Schwartz, the *Schwartz class* is the set $\mathcal{S} = \mathcal{S}(\mathbb{R})$ consisting of smooth functions $f \in C^\infty(\mathbb{R})$ which satisfy the following property: For each pair of natural numbers $k, n \in \mathbb{N}$, there is a positive number $M = M_{k,n,f} > 0$ for which

$$|x^k f^{(n)}(x)| \leq M \tag{1}$$

for all $x \in \mathbb{R}$. This condition says that, f and all of its derivatives decay faster than any rational function at $\pm\infty$. To be explicit,

$$\mathcal{S} = \left\{ f \in C^\infty(\mathbb{R}) : \forall k, n \in \mathbb{N}, \exists\, M > 0 \text{ such that } |x^k f^{(n)}(x)| \leq M \,\forall x \in \mathbb{R} \right\}.$$

The following is a useful characterization of $\mathcal{S}$.

Proposition 2. *A function $f \in \mathcal{S}$ if and only if, $f \in C^\infty(\mathbb{R})$ and, for all $k, n \in \mathbb{N}$,*

$$\lim_{x \to \pm\infty} |x^k f^{(n)}(x)| = 0. \tag{2}$$

Proof. Let $f \in \mathcal{S}$ and let $k, n \in \mathbb{N}$. In view of (1), there is a positive number M for which

$$|x||x^k f^{(n)}(x)| = |x^{k+1} f^{(n)}(x)| \leq M$$

for all $x \in \mathbb{R}$. Consequently,

$$0 \leq |x^k f^{(n)}(x)| \leq \frac{M}{|x|}$$

for all $x \in \mathbb{R}$ and so, by the squeeze theorem,

$$\lim_{x \to \pm\infty} |x^k f^{(k)}(x)| = \lim_{x \to \pm\infty} \frac{M}{|x|} = 0.$$

Conversely, assume that (2) holds and let $k, n \in \mathbb{N}$. Because $|x^k f^{(n)}(x)| \to 0$ and $x \to \pm\infty$, there is some number N for which

$$|x^k f^{(n)}(x)| \leq 1$$

whenever $|x| > N$ (you should think about why this is true). Since $x \mapsto x^n f^{(n)}(x)$ is a continuous function on $\mathbb{R}$, it is necessarily bounded on bounded interval $[-N, N]$ and we write

$$\widetilde{M} = \max_{-N \leq x \leq N} |x^k f^{(n)}(x)|.$$

Thus, by setting $M = \max\{1, \widetilde{M}\}$, we have

$$|x^k f^{(n)}(x)| \leq M$$

for all $x \in \mathbb{R}$ and so (1) holds. □

In view of the preceding proposition,

$$S = \left\{ f \in C^\infty(\mathbb{R}) : \forall k, n \in \mathbb{N}, \lim_{x \to \pm\infty} |x^k f^{(n)}(x)| = 0 \right\}.$$

So far, we've studied $\mathcal{S}$ but we haven't given a single example of a member of $\mathcal{S}$ (except the boring example of the zero function). The following definition gives a one-parameter family of functions in $\mathcal{S}$ which, as it turns out, is one of the most useful and important functions from the perspective of analysis and partial differential equations.

Definition 3. *The heat kernel on $\mathbb{R}$ is the function $G_{(\cdot)}(\cdot) : (0, \infty) \times \mathbb{R} \to \mathbb{R}$ given by*

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

for $t > 0$ and $x \in \mathbb{R}$.

Proposition 4. *For each $t > 0$, $x \mapsto G_t(x)$ is a member of $\mathcal{S}$. Further, it is an approximation of the identity in the sense that, for any continuous and bounded function[4] $f : \mathbb{R} \to \mathbb{C}$, we have*

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

for each $x \in \mathbb{R}$.

[4]This assumption can be significantly weakened.

Proof. We first verify that, for each $t > 0$, $G_t \in \mathcal{S}$. To this end, observe that, for each $k, n \in \mathbb{N}$, the chain and product rules ensure that

$$x^k G_t^{(n)}(x) = p(x) e^{-\frac{x^2}{4t}} = \frac{p(x)}{e^{\frac{x^2}{4t}}}$$

for $x \in \mathbb{R}$ where $p(x) = p_{n,k,t}(x)$ is a polynomial in x with degree at most $k + n$ with coefficients depending on t. It follows by L'Hôpital's rule (or the definition of $e^{x^2/4t}$ in terms of its Maclaurin series) that

$$\lim_{x \to \pm\infty} |x^k G_t^{(n)}(x)| = \lim_{x \to \pm\infty} \frac{|p(x)|}{|e^{\frac{x^2}{4t}}|} = 0$$

and so $G_t \in \mathcal{S}$ in view of the preceding proposition. To see that G_t forms an approximation to the identity we first make two observations.

1. For each $t > 0$, we can make the change of variables $\tilde{y} = x - y$ to see that

   $$\begin{aligned} \int_{\mathbb{R}} G_t(x-y) f(y)\, dy &= \int_{-\infty}^{\infty} G_t(x-y) f(y)\, dy \\ &= \int_{-\infty}^{\infty} G_t(\tilde{y}) f(x-\tilde{y})(-d\tilde{y}) = \int_{-\infty}^{\infty} G_t(\tilde{y}) f(x-\tilde{y})\, d\tilde{y} \end{aligned}$$

   and therefore

   $$\int_{\mathbb{R}} G_t(x-y) f(y)\, y = \int_{\mathbb{R}} G_t(y) f(x-y)\, dy.$$

   As we discussed, the above is a convolution of G_t with f and this property shows that the convolution product is commutative.

2. For each $t > 0$, we make the change of variables $x = y/\sqrt{4t}$ to see that

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

   where I have made use of Lemma 15 to compute the last integral.

With these two observations in hand, let's now show that

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

or, equivalently,

$$\lim_{t \to 0} D_t(x) = 0$$

for each x where

$$D_t(x) := \left| \int_{\mathbb{R}} G_t(x-y) f(y)\, dy - f(x) \right|$$

To this end, fix $x \in \mathbb{R}$ and $\epsilon > 0$. Given that f is continuous at x, there is a $\delta > 0$ for which $|f(x-y) - f(x)| < \epsilon/2$ for all $|y| < \delta$. Also, since f is bounded (say by M), $|f(x-y) - f(x)| \leq |f(x-y)| + |f(x)| \leq 2M$ in view of the triangle inequality. Using both observations above,

$$\begin{aligned}
D_t(x) &= \left|\int_{\mathbb{R}} G_t(x-y)f(y)\,dy - f(x)\right| \\
&= \left|\int_{\mathbb{R}} G_t(y)f(x-y)\,dy - f(x)\int_{\mathbb{R}} G_t(y)\,dy\right| \\
&= \left|\int_{\mathbb{R}} G_t(y)(f(x-y)-f(y))\,dy\right| \\
&\leq \int_{\mathbb{R}} G_t(y)\,|f(x-y)-f(y)|\,dy \\
&= \int_{-\delta}^{\delta} G_t(y)\,|f(x-y)-f(y)|\,dy + \int_{|y|>\delta} G_t(y)\,|f(x-y)-f(x)|\,dy \\
&\leq \frac{\epsilon}{2}\int_{-\delta}^{\delta} G_t(y)\,dy + 2M\int_{|y|>\delta} G_t(y)\,dy
\end{aligned}$$

Since $G_t(y) \geq 0$,

$$\int_{-\delta}^{\delta} G_t(y)\,dy \leq \int_{\mathbb{R}} G_t(y)\,dy = 1.$$

Also, using the change of variables $\tilde{y} = y/\sqrt{4t}$, we have

$$\int_{|y|>\delta} G_t(y)\,dy = \int_{|\tilde{y}|>\delta/\sqrt{4t}} e^{-\tilde{y}^2}\,d\tilde{y}$$

and since

$$\lim_{t\to 0}\frac{\delta}{\sqrt{4t}} = \infty,$$

from the above equation we see that

$$\lim_{t\to 0}\int_{|y|>\delta} G_t(y)\,dy = 0$$

or, equivalently, there is some $t_0 > 0$ for which the

$$\int_{|y|>\delta} G_t(y)\,dy < \epsilon/4M$$

whenever $t < t_0$. Putting these together, we see that

$$D_t(x) \leq \frac{\epsilon}{2} + 2M\frac{\epsilon}{4M} = \epsilon$$

whenever $t < t_0$. Consequently, $\lim_{t\to 0} D_t(x) = 0$ as claimed.

□

Exercise 3. *In this exercise, you will explore some functions in $\mathcal{S}$ and some functions not in $\mathcal{S}$.*

1. *Show that $x \mapsto \frac{1}{x^2+1}$, while smooth, is not a member of $\mathcal{S}$.*
2. *Generalizing the above, show that, if $p(x)$ and $q(x)$ are polynomials where p is not the zero polynomial and $q(x) \neq 0$ for all $x \in \mathbb{R}$, then $x \mapsto p(x)/q(x)$ is smooth but not a member of $\mathcal{S}$.*

3. *We say that a function $f : \mathbb{R} \to \mathbb{R}$ is compactly supported if there is $K > 0$ for which $f(x) = 0$ for all $|x| > K$ and we denote by $C_c^\infty(\mathbb{R})$ the collection of function $f \in C^\infty(\mathbb{R})$ which are compactly supported. Show that $C_c^\infty(\mathbb{R}) \subseteq \mathcal{S}$.*

Exercise 4. *A collection of function g_t for $t > 0$ is said to be an approximation to the identity (in a pointwise sense) if, for any continuous and bounded function f,*

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

for each $x \in \mathbb{R}$.. Let $g : \mathbb{R}$ be an improperly Riemann integrable function, i.e, $\int_{\mathbb{R}} |g|(x)\, dx < \infty$ (as an improper Riemann integral) and such that

$$\int_{\mathbb{R}} g(x) = 1.$$

Show that

$$g_t(x) = \frac{1}{t} g(x/t)$$

for $t > 0$ is an approximate identity. Note: This is essentially what we proved for the heat kernel (but where t was replaced by $\sqrt{4t}$) – all of the arguments go through without issue.

We could go on and on and on about the Schwartz class[5], but we won't. The following proposition is all we really need.

Proposition 5. *The Schwartz class $\mathcal{S}$, equipped with the usual definition of function addition and scalar multiplication, is a vector space over $\mathbb{C}$.*

Proof. As subspaces of vector spaces are vector spaces (over the same field) in their own right, it suffices to show that $\mathcal{S}$ is a subspace of $C^\infty(\mathbb{R})$. Since we've already shown that $\mathcal{S}$ is non-empty (e.g., $G_1 \in \mathcal{S}$)), it suffices to show that $\mathcal{S}$ is closed under linear combinations. To this end, let $f, g \in \mathcal{S}$ and let $\alpha, \beta \in \mathbb{C}$. Given any $k, n \in \mathbb{N}$, standard rules of calculus guarantee that

$$x^k(\alpha f + \beta g)^{(n)}(x) = \alpha x^k f^{(n)}(x) + \beta x^k g^{(n)}(x)$$

for $x \in \mathbb{R}$. Consequently,

$$\lim_{x \to \pm\infty} |x^k(\alpha f + \beta g)^{(n)}(x)| = \lim_{x \to \pm\infty} |x^k f^{(n)}(x)| + \lim_{x \to \pm\infty} |x^k g^{(n)}(x)| = 0 + 0 = 0$$

and therefore $\alpha f + \beta g \in \mathcal{S}$. $\square$

2 The Fourier Transform on $\mathcal{S}$

Given $f \in \mathcal{S}$, define[6] $\mathcal{F}(f) = \widehat{f} : \mathbb{R} \to \mathbb{C}$ by

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{ix\xi}\, dx$$

for $\xi \in \mathbb{R}$. This function is called the *Fourier transform of f*. By the end of this section, we will prove the following theorem.

[5] If you're interested, $\mathcal{S}$ is a topological vector space (its topology is given by a separating family of seminorms and it is meterizable). It is also a Frèchet space.

[6] In this note, our definition of the integral should be taken as an improper Riemann integral $\int_{\mathbb{R}} f(x)\, dx := \lim_{t \to \infty} \int_{-t}^{t} f(x)\, dx$ (which you will explore in the homework). As all functions in sight, for us, are continuous, the only possible impediment to such an integral converging/existing is that the limit above might not exist and this happens when $|f|$ contains an infinite amount of area under its graph, e.g., $\int_{\mathbb{R}} x^2\, dx$ does not converge.

Theorem 6. *The Fourier transform is an automorphism of $\mathcal{S}$ (i.e., $\mathcal{F}$ is an invertible linear transformation from $\mathcal{S}$ to itself). Its inverse is given by*

$$\mathcal{F}^{-1}(h)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} h(\xi)e^{-i\xi x}\,d\xi$$

for $x \in \mathbb{R}$.

In the course of proving this theorem, we will introduce and prove many beautiful and amazing properties of the Fourier transform. We begin with the following proposition which shows that, in particular, the Fourier transform is linear.

Proposition 7. *Given $f, g \in \mathcal{S}$ and $\alpha, \beta \in \mathbb{C}$,*

1. *For every $\xi \in \mathbb{R}$,*
$$\mathcal{F}(\alpha f + \beta g)(\xi) = \alpha\mathcal{F}(f)(\xi) + \beta\mathcal{F}(g)(\xi).$$

2. *For every $a > 0$,*
$$\mathcal{F}(f(x/a))(\xi) = a\mathcal{F}(f)(a\xi)$$
for all $\xi \in \mathbb{R}$.

3. *Given $x_0 \in \mathbb{R}$, denote τ_{x_0} the "shift" operator for which $(\tau_{x_0}f)(x) = f(x - x_0)$ for $x \in \mathbb{R}$. Then*
$$\mathcal{F}(\tau_{x_0}f)(x) = e^{ix_0\xi}\widehat{f}(\xi) = e^{ix_0\xi}\mathcal{F}(\xi)$$
for all $\xi \in \mathbb{R}$.

Proof. The first item is a standard consequence of the linearity of the integral. To see the second, we compute

$$\mathcal{F}(f(x/a))(\xi) = \int_{\mathbb{R}} f\left(\frac{x}{a}\right)e^{ix\xi}\,dx.$$

By making a change of variables[7] $\tilde{x} = x/a$ and so $ad\tilde{x} = dx$ we have

$$\mathcal{F}(f(x/a))(\xi) = \int_{\mathbb{R}} f(\tilde{x})e^{ia\tilde{x}\xi}a\,d\tilde{x} = a\int_{\mathbb{R}} f(\tilde{x})e^{i\tilde{x}(a\xi)}\,d\tilde{x} = a\widehat{f}(a\xi)$$

for $\xi \in \mathbb{R}$. To see the third property, we makes the change of variables $\tilde{x} = x - x_0$ (with $x = x_0 + \tilde{x}$ and $dx = d\tilde{x}$) to see that

$$\begin{aligned}
\mathcal{F}(\tau_{x_0}(f))(\xi) &= \int_{\mathbb{R}} f(x - x_0)e^{ix\xi}\,dx \\
&= \int_{\mathbb{R}} f(\tilde{x})e^{i(x_0+\tilde{x})\xi}\,d\tilde{x} \\
&= \int_{\mathbb{R}} f(\tilde{x})e^{ix_0\xi}e^{i\tilde{x}\xi}\,d\tilde{x} \\
&= e^{ix_0\xi}\int_{\mathbb{R}} f(\tilde{x})e^{i\tilde{x}\xi}\,d\tilde{x} \\
&= e^{ix_0\xi}\widehat{f}(\xi)
\end{aligned}$$

for $\xi \in \mathbb{R}$, as desired. □

[7] Truthfully, to verify that this change of variables can be done (and everything converges), one should invoke the definition of the improper Riemann integral.

With the preceding result at our fingertips, let's work out an important example.

Proposition 8. *We have*

$$\mathcal{F}(e^{-x^2})(\xi) = \widehat{e^{-x^2}}(\xi) = \sqrt{\pi}e^{-\xi^2/4}$$

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

Proof. The proof presented here is essentially that which appears in [2]. Define $g : \mathbb{R} \to \mathbb{C}$ by

$$g(s) = \int_{\mathbb{R}} e^{-(x-is)^2}\,dx$$

for $s \in \mathbb{R}$. It can be verified that the improper Riemann integral converges absolutely for each value of $s \in \mathbb{R}$ and so this expression makes sense. In fact, because the Gaussian $x \mapsto e^{-x^2}$ is such a well-behaved function, the function g is differentiable for all $s \in \mathbb{R}$ and one can differentiate through the integral sign[8] to see that

$$g'(s) = \frac{d}{ds}\int_{\mathbb{R}} e^{-(x-is)^2}\,dx = \int_{\mathbb{R}} \frac{d}{ds}e^{-(x-is)^2}\,dx = \int_{\mathbb{R}} 2i(x-is)e^{-(x-is)^2}\,dx$$

for each $s \in \mathbb{R}$ where we have made use of the chain rule. We note that

$$\frac{d}{dx}e^{-(x-is)^2} = -2(x-is)e^{-(x-is)^2}$$

and therefore

$$g'(s) = -i\int_{\mathbb{R}} \frac{d}{dx}\left(e^{-(x-is)^2}\right)\,dx$$

Using the definition of the above improper integral and the fundamental theorem of calculus, we have

$$\begin{aligned}
\int_{\mathbb{R}} \frac{d}{dx}\left(e^{-(x-is)^2}\right)\,dx &= \lim_{t\to\infty}\int_{-t}^{t} \frac{d}{dx}\left(e^{-(x-is)^2}\right)\,dx \\
&= \lim_{t\to\infty} e^{-(x-is)^2}\Big|_{x=-t}^{x=t} \\
&= \lim_{t\to\infty}\left(e^{-(t-is)^2} - e^{-(-t-is)^2}\right) \\
&= \lim_{t\to\infty}\left(e^{-t^2+2its-i^2s^2} - e^{-(-t)^2-2its-i^2s^2}\right) \\
&= \lim_{t\to\infty} e^{-t^2}e^{s^2}\left(e^{2its} - e^{-2its}\right) \\
&= e^{s^2}\lim_{t\to\infty} e^{-t^2}2i\sin(2ts) = 0
\end{aligned}$$

where we have used the identity $\sin(\theta) = (e^{i\theta} - e^{-i\theta}/)2i$ and the fact that $e^{-t^2} \to 0$ as $t \to \infty$. Consequently,

$$g'(s) = -i\int_{\mathbb{R}} \frac{d}{dx}\left(e^{-(x-is)^2}\right)\,dx = -i0 = 0$$

for each $s \in \mathbb{R}$. By the mean value theorem, it follows that, for any $\xi \in \mathbb{R}$,

$$g(\xi/2) = g(0).$$

[8] To differentiate under an integral sign, one needs to verify that the integrand is sufficiently well behaved and that this differentiation won't affect the convergence of the resultant integral. We won't worry about this – but you should know that there are details to be checked and, if you take a course in measure theory, you will check them.

Fixing $\xi \in \mathbb{R}$, we see that

$$e^{-(x-i\xi/2)^2} = e^{-x^2+ix\xi+\xi^2/4} = e^{\xi^2/4}e^{-x^2}e^{ix\xi}$$

and therefore

$$\begin{aligned} g(\xi/2) &= \int_{\mathbb{R}} e^{-(x-i\xi/2)^2}\,dx \\ &= \int_{\mathbb{R}} e^{\xi^2/4}e^{-x^2}e^{ix\xi}\,dx \\ &= e^{\xi^2/4}\int_{\mathbb{R}} e^{-x^2}e^{ix\xi}\,dx \\ &= e^{\xi^2/4}\mathcal{F}(e^{-x^2})(\xi). \end{aligned}$$

Also, by Lemma 15,

$$g(0) = \int_{\mathbb{R}} e^{-x^2}\,dx = \sqrt{\pi}$$

and so we obtain

$$\sqrt{\pi} = g(0) = g(\xi/2) = e^{\xi^2/2}\mathcal{F}(e^{-x^2})(\xi)$$

and, equivalently,

$$\mathcal{F}(e^{-x^2})(\xi) = \sqrt{\pi}e^{-\xi^2/4}.$$

□

Corollary 9. *For $t > 0$, define $G_t : \mathbb{R} \to \mathbb{R}$ by*

$$G_t(x) = \frac{1}{\sqrt{4\pi t}}\exp\left(-\frac{x^2}{4t}\right)$$

for $x \in \mathbb{R}$. Then, for each $t > 0$,

$$\mathcal{F}(G_t)(\xi) = \widehat{G_t}(\xi) = e^{-t\xi^2}$$

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

Proof. Let $t > 0$ and observe that, by virtue of Proposition 7,

$$\mathcal{F}(G_t)(\xi) = \frac{1}{\sqrt{4\pi t}}\mathcal{F}(e^{-(x/a)^2})(\xi) = \frac{a}{\sqrt{4\pi t}}\mathcal{F}(e^{-x^2})(a\xi) = \frac{1}{\sqrt{\pi}}\mathcal{F}(e^{-x^2})(a\xi)$$

for all $\xi \in \mathbb{R}$ where $a = \sqrt{4t}$. By invoking the preceding proposition, we find

$$\mathcal{F}(G_t)(\xi) = \frac{1}{\sqrt{\pi}}\sqrt{\pi}\exp\left(\frac{-(a\xi)^2}{4}\right) = \exp\left(\frac{-\left(\sqrt{4t}\right)^2\xi^2}{4}\right) = e^{-t\xi^2}$$

for $\xi \in \mathbb{R}$, as desired. □

The following proposition shows that the Fourier transform exchanges differentiation and polynomial multiplication.

Proposition 10. *For any $f \in \mathcal{S}$ and $k, n \in \mathbb{N}$,*

$$\mathcal{F}(x^k f^{(n)})(\xi) = (-i)^{n+k}\frac{d^k}{d\xi^k}\left(\xi^n\widehat{f}(\xi)\right)$$

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

Proof. Because $(-i)(i) = 1$, we have

$$\begin{aligned}
\mathcal{F}(x^k f^{(n)})(\xi) &= \int_{\mathbb{R}} x^k f^{(n)}(x) e^{ix\xi}\, dx \\
&= \int_{\mathbb{R}} (-i)^k f^{(n)} \left((ix)^k e^{ix\xi}\right)\, dx \\
&= (-i)^k \int_{\mathbb{R}} f^{(n)}(x) \frac{d^k}{d\xi^k}\left(e^{ix\xi}\right)\, dx \\
&= (-i)^k \int_{\mathbb{R}} \frac{d^k}{d\xi^k}\left(f^{(n)}(x) e^{ix\xi}\right)\, dx \\
&= (-i)^k \frac{d^k}{d\xi^k} \int_{\mathbb{R}} f^{(n)}(x) e^{ix\xi}\, dx
\end{aligned}$$

where we have moved the derivatives in ξ, which is not the variable of integration, through the integral sign; this move is justfied because f and its derivatives are smooth and rapidly decreasing at infinity [1, Exercise 5.R]. We now focus on the above integral. A single application of integration by parts yields

$$\begin{aligned}
\int_{\mathbb{R}} f^{(n)}(x) e^{ix\xi}\, dx &= \int_{\mathbb{R}} \frac{d}{dx}(f^{(n-1)}(x)) e^{ix\xi}\, dx \\
&= -\int_{\mathbb{R}} f^{(n-1)}(x) \frac{d}{dx}\left(e^{ix\xi}\right)\, dx \\
&= -\int_{\mathbb{R}} f^{(n-1)}(x)(i\xi) e^{ix\xi}\, dx \\
&= (-i\xi) \int_{\mathbb{R}} f^{(n-1)}(x) e^{ix\xi}\, dx
\end{aligned}$$

for $\xi \in \mathbb{R}$. By repeating this computation $n-1$ more times, we obtain

$$\int_{\mathbb{R}} f^{(n)}(x)\, dx = (-i\xi)^n \int_{\mathbb{R}} f^{(n-n)}(x) e^{ix\xi}\, dx = (-i\xi)^n \int_{\mathbb{R}} f(x) e^{ix\xi}\, dx = (-i\xi)^n \widehat{f}(\xi)$$

for $\xi \in \mathbb{R}$. Consequently,

$$\mathcal{F}\left(x^k f^{(n)}\right)(\xi) = (-i)^k \frac{d^k}{d\xi^k}\left((-i\xi)^n \widehat{f}(\xi)\right) = (-i)^{k+n} \frac{d^k}{d\xi^k}\left(\xi^n \widehat{f}(\xi)\right)$$

for $\xi \in \mathbb{R}$. □

We defined the Fourier transform $\mathcal{F}$ on the Schwartz class $\mathcal{S}$ and established that it was linear (Proposition 7). To conclude that $\mathcal{F}$ is truly a linear map, we have to identify a vector space to which $\mathcal{F}$ maps. The following proposition, whose proof makes use of the preceding proposition, shows that $\mathcal{F}$ is a linear map from the Schwartz class into itself (and so it is a so-called endomorphism). This proposition is an important step to proving Theorem 6 which says that, in fact, $\mathcal{F}$ is an invertible linear transformation from $\mathcal{S}$ onto itself (and so it is called an automorphism of $\mathcal{S}$).

Proposition 11. *The Fourier transform $\mathcal{F}$ is a linear transformation from $\mathcal{S}$ into itself and so we may write $\mathcal{F} : \mathcal{S} \to \mathcal{S}$.*

Before proving the proposition, we treat a technical lemma which is related to the generalized Leibniz (product) rule. In the statement and proof of the lemma, we use the symbol D to indicate the derivative operator $D(f) = f'$ and we observe that $D^j(f) = f^{(j)}$ for every $j \in \mathbb{N}$. Also, for each pair of integers $0 \le j \le k$,

$$\binom{k}{j} = \frac{k!}{j!(k-j)!}.$$

Lemma 12. *For each $k \in \mathbb{N}$,*

$$gf^{(k)} = gD^k f = \sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{k-j} \left(D^j(g) f\right)$$

whenever $g, f \in C^\infty(\mathbb{R})$.

Proof. We shall prove this statement by induction. When $k = 0$, we have

$$gf^{(0)} = gf = D^0(gf) = \sum_{j=0}^{0} D^{k-j} \left(D^j(g) f\right) = \sum_{j=0}^{0} (-1)^j \binom{k}{j} D^{k-j} \left(D^j(g) f\right)$$

$f, g \in C^\infty(\mathbb{R})$ where we have used the fact that $(-1)^0 = 1$ and $\binom{0}{0} = 1$. Thus, the base case holds. We now induct on k. We assume that the statement holds for $k \in \mathbb{N}$ and let $f, g \in C^\infty(\mathbb{R})$. By the product rule,

$$gf^{(k+1)} = gD(f^{(k)}) = D(gf^{(k)}) - g' f^{(k)}. \tag{3}$$

By the inductive hypotheses, we have

$$gf^{(k)} = \sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{k-j} \left(D^j(g) f\right) \tag{4}$$

and

$$\begin{aligned} g' f^{(k)} &= \sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{k-j} \left(D^j(g') f\right) \\ &= \sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{k-j} \left(D^{j+1}(g) f\right) \\ &= \sum_{j=1}^{k+1} (-1)^{j-1} \binom{k}{j-1} D^{k-(j-1)} \left(D^j(g) f\right) \end{aligned} \tag{5}$$

where we have made the change of index $j \mapsto j-1$ to obtain the last equality. By combining (3), (4), and (5), we obtain

$$\begin{aligned} gf^{(k+1)} &= D\left(\sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{k-j} \left(D^j(g) f\right)\right) - \sum_{j=1}^{k+1} (-1)^{j-1} \binom{k}{j-1} D^{k-(j-1)} \left(D^j(g) f\right) \\ &= \sum_{j=0}^{k} (-1)^j \binom{k}{j} D \left(D^{k-j} \left(D^j(g) f\right)\right) + \sum_{j=1}^{k+1} (-1)^j \binom{k}{j-1} D^{k-(j-1)} \left(D^j(g) f\right) \\ &= \sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{k+1-j} \left(D^j(g) f\right) + \sum_{j=1}^{k+1} (-1)^j \binom{k}{j-1} D^{k+1-j} \left(D^j(g) f\right) \end{aligned}$$

where we have used the fact that D is linear. With the aim combining the summations, we peel off the zeroth term of the first sum and the last term of the second sum to obtain

$$\begin{aligned}
gf^{(k+1)} &= (-1)^0\binom{k}{0}D^{k+1-0}\left(D^0(g)f\right)+\sum_{j=1}^{k}(-1)^j\binom{k}{j}D^{k+1-j}\left(D^j(g)f\right)\\
&\quad +\sum_{j=1}^{k}(-1)^j\binom{k}{j-1}D^{k+1-j}\left(D^j(g)f\right)\\
&\quad\quad +(-1)^{k+1}\binom{k}{k+1-1}D^{k+1-(k+1)}\left(D^{k+1}(g)f\right)\\
&= (-1)^0\binom{k}{0}D^{k+1}\left(D^0(g)f\right)+\sum_{j=1}^{k}(-1)^j\left[\binom{k}{j}+\binom{k}{j-1}\right]D^{k+1-j}\left(D^j(g)f\right)\\
&\quad\quad +(-1)^{k+1}\binom{k}{k}D^{k+1-(k+1)}\left(D^{k+1-0}(g)f\right).
\end{aligned}$$

We obseve that $\binom{k}{0}=1=\binom{k+1}{0}$, $\binom{k}{k}=1=\binom{k+1}{k+1}$, and

$$\begin{aligned}
\binom{k}{j}+\binom{k}{j-1} &= \frac{k!}{j!(k-j)!}+\frac{k!}{(j-1)!(k-(j-1))!}\\
&= \frac{k-(j-1)}{k-(j-1)}\frac{k!}{j!(k-j)!}+\frac{j}{j}\frac{k!}{(j-1)!(k-(j-1))!}\\
&= \frac{(k+1-j)k!+jk!}{j!(k+1-j)!}\\
&= \frac{(k+1)!}{j!(k+1-j)!}\\
&= \binom{k+1}{j}
\end{aligned}$$

for $1\le j\le k$. Thus

$$\begin{aligned}
gf^{(k+1)} &= (-1)^0\binom{k+1}{0}D^{k+1-0}\left(D^0(g)f\right)+\sum_{j=1}^{k}(-1)^j\binom{k+1}{j}D^{k+1-j}\left(D^j(g)f\right)\\
&\quad\quad +(-1)^{k+1}\binom{k+1}{k+1}D^{k+1-(k+1)}\left(D^{k+1}(g)f\right)\\
&= \sum_{j=0}^{k+1}(-1)^j\binom{k+1}{j}D^{k+1-j}\left(D^j(g)f\right)
\end{aligned}$$

which verifies the inductive step. Our proof is complete. □

Proof of Proposition 11. Let $f\in\mathcal{S}$. By virtue of Proposition 10, $\widehat{f}$ is infinitely differentiable and so it remains to show that, for each pair of integers n and k, there is some constant M for which

$$\left|\xi^n\widehat{f}^{(k)}(\xi)\right|\le M$$

for all $\xi\in\mathbb{R}$. Let $n,k\in\mathbb{N}$ be arbitrary but fixed. By virtue of the preceding lemma (applied to ξ^n and $\widehat{f}$),

$$\xi^n\widehat{f}^{(k)}(\xi)=\sum_{j=0}^{k}(-1)^j\binom{k}{j}D^{k-j}\left(D^j\left(\xi^n\right)\widehat{f}\right)$$

for all $\xi \in \mathbb{R}$. Upon noting that

$$D^j(\xi^n) = \begin{cases} \frac{n!}{(n-j)!}\xi^{n-j} & j \leq n \\ 0 & j > n \end{cases}$$

because $n \cdot (n-1) \cdot (n-(j-1)) = n!/(n-j)!$, we obtain

$$\xi^n \widehat{f}^{(k)}(\xi) = \sum_{j=0}^{m} (-1)^j \binom{k}{j} \frac{n!}{(n-j)!} D^{k-j}\left(\xi^{n-j}\widehat{f}(\xi)\right)$$

for all $\xi \in \mathbb{R}$ where $m = \min\{k, n\}$. Thus, by an appeal to the Proposition 10, we have

$$\xi^n \widehat{f}^{(k)}(\xi) = \sum_{j=0}^{m} a_j \mathcal{F}\left(x^{k-j} f^{(n-j)}(x)\right)(\xi) \tag{6}$$

where

$$\begin{aligned} a_j &:= (i)^{k+n-2j}(-1)^j \binom{k}{j} \frac{n!}{(n-j)!} \\ &= (i)^{k+n}(i)^{-2j}(-1)^j \frac{k!}{j!(k-j)!} \frac{n!}{(n-j)!} \\ &= \frac{(i)^{k+j} k! n!}{j!(k-j)!(n-j)!} \end{aligned}$$

for $j = 0, 1, 2, \ldots m$; here, we have used the fact that $(i)^{-2j} = (i^2)^{-j} = (-1)^{-j}$. Given that $f \in \mathcal{S}$, to each $j = 1, 2, \ldots, m$ we may find some constant $C_j > 0$ for which

$$\left| x^{k-j} f^{(n-j)}(x) \right| \leq \frac{C_j}{\pi(1+x^2)}$$

for all $x \in \mathbb{R}$. It follows that, for each $j = 1, 2 \ldots, m$,

$$\left| x^{k-j} f^{(n-j)}(x) e^{ix\xi} \right| = \left| x^{k-j} f^{(n-j)}(x) \right| \left| e^{ix\xi} \right| = \left| x^{k-j} f^{(n-j)}(x) \right| \leq \frac{C_j}{\pi(1+x^2)}$$

for all $x \in \mathbb{R}$ and $\xi \in \mathbb{R}$. Consequently, for each $j = 1, 2, \ldots, m$,

$$\begin{aligned} \left| \mathcal{F}(x^{k-j} f^{(n-j)}(x))(\xi) \right| &= \left| \int_{\mathbb{R}} x^{k-j} f^{(n-j)}(x) e^{ix\xi}\, dx \right| \\ &\leq \int_{\mathbb{R}} \left| x^{k-j} f^{(n-j)}(x) e^{ix\xi} \right| dx \\ &\leq \frac{C_j}{\pi} \int_{\mathbb{R}} \frac{1}{1+x^2}\, dx \end{aligned}$$

for all $\xi \in \mathbb{R}$. Since

$$\int_{\mathbb{R}} \frac{1}{1+x^2}\, dx = \lim_{t\to\infty} \int_{-t}^{t} \frac{1}{1+x^2}\, dx = \lim_{t\to\infty} \tan^{-1}(x)\Big|_{x=-t}^{x=t} = \pi,$$

we conclude that, for each $j = 0, 1, 2 \ldots, m$,

$$\left| \mathcal{F}(x^{k-j} f^{(n-j)}(x))(\xi) \right| \leq C_j$$

for all $\xi \in \mathbb{R}$. By virtue of (6),

$$\begin{aligned}
\left|\xi^n \widehat{f}^{(k)}(\xi)\right| &= \left|\sum_{j=0}^{m} a_j \mathcal{F}\left(x^{k-j} f^{(n-j)}\right)(\xi)\right| \\
&\leq \sum_{j=0}^{m} |a_j| \left|F\left(x^{k-j} f^{(n-j)}\right)(\xi)\right| \\
&\leq \sum_{j=0}^{m} |a_j| C_j \\
&\leq M
\end{aligned}$$

for all $\xi \in \mathbb{R}$ where

$$M = \sum_{j=0}^{m} |a_j| C_j = \sum_{j=0}^{m} \left|\frac{(i)^{k+n} k!n!}{j!(k-j)!(n-j)!}\right| C_j = \sum_{j=0}^{m} \frac{k!n!C_j}{j!(k-j)!(n-j)!}.$$

$\square$

Before stating our next result, we introduce an important operation between functions, called convolution, with which you are already family from our study of the heat equation. Given $f, g \in \mathcal{S}$, the convolution of f and g is the function $f * g : \mathbb{R} \to \mathbb{C}$ defined by

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y) g(y)\, dy$$

for $x \in \mathbb{R}$. Though I will not show it, it isn't terribly hard to see that $f * g \in \mathcal{S}$ whenever $f, g \in \mathcal{S}$. An important property of the Fourier transform is that it converts convolution (a fairly computationally intensive process) into pointwise multiplication. More precisely:

Proposition 13. *Let $f, g \in \mathcal{S}$, then*

$$\mathcal{F}(f * g)(\xi) = \widehat{f * g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi)$$

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

Proof. We have

$$\begin{aligned}
\mathcal{F}(f * g)(\xi) &= \int_{\mathbb{R}} (f * g)(x) e^{ix\xi}\, dx \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x-y) g(y)\, dy\right) e^{ix\xi}\, dx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} g(y) f(x-y) e^{ix\xi}\, dx\, dy \\
&= \int_{\mathbb{R}} g(y) \left(\int_{\mathbb{R}} f(x-y) e^{ix\xi}\, dx\right) dy
\end{aligned}$$

where we have made use of Fubini's theorem to exchange the order of integration. We will preform the x-integration first (where y is treated as a constant) but first making the change of variables $x \mapsto \widetilde{x} = x - y$ so that $x = \widetilde{x} + y$, $d\tilde{x} = dx$ and therefore

$$\int_{\mathbb{R}} f(x-y) e^{ix\xi}\, dx = \int_{\mathbb{R}} f(\tilde{x}) e^{i(\tilde{x}+y)\xi}\, d\tilde{x} = e^{iy\xi} \widehat{f}(\xi)$$

where we have used the fact that $e^{i(\tilde{x}+y)\xi} = e^{i\tilde{x}\xi + iy\xi} = e^{i\tilde{x}\xi}e^{iy\xi}$ and that $e^{iy\xi}$ is constant in the eyes of the integration variable $\tilde{x}$ (and so it can be factored out of the integral). Consequently,

$$\mathcal{F}(f*g)(\xi) = \int_{\mathbb{R}} g(y)e^{iy\xi}\widehat{f}(\xi)\,dy = \widehat{f}(\xi)\int_{\mathbb{R}} g(y)e^{iy\xi}\,dy = \widehat{f}(\xi)\widehat{g}(\xi)$$

as was asserted. $\square$

We recall our definition of $\mathcal{F}^{-1}$: For $h \in \mathcal{S}$,

$$\mathcal{F}^{-1}(h)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} h(\xi)e^{-ix\xi}\,dx$$

for $x \in \mathbb{R}$. Upon noting that $\mathcal{F}^{-1}(h)(x) = (2\pi)^{-1}\mathcal{F}(h)(-x)$, we easily obtain (from Propositions 7 and 11) that $\mathcal{F}^{-1}$ is a linear transformation from S into S. The following proposition guarantees that $\mathcal{F}^{-1}$, when applied to $\widehat{f} = \mathcal{F}(f)$, returns f. In other words, $\mathcal{F}^{-1}$ is a left inverse of $\mathcal{F}$.

Proposition 14. *For any $f \in \mathcal{S}$,*

$$\left(\mathcal{F}^{-1}\circ\mathcal{F}\right)(f)(z) = \mathcal{F}^{-1}(\widehat{f})(z) = f(z)$$

for all $z \in \mathbb{R}$.

Proof. Let $f \in \mathcal{S}$. Observe that, for each $\xi \in \mathbb{R}$,

$$\lim_{t\to 0}\mathcal{F}(G_t * f)(\xi) = \lim_{t\to 0} e^{-t\xi^2}\widehat{f}(\xi) = e^0\widehat{f}(\xi) = \widehat{f}(\xi)$$

by virtue of Corollary 9 and Proposition 2. For each $t \in \mathbb{R}$, we have

$$\begin{aligned}
\left(\mathcal{F}^{-1}\circ\mathcal{F}\right)(G_t * f)(z) &= \frac{1}{2\pi}\int_{\mathbb{R}}\mathcal{F}(G_t * f)(\xi)e^{-i\xi z}\,d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}(G_t * f)(x)e^{i\xi x}\,dx\right)e^{-i\xi z}\,d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\left(\int_{\mathbb{R}}G_t(x-y)f(y)\,dy\right)e^{i\xi x}\,dx\right)e^{-i\xi z}\,d\xi \\
&= \int_{\mathbb{R}} f(y)\left(\frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}G_t(x-y)e^{i(x-z)\xi}\,dx\,d\xi\right)dy \qquad (7)
\end{aligned}$$

where we have made use of Fubini's theorem. We shall focus on the inner double integral (the term in parantheses). Observe that, though the change of variables $x \mapsto \widetilde{x} = x - y$, $dx = d\widetilde{x}$, $x - z = \widetilde{x} + y - z$ and so

$$\begin{aligned}
\int_{\mathbb{R}} G_t(x-y)e^{i(x-z)\xi}\,dx &= \int_{\mathbb{R}} G_t(\widetilde{x})e^{i(\widetilde{x}+y-z)\xi}\,d\widetilde{x} \\
&= e^{i(y-z)\xi}\int_{\mathbb{R}} G_t(\widetilde{x})e^{i\widetilde{x}\xi}\,dx \\
&= e^{i(y-z)\xi}e^{-t\xi^2}
\end{aligned}$$

where we have made use of Corollary 9. Consequently, for each $y \in \mathbb{R}$,

$$\frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}}G_t(x-y)e^{i(x-z)\xi}\,dx\,d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-t\xi^2}e^{i(y-z)\xi}\,d\xi. \qquad (8)$$

By making the change of variables $\xi \mapsto x = \sqrt{t}\xi$ so that $(y-z)\xi = \frac{y-z}{\sqrt{t}}x$ and $d\xi = dx/\sqrt{t}$, an appeal to Proposition 8 gives

$$\begin{aligned}
\frac{1}{2\pi}\int_{\mathbb{R}} e^{-t\xi^2} e^{i(y-z)\xi}\,d\xi &= \frac{1}{2\pi}\int_{\mathbb{R}} e^{-x^2} e^{i\frac{(y-z)}{\sqrt{t}})x}\frac{1}{\sqrt{t}}\,dx \\
&= \frac{1}{2\pi}\frac{1}{\sqrt{t}}\sqrt{\pi}\exp\left(-\frac{1}{4}\left(\frac{y-z}{\sqrt{t}}\right)^2\right) \\
&= \frac{1}{4\pi t}\exp\left(-\frac{(z-y)^2}{4t}\right) \\
&= G_t(z-y). \qquad (9)
\end{aligned}$$

By combining (7), (8), and (9), we obtain the identity

$$\left(\mathcal{F}^{-1}\circ\mathcal{F}\right)(G_t * f)(z) = \int_{\mathbb{R}} G_t(z-y)f(y)\,dy.$$

for $t > 0$ and $z \in \mathbb{R}$. Thanks to Proposition 4 the right hand side converges to $f(z)$ as $t \to 0$ for each $z \in \mathbb{R}$ (this uses the fact that f is continuous at every z). Consequently, we have

$$f(z) = \lim_{t\to 0}(\mathcal{F}^{-1}\circ\mathcal{F})(G_t * f)(z) = \lim_{t\to 0}\mathcal{F}^{-1}(e^{-t\xi^2}\widehat{f}(\xi))(z)$$

If we push the limit as $t \to 0$ through $\mathcal{F}^{-1}$, which is valid by virtue of a big theorem called the Dominated convergence theorem, we obtain

$$\mathcal{F}^{-1}(\widehat{f})(z) = \mathcal{F}^{-1}\left(\lim_{t\to 0} e^{-t\xi^2}\widehat{f}(\xi)\right)(z) = \lim_{t\to 0}\mathcal{F}^{-1}(e^{-t\xi^2}\widehat{f}(\xi))(z) = f(z)$$

for every $z \in \mathbb{R}$, as asserted. □

With the "heavy lifting" done in the proof above, we are finally able to prove Theorem 6.

Proof of Theorem 6. By virtue of Proposition 11, $\mathcal{F}$ is a linear transformation from $\mathcal{S}$ into itself. Since, the preceding proposition shows that $\mathcal{F}^{-1}$ is a left inverse for $\mathcal{F}$, to prove that $\mathcal{F}^{-1}$ is the *bona fide* inverse of $\mathcal{F}$ (and hence justifying the symbol $^{-1}$), it remains to prove that $\mathcal{F}^{-1}$ is also a right inverse for $\mathcal{F}$, i.e., for each $f \in \mathcal{S}$, $(\mathcal{F}\circ\mathcal{F}^{-1})(f) = f$. Fix $f \in \mathcal{S}$ and observe, by the definition of $\mathcal{F}^{-1}$,

$$\mathcal{F}^{-1}(f)(\xi) = \frac{1}{2\pi}\mathcal{F}(f)(-\xi)$$

for every $\xi \in \mathbb{R}$. Thus, for each $x \in \mathbb{R}$,

$$\begin{aligned}
\left(\mathcal{F}\circ\mathcal{F}^{-1}\right)(f)(x) &= \mathcal{F}\left(\frac{1}{2\pi}\mathcal{F}(f)(-\xi)\right)(x) \\
&= \mathcal{F}\left(\frac{1}{2\pi}\int_{\mathbb{R}} f(y)e^{i(-\xi)y}\,dy\right)(x) \\
&= \int_{\mathbb{R}}\left(\frac{1}{2\pi}\int_{\mathbb{R}} f(y)e^{i(-\xi)y}\,dy\right)e^{i\xi x}\,d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(y)e^{-i\xi y}e^{i\xi x}\,dy\right)d\xi \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(y)e^{i\zeta y}e^{-i\zeta x}\,dy\right)d\zeta \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} f(y)e^{i\zeta y}\,dy\right)e^{-i\zeta x}\,d\zeta \\
&= \mathcal{F}^{-1}\left(\mathcal{F}(f)\right)(x) \\
&= f(x)
\end{aligned}$$

where we have made the change of variables $\xi \mapsto \zeta = -\xi$ and made an appeal to the preceding proposition. $\square$

In your homework (Maybe?), you will establish many useful facts about the Fourier transform an its interplay with the so-called L^2-inner product (and its associated norm) on $\mathcal{S}$. In particular, you will prove that $\mathcal{F}$ is (modulo a multiplicative constant) is a unitary transformation (and hence an isometry). I will end this "primer" by discussing (with unfortunate extreme brevity) the application of the Fourier transform to partial differential equations.

3 Application to partial differential equations

In Proposition 10, we saw that the Fourier transform exchanges differentiation and polynomial multiplication. As a very straightforward application of that proposition, we observe that, for each $f \in \mathcal{F}$,

$$\mathcal{F}\left(\Delta f\right)(\xi) = \mathcal{F}(f^{(2)})(\xi) = (-i)^2\xi^2\widehat{f}(\xi) = -\xi^2\widehat{f}(\xi) \tag{10}$$

for all $\xi \in \mathbb{R}$. As a consequence, for $f \in \mathcal{S}$,

$$\Delta f(x) = \mathcal{F}^{-1}\left(-\xi^2\widehat{f}(\xi)\right)(x)$$

for all $x \in \mathbb{R}$. This simple observation is the basis for an entire subfield of mathematics called *pseudodifferential operator theory.* The essential idea therein is that applying a differential operator (or, more generally, a pseudodifferential operator) to a function is equivalent to multiplying that function's Fourier transform by a nice function. To see the basic idea, you should work the following exercise.

Exercise 5. *Consider an nth-degree polynomial $P(\xi) = a_0 + a_1\xi + a_2\xi^2 + \cdots a_n\xi^n$ defined for $\xi \in \mathbb{R}$ and consider*

$$\begin{aligned} P(iD) &= a_0(iD)^0 + a_1 iD + a_2(i)^2D^2 + \cdots a_n(i)^nD^n \\ &= a_0 + ia_1D - a_2D^2 + \cdots i^na_nD^n. \end{aligned}$$

As before, D is the (first-order) derivative operator with $D^jf = f^{(j)}$ for each $f \in \mathcal{S}$ and $j \in \mathbb{N}$. With this we may recognize that $P(iD)$ is a so-called partial differential operator $\Lambda = P(iD)$ defined, for $f \in \mathcal{S}$, by

$$\Lambda f(x) = P(iD)f(x) = a_0f(x) + ia_1f^{(1)}(x) - a_2f^{(2)}(x) + \cdots i^na_nf^{(n)}(x)$$

for $x \in \mathbb{R}$. In this way, the polynomial P is said to be the symbol of the differential operator $\Lambda = p(iD)$. Show that

1. *$\Lambda : \mathcal{S} \to \mathcal{S}$ is linear. For this, you must verify that, for each $f \in \mathcal{S}$, $\Lambda f \in \mathcal{S}$ and, for any constant $\alpha, \beta \in \mathbb{C}$ and $f, g \in \mathcal{S}$, $\Lambda(\alpha f + \beta g) = \alpha\Lambda f + \beta\Lambda g$.*
2. *For each $f \in \mathcal{S}$,*
$$\mathcal{F}(\Lambda f)(\xi) = P(\xi)\widehat{f}(\xi)$$
for all $\xi \in \mathbb{R}$.
3. *What's the symbol of Δ? Is (10) a consequence of the previous item?*

We now recognize this great utility of the Fourier transform and its property (10) to solve the Cauchy problem for the heat equation. Given $u_0 \in \mathcal{S}$, consider the Cauchy problem: Find $u(t,x)$ for which

$$\begin{cases} u_t = ku_{xx} & x \in \mathbb{R} \text{ and } t > 0 \\ u(0,x) = u_0(x) & x \in \mathbb{R}. \end{cases}$$

By assuming that a solution $u(t,x)$ is sufficiently nice[9], we apply the Fourier transform[10] to the partial differential equation

$$u_t = ku_{xx} = k\Delta u,$$

and find that

$$\begin{aligned} \frac{\partial}{\partial t} v(t,\xi) &= \frac{\partial}{\partial t} \mathcal{F}(u(t,x))(\xi) \\ &= \mathcal{F}(u_t(t,x)) \\ &= \mathcal{F}(ku_{xx}(t,x))(\xi) \\ &= k\mathcal{F}\left(\Delta u(t,x)\right)(\xi) \\ &= -k\xi^2 \mathcal{F}(u(t,x))(\xi) \\ &= -k\xi^2 v(t,\xi) \end{aligned}$$

for $t > 0$ and $\xi \in \mathbb{R}$ where we have made use of (10) and put

$$v(t,\xi) = \mathcal{F}(u(t,x))(\xi) = \int_{\mathbb{R}} u(t,x) e^{ix\xi}\, dx$$

for $t \geq 0$ and $\xi \in \mathbb{R}$. Also, given that $u(t,x) = u_0(x) \in \mathcal{S}$, we have

$$v(0,\xi) = \mathcal{F}(u(0,x))(\xi) = \mathcal{F}(u_0)(\xi) = \widehat{u_0}(\xi).$$

Together, this says that $v(t,\xi)$ satisfies the family of ODE-IVPs

$$\begin{cases} \frac{\partial}{\partial t} v(t,\xi) = -k\xi^2 v(t,\xi) & t > 0 \text{ and } \xi \in \mathbb{R} \\ v(0,\xi) = \widehat{u_0}(\xi) & \xi \in \mathbb{R}. \end{cases}$$

It is easy to see that this family of ODE-IVPs has the solution[11]

$$v(t,\xi) = \widehat{u_0}(\xi) e^{-k\xi^2 t} = e^{-tk\xi^2} \widehat{u_0}(\xi)$$

for $t \geq 0$ and $\xi \in \mathbb{R}$. By virtue of Corollary 9, we recall that

$$\widehat{G_t}(\xi) = e^{-tk\xi^2}$$

for $t > 0$ and $\xi \in \mathbb{R}^d$ where G_t is the heat kernel. Consequently, for $t > 0$ and $\xi \in \mathbb{R}$,

$$v(t,\xi) = \widehat{G_t}(\xi)\widehat{u_0}(\xi) = \mathcal{F}(G_t * u_0)(\xi)$$

in view of Proposition 2. By recognizing that that Fourier transform is invertible, we find that

$$u(t,x) = \mathcal{F}^{-1}(v(t,\xi))(x) = \left(G_t * u_0\right)(x)$$

[9] Say, for each $t > 0$, $x \mapsto u(t,x) \in \mathcal{S}$

[10] In much of the literature, they call this the spatial Fourier transform because it treats the independent/integration variable as x and leaves t undisturbed. Sometimes one even writes Δ_x and $\mathcal{F}_x$ to emphasize this.

[11] You should verify this by hand.

for $t > 0$ and $x \in \mathbb{R}$. In other words, the Cauchy problem is solved by

$$u(t,x) = (G_t * u_0)(x) = \frac{1}{\sqrt{4\pi kt}} \int_{\mathbb{R}} e^{\frac{(x-y)^2}{4kt}} u_0(y)\, dy$$

for $t > 0$ and $x \in \mathbb{R}$. We recall, that we also were led to this solution by the "super duper" principle of superposition and the natural appearance of G_t in the central limit theorem.

Remark 1. It should be noted that the above solution isn't defined for $t = 0$ and, if you pay close attention, I was very careful about the quantifier $t > 0$ versus $t \geq 0$. You may recall our interpretation that $u(t,x)$ satisfies the $u(0,x) = u_0(x)$ by asking instead that $u(t,x)$ recovers u_0 in the sense that

$$\lim_{t \searrow 0} u(t,x) = \lim_{t \searrow 0} (G_t * u_0)(x) = u_0(x)$$

for each $x \in \mathbb{R}$; here, $t \searrow 0$ means that we take t to zero from the right. As you know, we proved this important and fundamental property of the heat kernel in lecture[12]. In fact, this property can also be seen using the Fourier transform[13]: We have (as long as we can pass the limit through the Fourier transform),

$$\begin{aligned}
\lim_{t \searrow 0} u(t,x) &= \lim_{t \searrow 0} \mathcal{F}^{-1}(v(t,\xi))(\xi) \\
&= \mathcal{F}^{-1}\left(\lim_{t \to 0} v(t\xi)\right)(x) \\
&= \mathcal{F}^{-1}\left(\lim_{t \searrow 0} e^{-tk\xi^2}\widehat{u_0}(\xi))\right)(x) \\
&= \mathcal{F}^{-1}\left(\widehat{u_0}(\xi)\right)(x) \\
&= u_0(x)
\end{aligned}$$

by virtue of Theorem 6.

A A Gaussian Integral

Lemma 15.

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}.$$

Proof. Observe that

$$I = \int_{-\infty}^{\infty} e^{-x^2}\, dx = \int_{-\infty}^{\infty} e^{-y^2}\, dy$$

and therefore

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2}\, dy\right).$$

Using the linearity of the integral (e.g., that $\int \alpha f = \alpha \int f$) and the Fubini-Tonelli theorem,

[12] One therefore says that the heat kernel is an *approximation to the identity*

[13] However, we did use the property to show that the Fourier transform was invertible, so we should be careful to not overstate things lest we fall into circular logic.

we have

$$
\begin{aligned}
I^2 &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx \right) e^{-y^2}\, dy \\
&= \int_{-\infty}^{\infty} \left(e^{-y^2} \int_{-\infty}^{\infty} e^{-x^2}\, dx \right) dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-y^2} e^{-x^2}\, dx \right) dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2}\, dx\, dy \\
&= \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\, dA.
\end{aligned}
$$

We now make a change to polar coordinates (r, θ) by putting $x = r\cos(\theta)$, $y = r\sin(\theta)$ so that the integration domain in polar coordinates is $\{(r, \theta) : 0 \leq r < \infty,\ 0 \leq \theta \leq 2\pi\}$, and the area element $dA = dx\, dy$ becomes $r\, dr\, d\theta$. Under this change to polar coordinates, we have

$$
\begin{aligned}
I^2 &= \int_0^{2\pi} \int_0^{\infty} e^{-((r\cos(\theta))^2+(r\sin(\theta))^2)} r\, dr\, d\theta \\
&= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r\, dr\, d\theta \\
&= \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r\, d\theta\, dr \\
&= \int_0^{\infty} 2\pi e^{-r^2} r\, dr \\
&= \pi \int_0^{\infty} 2re^{-r^2}\, dr
\end{aligned}
$$

where we have used the fact that $\cos^2(\theta) + \sin^2(\theta) = 1$. Since $\frac{d}{dr}(-e^{-r^2}) = 2re^{-r^2}$, we have

$$
\int_0^{\infty} 2re^{-r^2}\, dr = \lim_{t\to\infty} \int_0^{t} 2re^{-r^2}\, dr = \lim_{t\to\infty} (-e^{-r^2})\Big|_0^t = \lim_{t\to\infty} (1 - e^{-t^2}) = 1 - 0 = 1.
$$

Thus

$$
I^2 = \pi \int_0^{\infty} 2re^{-r^2}\, dr = \pi \cdot 1 = \pi
$$

or, equivalently, $I = \sqrt{\pi}$. $\square$

References

[1] Robert G. Bartle *The Elements of Integration,* John Wiley & Sons, Inc., 1966.

[2] Elliot H. Lieb and Michael Loss. *Analysis, 2nd Ed.* Graduate Studies in Mathematics, Vol. 14, American Mathematical Society, Providence, Rhode Island, 2001.