MATH 172: THE FOURIER TRANSFORM – BASIC PROPERTIES AND THE INVERSION FORMULA

ANDRAS VASY

The Fourier transform is the basic and most powerful tool for studying translation invariant analytic problems, such as constant coefficient PDE on $\mathbb{R}^n$. It is based on the following simple observation: for $\xi \in \mathbb{R}^n$, the functions

$$v_\xi(x) = e^{ix\cdot\xi} = e^{ix_1\xi_1} \cdots e^{ix_n\xi_n}$$

are joint eigenfunctions of the operators ∂_{x_j}, namely for each j,

$$\partial_{x_j} v_\xi = i\xi_j v_\xi. \tag{1}$$

It would thus be desirable to decompose an 'arbitrary' function u as an (infinite) linear combination of the v_ξ, namely write it as

$$u(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} \hat{u}(\xi) e^{ix\cdot\xi}\, d\xi, \tag{2}$$

where $\hat{u}(\xi)$ is the 'amplitude' of the harmonic $e^{ix\cdot\xi}$ in u. (The factor $(2\pi)^{-n}$ is here due to a convention, it could also be moved to other places.) It turns out that this identity, (2), holds provided that we define

$$\hat{u}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x)\, dx;$$

(2) is then the *Fourier inversion formula.*

Rather than showing this at once, we start with a step-by-step approach. We first *define* the *Fourier transform* as

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

for $u \in L^1(\mathbb{R}^n)$, thus in particular for for $u \in C(\mathbb{R}^n)$ with $|x|^N |u|$ bounded for some $N > n$ (i.e. $|u(x)| \leq M|x|^{-N}$ for some M in say $|x| > 1$, the point being that in this case u is absolutely integrable as $\int_{|x|>1} M|x|^{-N}\, dx$ converges). Note that for such functions

$$|(\mathcal{F}u)(\xi)| \leq \int_{\mathbb{R}^n} |e^{-ix\cdot\xi}|\, |u(x)|\, dx = \int_{\mathbb{R}^n} |u(x)|\, dx,$$

so $\mathcal{F}u$ is bounded, and if we have a sequence $\xi_k \to \xi$ then $(\mathcal{F}u)(\xi) \to (\mathcal{F}u)(\xi_k)$ by the dominated convergence theorem, so $\mathcal{F}u$ is actually a *bounded continuous* function.

We can similarly define the *inverse Fourier transform*

$$(\mathcal{F}^{-1}\psi)(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix\cdot\xi} \psi(\xi)\, d\xi;$$

then $\mathcal{F}^{-1}$ maps $u \in L^1(\mathbb{R}^n)$, and in particular $u \in C(\mathbb{R}^n)$ with $|x|^N |u|$ bounded for some $N > n$ to bounded continuous functions. With these definition it is of course not clear whether $\mathcal{F}^{-1}$ is indeed the inverse of $\mathcal{F}$, and worse, not even clear whether $\mathcal{F}^{-1}\mathcal{F}\phi$ makes sense for $\phi \in L^1(\mathbb{R}^n)$, since $\mathcal{F}\phi$ is then a bounded continuous function, which is not sufficient to ensure that the integral defining $\mathcal{F}^{-1}$ actually converges! We thus proceed to study properties of $\mathcal{F}$ and $\mathcal{F}^{-1}$.

However, first we comment on the normalization. While our normalization is a very common one in analysis/PDE theory, it is by no means the only one. One could normalize the Fourier transform differently in two ways: change the constant 1 in front, and change the exponent $-ix\cdot\xi$. Thus, consider a transform

$$(\mathcal{F}_{\alpha,C}\phi)(\xi) = C\int_{\mathbb{R}^n} e^{-ix\cdot\xi/\alpha}\phi(x)\,dx = C(\mathcal{F}\phi)(\xi/\alpha),$$

which is thus simply a rescaled (by $\alpha > 0$) version of the Fourier transform, multiplied by a constant ($C > 0$). To see what the analogue of $\mathcal{F}^{-1}$ should be, note that if $\mathcal{F}^{-1}$ is indeed the inverse of $\mathcal{F}$ then the inverse of $\mathcal{F}_{\alpha,C}$, applied to some ψ, is to take the function $\Psi(\xi) = C^{-1}\psi(\alpha\xi)$, and inverse Fourier transform it, since for $\psi = \mathcal{F}_{\alpha,C}\phi$, Ψ is exactly $\mathcal{F}\phi$. Thus, the inverse transform is defined as

$$(\mathcal{F}^{-1}_{\alpha,C}\psi)(x) = C^{-1}(2\pi)^{-n}\int_{\mathbb{R}^n} e^{ix\cdot\tilde\xi}\psi(\alpha\tilde\xi)\,d\tilde\xi = C^{-1}\alpha^{-n}(2\pi)^{-n}\int_{\mathbb{R}^n} e^{ix\cdot\xi/\alpha}\psi(\xi)\,d\xi,$$

where in the last step we changed the variable of integration from $\tilde\xi$ to $\xi = \alpha\tilde\xi$.

Some common choices are $\alpha = 1$, $C = (2\pi)^{-n/2}$ and $\alpha = (2\pi)^{-1}$, $C = 1$. With the former, the formulae look as before except both the Fourier transform and the inverse Fourier transform have a $(2\pi)^{-n/2}$ in front, in a symmetric manner. With the latter, one has

$$\phi \mapsto \int e^{-2\pi ix\cdot\xi}\phi(x)\,dx$$

as the transform, and

$$\psi \mapsto \int e^{2\pi ix\cdot\xi}\psi(x)\,dx$$

as the inverse transform, which is also symmetric, though now at the cost of making the exponent a bit longer. The latter is the convention used in our textbook; the former is often used in quantum mechanics.

Returning to properties of $\mathcal{F}$ and $\mathcal{F}^{-1}$, first we note a property of $\mathcal{F}$ which is the main reason for its usefulness in studying PDE, and which is an immediate consequence of (1). Namely, suppose that $\phi \in C^1(\mathbb{R}^n)$ and both ϕ and all first derivatives $\partial_j\phi$, $j = 1,\ldots,n$, decay at infinity in the same sense as above (so $|x|^N\partial_j\phi$ is bounded for some $N > n$). Then integration by parts gives

$$\begin{aligned}(\mathcal{F}(\partial_{x_j}\phi))(\xi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\partial_{x_j}\phi(x)\,dx = -\int_{\mathbb{R}^n}\partial_{x_j}(e^{-ix\cdot\xi})\phi(x)\,dx\\ &= i\xi_j\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\phi(x)\,dx = i\xi_j(\mathcal{F}\phi)(\xi).\end{aligned}$$

In other words, the operators $\mathcal{F}$, ∂_{x_j} and multiplication by ξ_j (usually just written as ξ_j) satisfy

$$\mathcal{F}\partial_{x_j} = i\xi_j\mathcal{F}.$$

In order to remove the factor of i, we let

$$D_{x_j} = \frac{1}{i}\partial_{x_j},$$

so

$$\mathcal{F}D_{x_j} = \xi_j\mathcal{F}.$$

Note, in particular, that this gives that for ϕ as above,

$$\xi_j\mathcal{F}\phi(\xi) = \mathcal{F}D_{x_j}\phi$$

is bounded for all j, so as $\mathcal{F}\phi$ is also bounded so

$$\Big(1+\sum_{j=1}^n \xi_j^2\Big)|\mathcal{F}\phi(\xi)|^2$$

is bounded, we deduce that

$$|\mathcal{F}\phi(\xi)| \leq C/(1+|\xi|^2)^{1/2}, \tag{3}$$

i.e. the Fourier transform of ϕ actually decays, and is not merely bounded. This gives us some hope that perhaps under some additional assumptions $\mathcal{F}^{-1}(\mathcal{F}\phi)$ actually makes sense.

Before proceeding, we note that $\frac{(1+|\xi|^2)^{1/2}}{1+|\xi|}$ is bounded from below and above by positive constants – indeed, it is certainly a positive continuous function, and as $|\xi| \to \infty$, it converges to 1 (since the summand 1 is negligible in the limit in both the numerator and the denominator). Thus, (3) is equivalent to, for some $C' > 0$,

$$|\mathcal{F}\phi(\xi)| \leq C'/(1+|\xi|).$$

There is an analogous formula for $\mathcal{F}(x_j\phi)$ if we instead assume that $\phi \in C(\mathbb{R}^n)$ and $|x|^N|\phi|$ is bounded for $N > n+1$, namely

$$\begin{aligned}\mathcal{F}(x_j\phi) &= \int_{\mathbb{R}^n} e^{-ix\cdot\xi} x_j\phi(x)\,dx = \int_{\mathbb{R}^n}(x_j e^{-ix\cdot\xi})\phi(x)\,dx = \int_{\mathbb{R}^n}(i\partial_{\xi_j} e^{-ix\cdot\xi})\phi(x)\,dx\\ &= i\partial_{\xi_j}\left(\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\phi(x)\,dx\right) = -D_{\xi_j}(\mathcal{F}\phi)(\xi),\end{aligned}$$

where $D_{\xi_j} = \frac{1}{i}\partial_{\xi_j}$. In operator notation,

$$\mathcal{F}x_j = -D_{\xi_j}\mathcal{F}.$$

In particular, this tells us that if $\phi \in C(\mathbb{R}^n)$ and $|x|^N|\phi|$ is bounded for $N > n+1$ then $\mathcal{F}\phi$ is continuously differentiable, and its derivatives $D_{\xi_j}\mathcal{F}\phi$ are bounded.

In summary, the Fourier transform interchanges differentiation and multiplication by the coordinate functions (up to a $-$ sign), and correspondingly it interchanges differentiability and decay at infinity. If we only care about differentiation, the natural class of 'very nice' functions is $\mathcal{C}^\infty$, since we can differentiate its elements arbitrary many times. In view of the properties of the Fourier transform, the relevant class of 'very nice' functions consists of functions which are $\mathcal{C}^\infty$ and decay rapidly at infinity.

Definition 1. The set $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$, called the set of Schwartz functions, consists of functions $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$ such that for all $N \geq 0$ and all multiindices $\alpha \in \mathbb{N}^n$, $|x|^N D^\alpha\phi$ is bounded on $\mathbb{R}^n$.

Here we used the multiindex notation:

$$D^\alpha = D_{x_1}^{\alpha_1}\ldots D_{x_n}^{\alpha_n}.$$

The functions $\phi \in \mathcal{S}(\mathbb{R}^n)$ *decay rapidly at infinity* with all derivatives.

We can put this into a more symmetric form by noting that it suffices to consider N even, and indeed merely ask if $(1+|x|^2)^N D^\alpha\phi$ is bounded for all N and α. Expanding the first term, using $|x|^2 = x_1^2+\ldots+x_n^2$, one easily sees that this in turn is equivalent to the statement that for all multiindices $\alpha,\beta \in \mathbb{N}^n$, $x^\alpha D^\beta\phi$ is bounded. Here we wrote

$$x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_n^{\alpha_n},$$

in analogy with the notation for D^β. Note that by Leibniz' rule (i.e. the product rule for differentiation), one can write $D^\beta x^\alpha\phi$ as a finite sum of powers $\leq \alpha$ of x

times derivatives of order $\leq \beta$ of ϕ, and conversely, so in fact $x^\alpha D^\beta \phi$ being bounded for all multiindices α, β is equivalent to $D^\beta x^\alpha \phi$ being bounded for all multiindices α, β.

With this definition, using the properties above, we conclude that if $\phi \in \mathcal{S}(\mathbb{R}^n)$ then $\mathcal{F}\phi \in \mathcal{S}(\mathbb{R}^n)$ as well. Indeed,

$$\xi^\alpha D_\xi^\beta \mathcal{F}\phi = (-1)^{|\beta|} \mathcal{F} D_x^\alpha x^\beta \phi,$$

and $D_x^\alpha x^\beta \phi \in \mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$ if $\phi \in \mathcal{S}(\mathbb{R}^n)$, so the right hand side is indeed bounded.

Similar calculations show that the inverse Fourier transform satisfies

$$\mathcal{F}^{-1} D_{\xi_j} \psi = -x_j \mathcal{F}\psi, \ D_{x_j} \mathcal{F}^{-1}\psi = \mathcal{F}^{-1}(\xi_j \psi), \tag{4}$$

so

$$\mathcal{F} : \mathcal{S} \to \mathcal{S}, \ \mathcal{F}^{-1} : \mathcal{S} \to \mathcal{S}.$$

In particular, $\mathcal{F}\mathcal{F}^{-1} : \mathcal{S} \to \mathcal{S}$ and $\mathcal{F}^{-1}\mathcal{F} : \mathcal{S} \to \mathcal{S}$; the Fourier inversion formula states that these are both the identity map on $\mathcal{S}(\mathbb{R}^n)$.

Of course, we would like to know that $\mathcal{S}(\mathbb{R}^n)$ is not a trivial vector space! One example of elements of $\mathcal{S}(\mathbb{R}^n)$ is

$$\phi(x) = e^{-\langle Ax, x\rangle}, \ x \in \mathbb{R}^n,$$

where A is a positive definite operator on $\mathbb{R}^n$. Indeed, in this case $\langle Ax, x\rangle \geq a|x|^2$ for some $a > 0$, and one easily checks the membership of ϕ in $\mathcal{S}(\mathbb{R}^n)$. (Here $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$; note that $\langle A\cdot, \cdot\rangle$ is simply another inner product on $\mathbb{R}^n$.) Note also that $\mathcal{S}(\mathbb{R}^n)$ is invariant under translations, so for $x_0 \in \mathbb{R}^n$,

$$\phi(x) = e^{\langle A(x-x_0), x-x_0\rangle}, \ x \in \mathbb{R}^n,$$

gives another example. These Gaussians play an important role below since their Fourier transform is easy to compute explicitly. Notice also that in these examples we could even take a complex linear operator $A : \mathbb{C}^n \to \mathbb{C}^n$, $A = \operatorname{Re} A + i \operatorname{Im} A$, with $\operatorname{Re} A$ positive definite, to obtain examples of Schwartz functions, so e.g. on $\mathbb{R}$ the function $\phi(x) = e^{-(a+ib)x^2}$, $a > 0$, is such an example.

Another class of examples is $\mathcal{C}_c^\infty(\mathbb{R}^n)$, consisting of $\mathcal{C}^\infty$ functions of compact support, where the support of a continuous function is the closure of the set where it takes non-zero values.

Lemma 0.1. *For all $x_0 \in \mathbb{R}^n$ and $\epsilon > 0$ there is a function $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\phi(x_0) > 0$, $\phi \geq 0$ and $\operatorname{supp} \phi \subset \{x : \ |x - x_0| < \epsilon\}$.*

Proof. First one checks that the function χ defined by

$$\chi(t) = e^{-1/t}, \ t > 0; \ \chi(t) = 0, \ t \leq 0,$$

is in $\mathcal{C}^\infty(\mathbb{R})$. Then we let

$$\phi(x) = \chi\left(\frac{\epsilon^2}{2} - |x - x_0|^2\right).$$

This has all the desired properties. □

It is also useful to have bump functions that are identically 1 near x_0, but still have compact support, with $\operatorname{supp} \phi \subset \{x : \ |x - x_0| < \epsilon\}$. There are various ways of obtaining these. One is the following: let ϕ, χ be as in the proof of the lemma. Then $\tilde{\phi}(x) = \phi(x_0)/2 - \phi(x)$ is ≤ 0 near x_0, and is equal to $\phi(x_0)/2$ if

$|x - x_0| \geq \epsilon/\sqrt{2}$. Correspondingly, $\chi(\tilde{\phi}(x))$ takes the value 0 near x_0, and the constant value $\chi(\phi(x_0)/2) > 0$ if $|x - x_0| \geq \epsilon/\sqrt{2}$. Now let

$$\psi(x) = 1 - \chi(\phi(x_0)/2)^{-1}\chi(\tilde{\phi}(x));$$

then $\psi \equiv 1$ near x_0, and vanishes if $|x - x_0| \geq \epsilon/\sqrt{2}$. In summary:

Lemma 0.2. *For all $x_0 \in \mathbb{R}^n$ and $\epsilon > 0$ there is a function $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\psi(x) = 1$ for x near x_0, $\psi \geq 0$ and* $\operatorname{supp}\psi \subset \{x : \ |x - x_0| < \epsilon\}$.

As a first step towards the inversion formula, we calculate the Fourier transform of the Gaussian $\phi(x) = e^{-a|x|^2}$, $a > 0$, on $\mathbb{R}^n$ (note that $\phi \in \mathcal{S}$!) by writing it as

$$\hat{\phi}(\xi) = \left(\int_{\mathbb{R}} e^{-ax_1^2} e^{-ix_1\xi_1}\, dx_1\right) \cdots \left(\int_{\mathbb{R}} e^{-ax_n^2} e^{-ix_n\xi_n}\, dx_n\right),$$

hence reducing it to one-dimensional integrals which can be calculated by a change of variable and shift of contours. We can also proceed as follows. Write x for the one-dimensional variable, ξ for its Fourier transform variable for simplicity, and $\psi(x) = e^{-ax^2}$,

$$\hat{\psi}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} e^{-ax^2}\, dx = e^{-\xi^2/4a} \int_{\mathbb{R}} e^{-a(x + i\xi/(2a))^2}\, dx,$$

where we simply completed the square. We wish to show that

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

is a constant, i.e. is independent of ξ, and in fact it is equal to $\sqrt{\pi/a}$. But that is easy: differentiating f, we obtain $f'(\xi) = -i\int_{\mathbb{R}}(x + i\xi/(2a))e^{-a(x+i\xi/(2a))^2}\, dx$. The integrand is the derivative of $(-1/(2a))e^{-a(x+i\xi/(2a))^2}$ with respect to x, so by the fundamental theorem of calculus, $f'(\xi) = (i/(2a))e^{-a(x+i\xi/(2a))^2}|_{x=-\infty}^{+\infty} = 0$, due to the rapid decay of the Gaussian at infinity. This says that f is a constant, so for all ξ, $f(\xi) = f(0) = \int_{\mathbb{R}} e^{-ax^2}\, dx$ which can be evaluated by the usual polar coordinate trick, giving $\sqrt{\pi/a}$. Returning to $\mathbb{R}^n$, the final result is thus that

$$\hat{\phi}(\xi) = (\pi/a)^{n/2} e^{-|\xi|^2/4a}, \tag{5}$$

which is hence another Gaussian. A similar calculation shows that for such Gaussians $\mathcal{F}^{-1}\hat{\phi} = \phi$, i.e. for such Gaussians $T = \mathcal{F}^{-1}\mathcal{F}$ is the identity map. Indeed with $\psi(\xi) = e^{-b|\xi|^2}$, $b > 0$,

$$\mathcal{F}^{-1}\psi(x) = (2\pi)^{-n}(\pi/b)^{n/2} e^{-|x|^2/4b} = (4\pi b)^{-n/2} e^{-|x|^2/4b}, \tag{6}$$

so

$$\mathcal{F}^{-1}(\hat{\phi})(x) = (\pi/a)^{n/2}(4\pi/(4a))^{-n/2} e^{-4a|x|^2/4} = e^{-a|x|^2} = \phi(x).$$

Before proeeding let's recall Taylor theorem with an integral remainder formula: if f is a C^{k+1} function then

$$f(x) = \sum_{j \leq k} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j + (x - x_0)^{k+1} \int_0^1 \frac{(1-s)^k}{k!} f^{(k+1)}(x_0 + s(x - x_0))\, ds.$$

Notice that

$$\int_0^1 \frac{(1-s)^k}{k!} f^{(k+1)}(x_0 + s(x - x_0))\, ds$$

is a continuous function of x; if f was C^∞, it is in fact a C^∞ function of x. This formula can be seen in the $k=0$ case (which is what we use below) by the fundamental theorem of calculus:

$$f(x)=f(x_0)+\int_{x_0}^x f'(t)\,dt=f(x_0)+(x-x_0)\int_0^1 f'(x_0+s(x-x_0))\,ds,$$

where we wrote $t=x_0+s(x-x_0)$, and changed variables, so $dt=(x-x_0)\,ds$. To continue one writes the integrand as $1\cdot f'(x_0+s(x-x_0))$ and integrates by parts, making the indefinite integral of 1 to be $s-1$ (and after k steps, starting from the above expression, one gets $\frac{1}{k!}(s-1)^k$ for this).

Now we can show that T is the identity map on all Schwartz functions using the following lemma, which is due to Hörmander.

Lemma 0.3. *Suppose $T:\mathcal{S}\to\mathcal{S}$ is linear, and commutes with x_j and D_{x_j}. Then T is a scalar multiple of the identity map, i.e. there exists $c\in\mathbb{C}$ such that $Tf=cf$ for all $f\in\mathcal{S}$.*

Proof. Let $y\in\mathbb{R}^n$. We show first that if $\phi(y)=0$ and $\phi\in\mathcal{S}$ then $(T\phi)(y)=0$. Indeed, we can write, essentially by Taylor's theorem, $\phi(x)=\sum_{j=1}^n(x_j-y_j)\phi_j(x)$, with $\phi_j\in\mathcal{S}$ for all j. In one dimension this is just a statement that if ϕ is Schwartz and $\phi(y)=0$, then $\phi_1(x)=\phi(x)/(x-y)=(\phi(x)-\phi(y))/(x-y)$ is Schwartz: smoothness near y follows from Taylor's theorem, while the rapid decay with all derivatives from $\phi_1(x)=\phi(x)/(x-y)$. For the multi-dimensional version, one can take $\phi_j(x)=(x_j-y_j)\phi(x)/|x-y|^2$ for $|x-y|\geq 2$, say, suitably modified inside this ball. Namely, let $\rho\in\mathcal{C}_c^\infty(\mathbb{R}^n)$ identically 1 near y, supported in $\{x:|x-y|<2\}$. Then one has $\phi_{j,1}\in\mathcal{C}^\infty(\mathbb{R}^n)$ from Taylor's theorem with $\phi(x)=\sum_{j=1}^n(x_j-y_j)\phi_{j,1}(x)$. Letting $\phi_{j,2}(x)=(x_j-y_j)\phi(x)/|x-y|^2$, we have

$$\phi(x)=\sum_{j=1}^n(x_j-y_j)\phi_j(x),\qquad \phi_j(x)=\rho(x)\phi_{j,1}(x)+(1-\rho(x))\phi_{j,2}(x),$$

and ϕ_j is in $\mathcal{S}$ since the first term is $\mathcal{C}^\infty$ and has compact support, while the second is in $\mathcal{S}$ since the only potential issue is a singularity at $x=y$, but $1-\rho$ vanishes near there. Thus,

$$T\phi=\sum_{j=1}^n(x_j-y_j)(T\phi_j),$$

where we used that T is linear and commutes with multiplication by x_j for all j. Substituting in $x=y$ yields $(T\phi)(y)=0$ indeed.

Thus, fix $y\in\mathbb{R}^n$, and some $g\in\mathcal{S}$ such that $g(y)=1$. Let $c(y)=(Tg)(y)$. We claim that for $f\in\mathcal{S}$, $(Tf)(y)=c(y)f(y)$. Indeed, let $\phi(x)=f(x)-f(y)g(x)$, so $\phi(y)=f(y)-f(y)g(y)=0$. Thus, $0=(T\phi)(y)=(Tf)(y)-f(y)(Tg)(y)=(Tf)(y)-c(y)f(y)$, proving our claim.

We have thus shown that there exists $c:\mathbb{R}^n\to\mathbb{C}$ such that for all $f\in\mathcal{S}$, $y\in\mathbb{R}^n$, $(Tf)(y)=c(y)f(y)$, i.e. $Tf=cf$. Taking $f\in\mathcal{S}$ such that f never vanishes, e.g. a Gaussian as above, shows that $c=Tf/f$ is $\mathcal{C}^\infty$, since Tf and f are such.

We have not used that T commutes with D_{x_j} so far. But

$$\begin{aligned}c(y)(D_{x_j}f)(y)=T(D_{x_j}f)(y)=D_{x_j}(Tf)|_{x=y}&=D_{x_j}(c(x)f(x))|_{x=y}\\&=(D_{x_j}c)(y)f(y)+c(y)(D_{x_j}f)(y).\end{aligned}$$

Comparing the two sides, and taking f such that f never vanishes, yields

$$(D_{x_j}c)(y)=0$$

for all y and for all j. Since all partial derivatives of c vanish, c is a constant, proving the lemma. □

The actual value of c can be calculated by applying T to a single Schwartz function, e.g. a Gaussian, and then the explicit calculation from above shows that $c = 1$, so $\mathcal{F}^{-1}\mathcal{F} = \mathrm{Id}$ indeed.

Let's see how we can use the Fourier transform to solve a constant coefficient PDE. Suppose that $a_\alpha \in \mathbb{C}$ and

$$P = \sum_{|\alpha| \le m} a_\alpha D_x^\alpha$$

is an mth order constant coefficient differential operator, and consider the PDE

$$Pu = f,\ f \in \mathcal{S}(\mathbb{R}^n).$$

Then for $u \in \mathcal{S}$ (for now),

$$\mathcal{F}Pu = \mathcal{F}(\sum_{|\alpha| \le m} a_\alpha D_x^\alpha u) = \sum_{|\alpha| \le m} a_\alpha \xi^\alpha \mathcal{F}u(\xi) = p(\xi)\mathcal{F}u,$$

where we let

$$p(\xi) = \sum_{|\alpha| \le m} a_\alpha \xi^\alpha$$

the *full symbol* of P. Thus, if p never vanishes, then

$$\mathcal{F}u = \frac{\mathcal{F}f}{p(\xi)},$$

which is in $\mathcal{S}(\mathbb{R}^n)$ provided p has a lower bound like $|p(\xi)| \ge C(1+|\xi|)^{-N}$ for some N and $C > 0$, hence we get (using the Fourier inversion formula)

$$u = \mathcal{F}^{-1}\left(\frac{\mathcal{F}f}{p(\xi)}\right),$$

solving the PDE. There are some issues we would like to understand better, e.g. the non-vanishing of p and also whether we really need $u, f \in \mathcal{S}$, but before getting further into this we need to investigate the Fourier inversion formula. To give an indication of what we'll see though, note the following examples:

- Laplace's equation: $P = \sum_{j=1}^n \partial_{x_j}^2$. Then $p(\xi) = -|\xi|^2$, so p vanishes at just one point, $\xi = 0$. Note that near infinity (well, for say $|\xi| > 1$), though, $|p(\xi)| > C(1+|\xi|^2)$, for some $C > 0$.
- Helmholtz equation: $P = \sum_{j=1}^n \partial_{x_j}^2 + \lambda$. Then $p(\xi) = -|\xi|^2 + \lambda$, so if $\lambda < 0$, then p never vanishes, and indeed $|p(\xi)| \ge C(1+|\xi|^2)$, for some $C > 0$.
- Wave equation: $P = -\sum_{j=1}^{n-1} \partial_{x_j}^2 + \partial_{x_n}^2$. Then $p(\xi) = |\xi'|^2 - \xi_n^2$, where $\xi' = (\xi_1, \ldots, \xi_{n-1})$, so p vanishes on the (light) cone $|\xi'| = |\xi_n|$.
- Heat equation: $P = -\sum_{j=1}^{n-1} \partial_{x_j}^2 + \partial_{x_n}$, then $p(\xi) = |\xi'|^2 + i\xi_n$, so p only vanishes at the origin. Moreover, for $|\xi| \ge 1$, $|p(\xi)| > C(1+|\xi|^2)^{1/2}$, for some $C > 0$ – this is an *weaker* estimate than the one for Laplace's equation.

For local result, i.e. whether you can solve a PDE locally, without regard to the behavior of the solution at infinity, what matters is whether $p(\xi)$ vanishes for large ξ: this is a reflection of the fact that the Fourier transform interchanges differentiability and decay. Thus, elliptic PDE, i.e. PDE of order m such that for sufficiently large $|\xi|$, $|p(\xi)| > C(1+|\xi|^2)^{m/2}$ for some $C > 0$ are the best behaved PDE; parabolic PDE like the heat equation where a weaker estimate holds are in certain aspects almost as well behaved, while hyperbolic PDE are most interesting!

We already saw a use of the inversion formula in solving $\Delta u - u = f$. For PDEs with initial or boundary conditions, it is often best to use the partial Fourier transform. This is defined as follows. Let $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^k$, and write $\mathbb{R}^n \ni x = (y,z) \in \mathbb{R}^m \times \mathbb{R}^k$. Suppose that $f \in C^1(\mathbb{R}^m \times \mathbb{R}^k)$ and $|z|^K f$, $|z|^K \partial_{x_j} f$ are bounded for all $j = 1, \ldots, n$, and $K > k$. Define the *partial Fourier transform* of f by

$$(\mathcal{F}_z f)(y,\zeta) = \int_{\mathbb{R}^k} e^{-iz\cdot\zeta} f(y,z)\,dz,\ y \in \mathbb{R}^m,\ \zeta \in \mathbb{R}^k.$$

By arguments as for the (full) Fourier transform, one can show easily (see the problem set) that

(i) $(\mathcal{F}_z D_{z_j} f)(y,\zeta) = \zeta_j (\mathcal{F}_z f)(y,\zeta)$.
(ii) $(\mathcal{F}_z D_{y_j} f)(y,\zeta) = (D_{y_j}(\mathcal{F}_z f))(y,\zeta)$.

Similarly, as for the full Fourier transform, we have that if $f \in C^0(\mathbb{R}^m \times \mathbb{R}^k)$ and $|z|^K f$ is bounded for some $K > k+1$, then

$$\mathcal{F}_z(z_j f) = -D_{\zeta_j} \mathcal{F}_z f,\ \mathcal{F}_z(y_j f) = y_j \mathcal{F}_z f.$$

Analogous formulae also hold for

$$(\mathcal{F}_\zeta^{-1}\psi)(y,z) = (2\pi)^{-k} \int_{\mathbb{R}^k} e^{iz\cdot\zeta} \psi(y,\zeta)\,dz.$$

An iterated application of these results also shows that

$$\mathcal{F}_z, \mathcal{F}_\zeta^{-1} : \mathcal{C}^\infty(\mathbb{R}^m; \mathcal{S}(\mathbb{R}^k)) \to \mathcal{C}^\infty(\mathbb{R}^m; \mathcal{S}(\mathbb{R}^k)),$$

where $\mathcal{C}^\infty(\mathbb{R}^m; \mathcal{S}(\mathbb{R}^k))$ stands for $\mathcal{C}^\infty$ functions on $\mathbb{R}^m$ with values in $\mathcal{S}(\mathbb{R}^k)$, which means that its elements are $\mathcal{C}^\infty$ functions f on $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^k$ such that locally in y (i.e. for $|y| < R$, $R > 0$ arbitrary), $|z|^N D_x^\alpha f$ is bounded for all $N \geq 0$ and all $\alpha \in \mathbb{N}^n$.

As an application of these results, let's solve the heat equation on $(0,\infty)_t \times \mathbb{R}^n_x$:

$$u_t = k\Delta u,\ u(0,x) = \phi(x),$$

with $\phi \in \mathcal{S}(\mathbb{R}^n)$ given. Taking the partial Fourier transform in x, and writing $\mathcal{F}_x u(t,\xi) = \hat{u}(t,\xi)$, gives

$$\frac{\partial \hat{u}}{\partial t}(t,\xi) = -k|\xi|^2 \hat{u}(t,\xi),\ \hat{u}(0,\xi) = (\mathcal{F}\phi)(\xi).$$

Solving the ODE for each fixed ξ yields

$$\hat{u}(t,\xi) = e^{-k|\xi|^2 t}(\mathcal{F}\phi)(\xi),$$

hence

$$u(t,x) = \mathcal{F}_\xi^{-1}\left(e^{-k|\xi|^2 t}(\mathcal{F}\phi)(\xi)\right). \tag{7}$$

We would like to rewrite this to have a more explicit expression for u in terms of ϕ. This can be done via convolutions.

Suppose first that $f, g \in L^1(\mathbb{R}^n)$ (so e.g. f, g continuous and $|x|^N f$, $|x|^N g$ are bounded for some $N > n$.) Then $\mathcal{F}f, \mathcal{F}g$ are bounded continuous functions, hence $(\mathcal{F}f)(\mathcal{F}g)$ is a bounded continuous function as well. We cannot take its inverse Fourier transform (yet) directly, except under stronger assumptions (such as $f, g \in \mathcal{S}(\mathbb{R}^n)$), but we can ask whether $(\mathcal{F}f)(\mathcal{F}g)$ is the Fourier transform of some $\chi \in$

$L^1(\mathbb{R}^n)$). So we compute:

$$\begin{aligned}(\mathcal{F}f)(\xi)(\mathcal{F}g)(\xi) &= \left(\int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\,dx\right)\left(\int_{\mathbb{R}^n} e^{-iy\cdot\xi} g(y)\,dy\right)\\ &= \int_{\mathbb{R}^{2n}} e^{-ix\cdot\xi} e^{-iy\cdot\xi} f(x)g(y)\,dx\,dy;\end{aligned}$$

where the last equality is Fubini's theorem using that $(x,y)\mapsto f(x)g(y)$ is in $L^1(\mathbb{R}^{2n})$, which in turn follows from the measurability of $(x,y)\mapsto f(x)$, the similar statement for g, hence of their product, and an application of Tonelli's theorem. We now change variables to make the exponent of the form $e^{-iz\cdot\xi}$; we thus let $z=x+y$, while keeping x, so $y=z-x$. Then we deduce

$$\begin{aligned}(\mathcal{F}f)(\xi)(\mathcal{F}g)(\xi) &= \int_{\mathbb{R}^{2n}} e^{-iz\cdot\xi} f(x)g(z-x)\,dx\,dz\\ &= \int_{\mathbb{R}^n} e^{-iz\cdot\xi}\left(\int_{\mathbb{R}^n} f(x)g(z-x)\,dx\right)dz = (\mathcal{F}(f*g))(\xi),\end{aligned} \tag{8}$$

where we let

$$(f*g)(z)=\int_{\mathbb{R}^n} f(x)g(z-x)\,dx$$

be the *convolution* of f and g. A change of variables shows that $(f*g)(z)=(g*f)(z)$, which is consistent with $(\mathcal{F}f)(\mathcal{F}g)=(\mathcal{F}g)(\mathcal{F}f)$. A simple calculation shows that if $f,g\in\mathcal{S}(\mathbb{R}^n)$ then $f*g\in\mathcal{S}(\mathbb{R}^n)$ as well – again, this is consistent with (and indeed follows from, for here we can use the inverse Fourier transform already) $(\mathcal{F}f)(\mathcal{F}g)\in\mathcal{S}(\mathbb{R}^n)$.

If we write $\mathbb{R}^n\ni x=(x',x'')\in\mathbb{R}^m\times\mathbb{R}^k$ as above, we can talk about partial convolutions, and we still have the analogue of (8): we let

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

and then

$$(\mathcal{F}_{x''}f)(\mathcal{F}_{x''}g)=\mathcal{F}_{x''}(f*_{x''}g). \tag{9}$$

We now use this to rewrite the solution formula for the heat equation. By (7), (9) and the Fourier inversion formula, if we write

$$e^{-k|\xi|^2t}=(\mathcal{F}_xf)(t,\xi)$$

for some $f\in\mathcal{C}^\infty((0,\infty)_t;\mathcal{S}(\mathbb{R}^n_x))$, then

$$u(t,x)=(f*_x\phi)(t,x)=\int f(t,x-y)\phi(y)\,dy.$$

But this is straightforward: we have computed the inverse Fourier transform of a Gaussian in (6), so with $b=kt$,

$$f(t,x)=(4\pi kt)^{-n/2}e^{-|x|^2/(4kt)},$$

and hence

$$u(t,x)=(4\pi kt)^{-n/2}\int e^{-|x-y|^2/(4kt)}\phi(y)\,dy,$$

yielding a more explicit solution formula for the heat equation.

In fact, the heat kernel provides an alternative way of showing the Fourier inversion formula. The point is that

$$K_t(x)=(4\pi kt)^{-n/2}e^{-|x|^2/(4kt)},\ t>0,$$

is a family of good kernels on $\mathbb{R}^n$, i.e. have integral 1, are uniformly bounded, by a constant M, in $L^1(\mathbb{R}^n)$ for $t > 0$ (which follows immediately from the previous statement and that they are positive functions, so one can take $M = 1$), and finally, for any $\delta > 0$, $K_t(x)\chi_{\mathbb{R}^n\setminus B_\delta(0)} \to 0$ in $L^1(\mathbb{R}^n)$ as $t \to 0$. Thus, if h is a bounded continuous function, then for every x, using $*$ simply to denote partial convolution, $(K_t * h)(x) \to h(x)$, and the convergence is uniform on sets on which h is uniformly continuous (in particular, on compact subsets). To see this, let $A \subset \mathbb{R}^n$ be such that h is uniformly continuous on A, and for $\epsilon > 0$ let $\delta > 0$ be such that $|y| < \delta$ implies $|h(x-y) - h(x)| < \epsilon/(2M)$. Then, using $\int K_t(y)\,dy = 1$, for $x \in A$,

$$\begin{aligned}(K_t * h)(x) - h(x) &= \int K_t(y)(h(x-y) - h(x))\,dy \\ &= \int_{B_0(\delta)} K_t(y)(h(x-y) - h(x))\,dy + \int_{\mathbb{R}^n\setminus B_0(\delta)} K_t(y)(h(x-y) - h(x))\,dy.\end{aligned}$$

Now, the absolute value of the first integral is

$$\leq \int_{B_0(\delta)} |K_t(y)|\,|h(x-y) - h(x)|\,dy \leq \frac{\epsilon}{2M}\int_{B_0(\delta)} |K_t(y)|\,dy \leq \frac{\epsilon}{2M}\int_{\mathbb{R}^n} |K_t(y)|\,dy \leq \frac{\epsilon}{2},$$

while that of the second intergral is

$$\leq \int_{\mathbb{R}^n\setminus B_0(\delta)} |K_t(y)|(|h(x-y)| + |h(x)|)\,dy \leq 2\sup|h|\|K_t\|_{L^1(\mathbb{R}^n\setminus B_0(\delta))},$$

so it goes to 0 as $t \to 0$, and in particular there is $t_0 > 0$ such that this is $< \epsilon/2$ for $0 < t < t_0$. In summary, $\sup\{|(K_t * h)(x) - h(x)| :\ x \in A\} \leq \epsilon$ for $0 < t < t_0$, proving the uniform convergence on A.

In particular, if $h \in \mathcal{S}(\mathbb{R}^n)$, then h is uniformly continuous on $\mathbb{R}^n$: first, given $\epsilon > 0$, choose $R > 0$ such that $|h(x)| < \epsilon/2$ for $|x| > R$ (which is possible by the decay of h at infinity), so if $|x| > R+1$, $|y| < 1$, then $|x - y| > R$ shows $|h(x-y) - h(x)| \leq |h(x-y)| + |h(y)| < \epsilon$. On the other hand, h is continuous, thus uniformly continuous, on the compact set $\{x :\ |x| \leq R+1\}$, so there is $\delta' > 0$ such that $|y| < \delta'$ implies $|h(x-y) - h(x)| < \epsilon$. Now simply let $\delta = \min(\delta', 1)$ to conclude the uniform continuity on $\mathbb{R}^n$. Correspondingly, for Schwartz functions h, $K_t * h \to h$ uniformly on $\mathbb{R}^n$.

Now, the Fourier transform satisfies the relation

$$\int \hat{\phi}(\xi)\psi(\xi)\,d\xi = \int \phi(x)\hat{\psi}(x)\,dx, \qquad \phi, \psi \in \mathcal{S}. \tag{10}$$

(Of course, we could have denoted the variable of integration by x on both sides.) Indeed, explicitly writing out the Fourier transforms,

$$\begin{aligned}\int \left(\int e^{-ix\cdot\xi}\phi(x)\,dx\right)\psi(\xi)\,d\xi &= \int_{\mathbb{R}^{2d}} e^{-ix\cdot\xi}\phi(x)\psi(\xi)\,dx\,d\xi \\ &= \int \phi(x)\left(\int e^{-ix\cdot\xi}\psi(\xi)\,d\xi\right)dx,\end{aligned}$$

where the middle integral's integrand is in $L^1(\mathbb{R}^{2n})$, so we can apply Fubini's theorem. Of course, this argument does not really require $\phi, \psi \in \mathcal{S}$, it suffices if $\phi, \psi \in L^1(\mathbb{R}^n)$.

We now apply this result with ψ replaced by the inverse Fourier transform of K_t, which is $\psi(\xi) = (2\pi)^{-n}e^{-k|\xi|^2 t}$ as we have already calculated the Fourier and inverse Fourier transform of Gaussians; this means that $\hat{\psi} = K_t$. Thus,

$$(2\pi)^{-n}\int \hat{\phi}(\xi)e^{-k|\xi|^2 t}\,d\xi = \int \phi(x)K_t(x)\,dx,$$

and the right hand side converges to $\phi(0)$ by our previous discussion (it is $K_t * \phi$ evaluated at 0). On the other hand, as $\hat{\phi} \in L^1(\mathbb{R}^n)$ and $0 < e^{-k|\xi|^2 t} \leq 1$, and for each ξ, $e^{-k|\xi|^2 t} \to 1$ as $t \to 0$, the dominated convergence theorem shows that the left hand side converges to $(2\pi)^{-n} \int \hat{\phi}(\xi)\, d\xi$, which is the inverse Fourier transform of $\hat{\phi}$ evaluated at 0. This shows that the Fourier inversion formula holds at 0.

For general $a \in \mathbb{R}^n$, let $\Phi(x) = \phi(x+a)$, so

$$\phi(a) = \Phi(0) = (2\pi)^{-n} \int \hat{\Phi}(\xi)\, d\xi,$$

but

$$\hat{\Phi}(\xi) = \int e^{-i\xi\cdot x} \phi(x+a)\, dx = e^{i\xi\cdot a} \int e^{-i\xi\cdot(x+a)} \phi(x+a)\, dx = e^{i\xi\cdot a} (\mathcal{F}\phi)(\xi),$$

which when substituted in, yields the Fourier inversion formula:

$$\phi(a) = (2\pi)^{-n} \int e^{ix\cdot a} \hat{\phi}(\xi)\, d\xi.$$

An alternative way of achieving this at once (without reducing to the $a = 0$ case) is using $K_t(x-a)$ in place of $K_t(x)$ in the argument above; then the inverse Fourier transform of $K_t(.-a)$ is $(2\pi)^{-n} e^{i\xi\cdot a} e^{-ik|\xi|^2 t}$, which is still bounded by $(2\pi)^{-n}$ in absolute value, but now converges to $(2\pi)^{-n} e^{i\xi\cdot a}$ pointwise, so the dominated convergence theorem gives

$$\mathcal{F}^{-1}\hat{\phi}(a) = \lim_{t\to 0} \int \phi(x) K_t(x-a)\, dx = \lim_{t\to 0} \int \phi(x) K_t(a-x) = \lim_{t\to 0} (K_t * \phi)(a),$$

and the proof is finished as above

Notice that our argument only used $\phi \in L^1$ and $\hat{\phi} \in L^1$, plus that $K_t * \phi \to \phi$ uniformly to get this conclusion. If instead of the last one of these we show that for $\phi \in L^1$, $K_t * \phi \to \phi$ in L^1, then we in fact obtain that the inverse Fourier transform of the Fourier transform of such ϕ is ϕ, for there is a sequence of $t_j \to 0$ then along which the convergence is a.e. pointwise. (Notice that $\mathcal{F}^{-1}\hat{\phi}$ is a continuous function, so under these assumptions ϕ is a.e. equal to a continuous function, so it is certainly *not* a typical L^1 function.)

But

$$(K_t * \phi)(x) - \phi(x) = \int (\phi(x-y) - \phi(x)) K_t(y)\, dy,$$

so if we denote the function on the left by Φ_t, then, as on the problem set, using Fubini's theorem (plus Tonelli to justify its application, i.e. to show that the middle integral's integrand is in $L^1(\mathbb{R}^{2n})$, being bounded by $|\phi(x-y)|K_t(y) + |\phi(x)|K_t(y)$, with both terms being such),

$$\|\Phi_t\|_{L^1} \leq \int_{\mathbb{R}^{2n}} |\phi(x-y) - \phi(x)| K_t(y)\, dx\, dy = \int \|\phi(\cdot - y) - \phi\|_{L^1} K_t(y)\, dy.$$

But we have already shown that $\phi(\cdot - y) \to \phi$ in L^1 as $y \to 0$, i.e. given $\epsilon > 0$ there exists $\delta > 0$ such that $|y| < \delta$ implies $\|\phi(\cdot - y) - \phi\|_{L^1} < \epsilon/2$. Now breaking up the integral into one over $B_\delta(0)$ and one over $\mathbb{R}^n \setminus B_\delta(0)$, much as in the continuous case above, the former is $\leq \epsilon/2$, while the latter is, using $\|\phi(\cdot - y) - \phi\|_{L^1} \leq \|\phi(\cdot - y)\|_{L^1} + \|\phi\|_{L^1}$,

$$\leq 2\|\phi\|_{L^1} \int_{\mathbb{R}^n \setminus B_\delta(0)} K_t,$$

which goes to 0 as $t \to 0$. Choosing $t_0 > 0$ such that for $0 < t < t_0$ this is $< \epsilon/2$, we deduce that $0 < t < t_0$ implies $\|\Phi_t\|_{L^1} \leq \epsilon$, giving that $K_t * \phi \to \phi$ in L^1. This completes the proof that if $\phi, \hat{\phi} \in L^1$, then $\mathcal{F}^{-1}\mathcal{F}\phi = \phi$.

One more topic we discuss is the Poisson summation formula. If we are given a function $\phi \in \mathcal{S}(\mathbb{R})$, we can form a 2π-periodic function by taking $\mathcal{F}^{-1}\phi$ and summing up its translates by multiples of 2π:

$$f(x) = \sum_{m \in \mathbb{Z}} (\mathcal{F}^{-1}\phi)(x + 2\pi m).$$

Note that this sum actually converges, and does so uniformly, hence the limit is continuous: since $|\mathcal{F}^{-1}\phi(y)| \leq C_N(1 + |y|)^{-N}$ for all N, this follows from the uniform convergence of

$$\sum_{m \in \mathbb{Z}} (1 + |x + 2\pi m|^2)^{-1},$$

which in turn can be checked by considering the sum only for $x \in [-\pi, \pi]$, using that for $|m| \geq 2$, the mth term is $\leq \frac{1}{4\pi^2(|m|-1)^2}$. Indeed, since the term-by-term differentiated series still has the same property, it follows that f is $\mathcal{C}^\infty$.

Another way of producing a 2π-periodic function is to regard the integer values of ϕ as Fourier series coefficients, and consider

$$g(x) = \sum_{m \in \mathbb{Z}} \phi(m) e^{imx}.$$

A natural question is how these two functions are related. To see this, let us find the Fourier coefficients of the 2π-periodic function f. These are

$$\begin{aligned} c_k &= (2\pi)^{-1} \int_0^{2\pi} e^{-ikx} f(x)\, dx = (2\pi)^{-1} \int_0^{2\pi} \sum_{m \in \mathbb{Z}} e^{-ikx} \mathcal{F}^{-1}\phi(x + 2\pi m)\, dx \\ &= (2\pi)^{-1} \sum_{m \in \mathbb{Z}} \int_0^{2\pi} e^{-ikx} \mathcal{F}^{-1}\phi(x + 2\pi m)\, dx \end{aligned}$$

Here the last equality holds by considering the sum as a limit: $\lim_{M\to\infty} \sum_{|m| \leq M}$, and noting that the limit can be brought through the integral by the dominated convergence theorem since

$$\sum_{|m| \leq M} |\mathcal{F}^{-1}\phi(x + 2\pi m)| \leq \sum_{|m| \leq M} (1 + |x + 2\pi m|^2)^{-1} \leq \sum_{m \in \mathbb{Z}} (1 + |x + 2\pi m|^2)^{-1},$$

which we saw converged uniformly to a continuous 2π-periodic, thus bounded, function, and $[0, 2\pi]$ has finite measure. In order to evaluate this integral, we use the translation invariance of the Lebesgue measure. This gives

$$c_k = (2\pi)^{-1} \sum_{m \in \mathbb{Z}} \int_{2\pi m}^{2\pi(m+1)} e^{-ikx} \mathcal{F}^{-1}\phi(x)\, dx.$$

Now, this is

$$\begin{aligned} c_k &= (2\pi)^{-1} \lim_{M\to\infty} \sum_{|m| \leq M} \int_{2\pi m}^{2\pi(m+1)} e^{-ikx} \mathcal{F}^{-1}\phi(x)\, dx \\ &= (2\pi)^{-1} \lim_{M\to\infty} \int_{-M}^{(M+1)2\pi} e^{-ikx} \mathcal{F}^{-1}\phi(x)\, dx = (2\pi)^{-1} \int_{\mathbb{R}} e^{-ikx} \mathcal{F}^{-1}\phi(x)\, dx \end{aligned}$$

again using the dominated convergence theorem and that $\mathcal{F}^{-1}\phi \in \mathcal{S} \subset L^1$. But this is $(2\pi)^{-1}$ times the Fourier transform of $\mathcal{F}^{-1}\phi$ evaluated at k, thus it is $(2\pi)^{-1}\phi(k)$! Since the Fourier coefficients uniquely determine a 2π-periodic C^1 function thanks to the Fourier inversion formula for the Fourier series, we conclude that $(2\pi)^{-1}g = f$, i.e. that

$$\sum_{m\in\mathbb{Z}} (\mathcal{F}^{-1}\phi)(x+2\pi m) = (2\pi)^{-1} \sum_{m\in\mathbb{Z}} \phi(m)e^{imx}.$$

An interesting application is obtained by taking $\phi(\xi) = e^{-k\xi^2 t}(\mathcal{F}\psi)(\xi)$, $\psi \in \mathcal{S}(\mathbb{R})$. Then $\mathcal{F}^{-1}\phi(x) = K_t * \psi(x)$, with $K_t(x) = (4\pi kt)^{-1}e^{-x^2/(4kt)}$ the heat kernel on the real line at time $t > 0$. Summing up the translates produces a 2π-periodic function which still solves the heat equation with initial data given by the 2π-periodicized version of ψ: $\sum_{m\in\mathbb{Z}} \psi(x+2\pi m)$. On the other hand, $(2\pi)^{-1}\sum_{m\in\mathbb{Z}} e^{-km^2 t}\mathcal{F}\psi(m)e^{imx}$ is the solution of the heat equation on the circle with initial data $(2\pi)^{-1}\sum_{m\in\mathbb{Z}} \mathcal{F}\psi(m)e^{imx}$, which is also $\sum_{m\in\mathbb{Z}} \psi(x+2\pi m)$. Thus, we have two methods for solving the heat equation on the circle, say for a $\mathcal{C}^\infty$ function ψ on $\mathbb{R}$ which is supported in $(0, 2\pi)$: we can either use the Fourier series, or we can use the solution of the heat equation on $\mathbb{R}$, and sum the translates. The latter is a version of the method of images. Also notice the nice identity one gets by applying the Poisson summation formula to the heat kernel directly:

$$\sum_{m\in\mathbb{Z}} (4\pi kt)^{-1}e^{-(x+2\pi m)^2/(4kt)} = (2\pi)^{-1} \sum_{m\in\mathbb{Z}} e^{-km^2 t}e^{imx}.$$

We finally show the *Parseval/Plancherel formula*:

Lemma 0.4. *For $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$,*

$$\int_{\mathbb{R}^n} \phi(x)\,\overline{\psi(x)}\,dx = (2\pi)^{-n} \int_{\mathbb{R}^n} (\mathcal{F}\phi)(\xi)\,\overline{(\mathcal{F}\psi)(\xi)}\,d\xi.$$

Thus, up to a constant factor, the Fourier transform preserves L^2-norms:

$$\|\mathcal{F}\phi\|_{L^2(\mathbb{R}^n)} = (2\pi)^{n/2}\|\phi\|_{L^2(\mathbb{R}^n)}.$$

Proof. Before proceeding note that following relationship between $\mathcal{F}$ and $\mathcal{F}^{-1}$: for $\varphi \in L^1(\mathbb{R}^n)$,

$$(\mathcal{F}^{-1}\overline{\varphi})(x) = (2\pi)^{-n}\int e^{ix\cdot\xi}\varphi(\xi)\,d\xi = (2\pi)^{-n}\overline{\int e^{-ix\cdot\xi}\varphi(\xi)\,d\xi} = (2\pi)^{-n}\overline{\mathcal{F}\varphi)(\xi)},$$

i.e.

$$\mathcal{F}^{-1}\overline{\varphi} = (2\pi)^{-n}\overline{\mathcal{F}\varphi}. \tag{11}$$

Now, for $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$,

$$\begin{aligned}\int_{\mathbb{R}^n} \phi(x)\,\overline{\psi(x)}\,dx &= \int_{\mathbb{R}^n} \phi(x)\,(\mathcal{F}(\mathcal{F}^{-1}\overline{\psi}))(x)\,dx = \int_{\mathbb{R}^n} \mathcal{F}\phi(\xi)\,(\mathcal{F}^{-1}\overline{\psi}))(\xi)\,d\xi \\ &= \int_{\mathbb{R}^n} \mathcal{F}\phi(\xi)\,(2\pi)^{-n}\overline{(\mathcal{F}\psi)(\xi)}\,d\xi = (2\pi)^{-n}\int_{\mathbb{R}^n} (\mathcal{F}\phi)(\xi)\,\overline{(\mathcal{F}\psi)(\xi)}\,d\xi,\end{aligned}$$

where the first equality follows from $\mathcal{F}\mathcal{F}^{-1} = \mathrm{Id}$ on $\mathcal{S}(\mathbb{R}^n)$, the second from (10) and the third from (11). Substituting in $\psi = \phi$ yields that

$$\|\mathcal{F}\phi\|^2_{L^2(\mathbb{R}^n)} = (2\pi)^n\|\phi\|^2_{L^2(\mathbb{R}^n)},$$

giving the claimed conclusion. □

We note that $\mathcal{S}(\mathbb{R}^n)$, and indeed compactly supported $\mathcal{C}^\infty$ functions are dense in $L^2(\mathbb{R}^n)$.

Lemma 0.5. *For all $f \in L^2(\mathbb{R}^n)$ and $\epsilon > 0$ there exists $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ such that $\|f - \phi\|_{L^2} < \epsilon$.*

Proof. Since continuous functions of compact support are dense in $L^2(\mathbb{R}^n)$, there exists $g \in C(\mathbb{R}^n)$, of compact support, say $\operatorname{supp} g \subset B_R(0)$ such that $\|g - f\|_{L^2} < \epsilon/2$. So it suffices to find $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, supported say in $B_{R+1}(0)$, such that $\|\phi - g\|_{L^2} < \epsilon/2$. But

$$\|\phi - g\|_{L^2} \leq m(B_{R+1}(0)) \sup |\phi - g|,$$

so it suffices to find $\phi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, supported in $B_{R+1}(0)$, that is close to g in the uniform norm.

For this purpose, let $\chi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$, supported in $B_1(0)$, $\chi \geq 0$, $\chi(0) > 0$. Multiplying χ by a positive constant we may assume that $\int \chi = 1$. Now for $\delta > 0$ let $\chi_\delta(x) = \delta^{-n}\chi(x/\delta)$. Then the family χ_δ, $\delta \in (0,1)$, is a family of good kernels as $\delta \to 0$, so it follows that $\chi_\delta * g \to g$ uniformly. Further, $\chi_\delta \in \mathcal{C}^\infty(\mathbb{R}^n)$ for $\delta > 0$. Note also that the convolution $(\chi_\delta * g)(x)$ vanishes for x with $|x| \geq R+1$, for in this case $|x - y| + |y| \geq |x| \geq R + 1$ shows that either $|y| \geq 1$ or $|x - y| \geq R$, and thus the integrand of $\int g(x-y)\chi_\delta(y)\,dy$ vanishes identically. Hence, for $\delta > 0$ small, $\phi = \chi_\delta * g$ satisfies all requirements, completing the proof. $\square$

An immediate corollary is the following:

Theorem 0.6. *The Fourier transform, defined a priori on $\mathcal{S}(\mathbb{R}^n)$, has a unique continuous extension to a map $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ which still satisfies*

$$\|\mathcal{F}\phi\|_{L^2(\mathbb{R}^n)} = (2\pi)^{n/2}\|\phi\|_{L^2(\mathbb{R}^n)},\ \phi \in L^2(\mathbb{R}^n). \tag{12}$$

The corresponding statement also holds for $\mathcal{F}^{-1}$, with

$$\|\mathcal{F}^{-1}\phi\|_{L^2(\mathbb{R}^n)} = (2\pi)^{-n/2}\|\phi\|_{L^2(\mathbb{R}^n)},\ \phi \in L^2(\mathbb{R}^n). \tag{13}$$

Finally, $\mathcal{F}$ and $\mathcal{F}^{-1}$ are inverses of each other on $L^2(\mathbb{R}^n)$.

Proof. We first show the unique extendability of $\mathcal{F}$ to L^2; the argument for $\mathcal{F}^{-1}$ is completely analogous.

The linear map $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ satisfies

$$\|\mathcal{F}\phi\|_{L^2(\mathbb{R}^n)} \leq C\|\phi\|_{L^2(\mathbb{R}^n)} \tag{14}$$

for $\phi \in \mathcal{S}(\mathbb{R}^n)$ (indeed, equality, with $C = (2\pi)^{n/2}$), and thus it has a unique continuous extension to the closure of $\mathcal{S}(\mathbb{R}^n)$ in the Hilbert space $L^2(\mathbb{R}^n)$ as stated. Since the maps $\phi \mapsto \|\mathcal{F}\phi\|_{L^2(\mathbb{R}^n)}$ and $(2\pi)^{n/2}\|\phi\|_{L^2(\mathbb{R}^n)}$ are continuous on $L^2(\mathbb{R}^n)$, and they agree on the dense subset $\mathcal{S}(\mathbb{R}^n)$, the identity (13) is valid on all of $L^2(\mathbb{R}^n)$.

For the sake of completeness of details, recall that a continuous map is determined by values on a dense subset, so the uniqueness statement of the theorem follows just by the density of $\mathcal{S}(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$. To get the existence, one shows that $\mathcal{F}$ maps sequences $\{\phi_j\}_{j=1}^\infty$ in $\mathcal{S}(\mathbb{R}^n)$ which are Cauchy sequences in the $L^2(\mathbb{R}^n)$ norm to L^2-Cauchy sequences (which is immediate from (14)), and thus to L^2-convergent sequences (which is where the completeness of the target L^2 is used). Moreover, equivalent Cauchy sequences can be combined by alternating the elements into a single Cauchy sequence, showing that the images are also equivalent (since the alternated version is still Cauchy). Thus, for $f \in L^2(\mathbb{R}^n)$, taking $\phi_j \in \mathcal{S}(\mathbb{R}^n)$, $\phi_j \to f$ in L^2, and letting $\mathcal{F}f = \lim_{j\to\infty} \mathcal{F}\phi_j$ means that $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is well-defined. As $\mathcal{F}$ is linear on $\mathcal{S}$, so e.g. $\mathcal{F}(\phi_j + \psi_j) = \mathcal{F}\phi_j + \mathcal{F}\psi_j$, the linearity of $\mathcal{F}$ on $L^2(\mathbb{R}^n)$ also follows by taking limits. Finally, we need to establish the bound

$$\|\mathcal{F}f\|_{L^2(\mathbb{R}^n)} \leq C\|f\|_{L^2(\mathbb{R}^n)}$$

even for $f \in L^2$, since this gives the continuity of $\mathcal{F} : L^2 \to L^2$. But this is easy: if $\phi_j \to f$ in L^2, then $\mathcal{F}\phi_j \to \mathcal{F}f$ in L^2 by the definition of $\mathcal{F}f$. Since the norm is a continuous map on any normed space, $\|\phi_j\|_{L^2} \to \|f\|_{L^2}$ and $\|\mathcal{F}\phi_j\|_{L^2} \to \|\mathcal{F}f\|_{L^2}$. Since $\|\mathcal{F}\phi_j\|_{L^2} \leq C\|\phi_j\|_{L^2}$, letting $j \to \infty$ gives the desired conclusion.

It remains to show $\mathcal{F}\mathcal{F}^{-1} = \mathrm{Id} = \mathcal{F}^{-1}\mathcal{F}$ on L^2. But $\mathcal{F}\mathcal{F}^{-1}, \mathcal{F}^{-1}\mathcal{F}, \mathrm{Id}$ are all continuous maps on L^2, they all agree on the dense subset $\mathcal{S}$, thus on all on L^2. $\square$