Lecture #9

Fourier Transform

The Fourier Transform is a transformation on functions into functions that "transforms" derivatives operators into multiplications by polynomial. This clearly reduces the difficulty in finding solutions for PDE.

Problem: Prove that if ψ belongs to $\mathcal{S}$ then $|\psi(x)| \le \frac{C_M}{(1+|x|)^M}$ for all $M \geq 1$

Set up:

Recall the definition of Schwartz functions $\mathcal{S}$

$$\mathcal{S} = \left\{ \psi(x_1, \dots, x_n) \;/\; (*) \right\}$$

$(*)$ For any multi indices $\alpha = (\alpha_1, \dots, \alpha_n)$

$\beta = (\beta_1, \dots, \beta_n)$

$$|X^\alpha \partial^\beta \psi(x_1, \dots, x_n)| \le C_{\alpha,\beta} \text{ for all } x \text{ in } \mathbb{R}^n$$

$$X^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \qquad \partial^\beta = \partial_{x_1}^{\beta_1} \cdots \partial_{x_n}^{\beta_n} \qquad (\Delta)$$

We recall the definition of tempered distributions

$$\mathcal{S}' = \left\{ u : \mathcal{S} \longrightarrow \mathbb{R} \,(\text{or } \mathbb{C}) \quad u \text{ linear continuous} \right\}$$

that is $u(a\psi + b\varphi) = a\,u(\psi) + b\,u(\varphi)$

$:= a(u,\psi) + b(u,\varphi)$. + continuity.

Recall that

u is in $\mathcal{S}'$ then u is also in $\mathcal{D}'$

~~This is known~~ that is $\mathcal{S}' \subset \mathcal{D}'$

because $\mathcal{D} \subseteq \mathcal{S}$

We also showed that if

$u(x) = e^{x^2}$ then u is not in $\mathcal{S}'$

because

$$\int u(x)\,\varphi(x)\,dx$$ does not necessarily converges for all φ in $\mathcal{S}$

for example take $\varphi(x) = \dfrac{1}{1+x^2}$!

On the other hand

$$\left|\int u(x)\,\psi(x)\,dx\right| < \infty \text{ for all } \psi \text{ in } \mathcal{D}$$

because

$$\wedge$$

$$\int_{|x|<M} e^{x^2} \underbrace{\max_{B(0,M)}|\psi(x)|}_{N}\,dx \le \left(\int_{B(0,M)} dx\right) e^{M^2} N < \infty$$

so e^{x^2} is in $\mathcal{D}'$

This example also explains why $\mathcal{S}'$ are called "tempered" distribution: because in order for a function $u(x)$ to be a "tempered" distribution $u(x)$ has to be a function "tempered at infinity", that is with a ~~certain~~ ~~[illegible]~~ "slow" growth at infinity.

<u>Definition:</u> For any ϕ in $\mathcal{S}$, we define the Fourier transform ($\mathcal{F}\phi$) of ϕ as the function

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

where $x\cdot\xi = \langle x,\xi\rangle = \sum_{i=1}^{n} x_i \xi_i$

<u>Remark:</u> $\hat{\phi}(\xi)$ is well defined for all ξ:

$$\left|\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\,\phi(x)\,dx\right| \leq \int_{\mathbb{R}^n} |e^{-ix\cdot\xi}|\,|\phi(x)|\,dx$$

$$= \int_{\mathbb{R}^n} |\phi(x)|\,dx \leq C \int \frac{1}{(1+|x|)^M} < \infty$$

if $M > n$

<u>Definition</u>

We define the inverse Fourier transform as

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

for any ϕ in $\mathcal{S}$.

We will show later that

a) $\mathcal{F} : \mathcal{S} \longrightarrow \mathcal{S}$

$\mathcal{F}^{-1} : \mathcal{S} \longrightarrow \mathcal{S}$

and b) $\mathcal{F} \circ \mathcal{F}^{-1} = \mathcal{F}^{-1} \circ \mathcal{F} = Id.$

Fact 1 : $\widehat{\partial_{x_i} \phi}\,(\xi) = i\xi_i\, \hat{\phi}\,(\xi)$

$$\widehat{\partial_{x_i}\phi}(\xi) = \int e^{-ix\cdot\xi} (\partial_{x_i}\phi)(x)\, dx = i\xi_i \int e^{-ix\cdot\xi}\phi(x)\, dx$$

$$= i\xi_i\, \hat{\phi}(\xi)$$

Use Fact 1 to prove a) above

Fact 2: For any ϕ, ψ in S

$$\int \hat{\phi}(\xi)\psi(\xi)\,d\xi = \int \phi(x)\hat{\psi}(x)\,dx$$

Proof:

$$RHS = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-ix\cdot\xi} \phi(x)\,dx \right) \psi(\xi)\,d\xi$$

$$= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-ix\cdot\xi} \phi(x)\psi(\xi)\,d\xi\,dx$$

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

$$= \int_{\mathbb{R}^n} \phi(x)\,\hat{\psi}(x)\,dx$$

all the integral signs can be moved thanks to the absolute convergence of the integrals

We want to use Fact 1 to extend the definition of Fourier transform to tempered distribution

Definition: for any u in $\mathcal{S}'$

$$(\hat{u}, \varphi) := (u, \hat{\varphi})$$

Check that $\hat{u}$ is a tempered distribution.

Note: Observe that this definition agrees when u is a function in $\mathcal{S}$.

Fact 3: We can summarize the above with $\mathcal{F} : \mathcal{S}' \longrightarrow \mathcal{S}'$

Now we already know how to take derivatives of distributions. So let's put together the definition of Fourier transform and derivatives for $\mathcal{S}'$ like in Fact 1:

Fact 4

$$\left(\mathcal{F}(\partial_{x_i} u), \varphi\right) = i\left(\xi_i \mathcal{F}(u), \varphi\right)$$

Proof:

$$(\mathcal{F}(\partial_{x_i} u), \varphi) := (\partial_{x_i} u, \hat{\varphi}) := -(u, \partial_{x_i} \hat{\varphi})$$

Now observe that

$$\partial_{x_i} \hat{\varphi}(x) = \partial_{x_i} \int e^{-ix\cdot\xi} \varphi(\xi)\, d\xi$$

$$= \int e^{-ix\cdot\xi} (-i\xi_i \varphi(\xi))\, d\xi$$

$$= -i\, \mathcal{F}(\xi_i \varphi(\xi))(x)$$

$$= i(u, \mathcal{F}(\xi_i \varphi)) = i(\mathcal{F}(u), \xi_i \varphi)$$

$$= i(\xi_i \mathcal{F}(u), \varphi)$$

$$\mathcal{F}(\partial_{x_i} u) = i\xi_i \mathcal{F}(u)$$

<u>Example</u>: Consider the distribution $u = 1$

We want to find $\mathcal{F}(1)$.

$$(\mathcal{F}(1), \varphi) = \int \hat{\varphi}(\xi)\, d\xi = \int e^{ix\cdot\xi} \hat{\varphi}(\xi)\, d\xi \Big|_{x=0}$$

$$= \mathcal{F}^{-1}(\hat{\varphi})(0)(2\pi)^n = (2\pi)^n \varphi(0) = ((2\pi)^n \delta_0, \varphi)$$

where we used $\mathcal{F}^{-1}\mathcal{F} = \mathrm{Id}$

so

$$\mathcal{F}(1) = (2\pi)^n \delta_0$$

Now consider the distribution for $a \in \mathbb{R} \setminus \{0\}$

$$u(x) = e^{ix\cdot a} \qquad (\text{notice that } a = 0 \Rightarrow u = 1)$$

$$(\mathcal{F}(e^{ix\cdot a}), \varphi) = \int e^{ix\cdot a} \hat{\varphi}(x)\, dx$$

$$= (2\pi)^n \mathcal{F}^{-1}(\hat{\varphi})(a) = (2\pi)^n \varphi(a) = (2\pi)^n \delta_a .$$

Ex 2: The Fourier transform of a Gaussian

$$\phi(x) = e^{-a|x|^2}, \quad a > 0$$

$$\hat{\phi}(\xi) = \int_{\mathbb{R}^n} e^{-|x|^2 a} e^{-ix\cdot\xi}\, dx$$

$$= \int_{\mathbb{R}} e^{-x_1^2 a} e^{-ix_1\xi_1}\, dx_1 \times \cdots \times \int_{\mathbb{R}} e^{-x_n^2 a - ix_n\xi_n}\, dx_n$$

We calculate one of it

$$\int_{\mathbb{R}} e^{-x^2 a - ix\xi}\, dx = e^{-\frac{\xi^2}{4a}} \int e^{-\left(x\sqrt{a} + \frac{i}{2}\frac{\xi}{\sqrt{a}}\right)^2} dx$$

$$-x^2 a - ix\xi + \left(\frac{i}{2}\frac{\xi}{\sqrt{a}}\right)^2 + \left(\frac{i}{2}\frac{\xi}{\sqrt{a}}\right)^2$$

$$-\left(x\sqrt{a} + \frac{i}{2}\frac{\xi}{\sqrt{a}}\right)^2 - \frac{\xi^2}{4a}$$

$$x\sqrt{a} + \frac{i}{2}\frac{\xi}{\sqrt{a}} = \zeta \qquad d\zeta = \sqrt{a}\,dx$$

$$= \frac{e^{-\frac{\xi^2}{4a}}}{\sqrt{a}} \underbrace{\int e^{-\zeta^2}\, d\zeta}_{=\sqrt{\pi}} = \sqrt{\frac{\pi}{a}}\, e^{-\frac{\xi^2}{4a}}$$

So putting everything together

$$\mathcal{F}\left(e^{-|x|^2 a}\right)(\xi) = \left(\frac{\pi}{a}\right)^{\frac{n}{2}} e^{-\frac{|\xi|^2}{4a}}$$

So a Gaussian transform into another gaussian!

Some Important Fourier transform in 1D for simplicity

~~Contant~~

	$f(x)$	$\hat{f}(\xi)$	
Constant	1	$2\pi\,\delta_0(\xi)$	
Oscillation	$e^{ix\cdot a}$	$2\pi\,\delta_a(\xi)$	
Gaussian	$e^{-x^2 a}$	$\sqrt{\frac{\pi}{2}}\, e^{-\frac{\xi^2}{4a}}$	
Delta function	$\delta_a(x)$	$e^{i\xi\cdot a}$	$a \in \mathbb{R}$
Exponential	$e^{-a\lvert x\rvert}$	$\frac{2a}{a^2+\xi^2}$	
Heaviside Funct.	$H(x)$	$\pi\,\delta(\xi) + \frac{1}{i\xi}$	
Sign	$H(x) - H(-x)$	$\frac{2}{i\xi}$	

Properties of Fourier Transform (also 1-D)

	Function	F.T.
i)	$\frac{d}{dx} f(x)$	$i\xi \hat{f}(\xi)$
ii)	$x f(x)$	$i \frac{d}{d\xi} \hat{f}(\xi)$
iii)	$f(x-a)$	$e^{-ia\xi} \hat{f}(\xi)$
iv)	$e^{iax} f(x)$	$\hat{f}(\xi - a)$
v)	$a f(x) + b g(x)$	$a \hat{f}(\xi) + b \hat{g}(\xi)$ (linearity)
vi)	$f(ax)$	$\frac{1}{a} \hat{f}\left(\frac{\xi}{a}\right)$

MATH 152: THE FOURIER TRANSFORM – TEMPERED DISTRIBUTIONS

Recall that the support of a continuous function f, denoted by $\operatorname{supp} f$, is the closure of the set $\{x : f(x) \neq 0\}$. Thus, $x \notin \operatorname{supp} f$ if and only if there exists a neighborhood U of x such that $y \in U$ implies $f(y) = 0$. Thus, $\operatorname{supp} f$ is closed by definition; so for continuous functions on $\mathbb{R}^n$, it is compact if and only if it is bounded.

The support of a distribution u is defined similarly. One says that $x \notin \operatorname{supp} u$ if there exists a neighborhood U of x, such that on U, u is given by the zero function. That is, $x \notin \operatorname{supp} u$ if there exists U as above such that for all $\phi \in \mathcal{D}$ with $\operatorname{supp} \phi \subset U$, $u(\phi) = 0$. For example, if $u = \delta_a$ is the delta distribution at a, then $\operatorname{supp} u = \{a\}$, since $u(\phi) = \phi(a)$, so if $x \neq a$, taking U as a neighborhood of x that is disjoint from a, $u(\phi) = 0$ follows for all $\phi \in \mathcal{D}$ with $\operatorname{supp} \phi \subset U$.

Note that if u is a distribution and $\operatorname{supp} u$ is compact, u, which is a priori a map $u : \mathcal{D} \to \mathbb{C}$, extends to a map $u : \mathcal{C}^\infty(\mathbb{R}^n) \to \mathbb{C}$, i.e. $u(\phi)$ is naturally defined if ϕ is just smooth, and does not have compact support. To see this, let $f \in \mathcal{D}$ be identically one in a neighborhood of $\operatorname{supp} u$, and for $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$ define $u(\phi) = u(f\phi)$, noting that $f\phi \in \mathcal{D}$. If $u = u_g$ is given by integration against a continuous function g of compact support, this just says that we defined for $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$

$$u_g(\phi) = \int_{\mathbb{R}^n} g(x)f(x)\phi(x)\,dx = \int_{\mathbb{R}^n} g(x)\phi(x)\,dx,$$

which is of course the standard definition if ϕ had compact support. Note that the second equality above holds since we are assuming that f is identically 1 on $\operatorname{supp} g$, i.e. wherever f is not 1, g necessarily vanishes. We should of course check that the definition of the extension of u does not depend on the choice of f (which follows from the above calculation if u is given by a continuous function g). But this can be checked easily, for if f_0 is another function in $\mathcal{D}$ which is identically one on $\operatorname{supp} u$, then we need to make sure that $u(f\phi) = u(f_0\phi)$ for all $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$, i.e. that $u((f - f_0)\phi) = 0$ for all $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$. But $f = f_0 = 1$ on a neighborhood of $\operatorname{supp} u$, so $(f - f_0)\phi$ vanishes there, hence $u((f - f_0)\phi) = 0$ indeed.

Recall that $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ is the space of Schwartz functions, i.e. the functions $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$ with the property that for any multiindices $\alpha, \beta \in \mathbb{N}^n$, $x^\alpha \partial^\beta \phi$ is bounded. Here we wrote $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$, and $\partial^\beta = \partial_{x_1}^{\beta_1} \dots \partial_{x_n}^{\beta_n}$; with $\partial_{x_j} = \frac{\partial}{\partial x_j}$. (This notation with α, β, is called the multiindex notation.) Convergence of a sequence $\phi_m \in \mathcal{S}$, $m \in \mathbb{N}$, to some $\phi \in \mathcal{S}$, in $\mathcal{S}$ is defined as follows. We say that ϕ_m converges to ϕ in $\mathcal{S}$ if for all multiindices α, β, $\sup |x^\alpha \partial^\beta (\phi_m - \phi)| \to 0$ as $m \to \infty$, i.e. if $x^\alpha \partial^\beta \phi_m$ converges to $x^\alpha \partial^\beta \phi$ uniformly.

A tempered distribution u is defined as a continuous linear functional on $\mathcal{S}$ (this is written as $u \in \mathcal{S}'$), i.e. as a map $u : \mathcal{S} \to \mathbb{C}$ which is linear: $u(a\phi + b\psi) = au(\phi) + bu(\psi)$ for all $a, b \in \mathbb{C}$, $\phi, \psi \in \mathcal{S}$, and which is continuous: if ϕ_m converges to ϕ in $\mathcal{S}$ then $\lim_{m\to\infty} u(\phi_m) = u(\phi)$ (this is convergence of complex numbers).

In particular any tempered distribution is a distribution, since $\phi \in \mathcal{D}$ implies $\phi \in \mathcal{S}$, and convergence of a sequence in $\mathcal{D}$ implies that in $\mathcal{S}$ (recall that convergence of a sequence in $\mathcal{D}$ means that the supports stay inside a fixed compact set and the convergence of all derivatives is uniform). The converse is of course not true; e.g. any continuous function f on $\mathbb{R}^n$ defines a distribution, but $\int_{\mathbb{R}^n} f(x)\phi(x)\,dx$ will not converge for all $\phi \in \mathcal{S}$ if f grows too fast at infinity; e.g. $f(x) = e^{|x|^2}$ does not define a tempered distribution. On the other hand, any continuous function f satisfying an estimate $|f(x)| \leq C(1+|x|)^N$ for some N and C defines a tempered distribution $u = u_f$ via

$$u(\psi) = u_f(\psi) = \int_{\mathbb{R}^n} f(x)\psi(x)\,dx, \quad \psi \in \mathcal{S}.$$

This is the reason for the 'tempered' terminology: the growth of f is 'tempered' at infinity. Moreover, any distribution u of compact support, e.g. δ_a for $a \in \mathbb{R}^n$, is tempered. Indeed, $\psi \in \mathcal{S}$ certainly implies that $\psi \in \mathcal{C}^\infty(\mathbb{R}^n)$, so $u(\psi)$ is defined, and it is easy to check that this gives a tempered distribution. In particular, $\delta_a(\psi) = \psi(a)$, and it is easy to see that this defines a tempered distribution.

We defined the Fourier transform on $\mathcal{S}$ as

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

and the inverse Fourier transform as

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

The Fourier transform satisfies the relation

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

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

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

where the middle integral converges absolutely (since ϕ, ψ decrease rapidly at infinity), hence the order of integration can be changed. Of course, this argument does not really require $\phi, \psi \in \mathcal{S}$, it suffices if they decrease fast enough at infinity, e.g. $|\phi(x)| \leq C(1+|x|)^{-s}$ for some $s > n$, and similarly for ψ.

In the language of distributional pairing this just says that the tempered distributions u_ϕ, resp. $u_{\hat{\phi}}$, defined by ϕ, resp. $\hat{\phi}$, satisfy

$$u_{\hat{\phi}}(\psi) = u_\phi(\hat{\psi}), \qquad \psi \in \mathcal{S}.$$

Motivated by this, we define the Fourier transform of an arbitrary tempered distribution $u \in \mathcal{S}'$ by

$$\hat{u}(\psi) = u(\hat{\psi}), \qquad \psi \in \mathcal{S}.$$

It is easy to check that $\hat{u}$ is indeed a tempered distribution, and as observed above, this definition is consistent with the original one if u is a tempered distribution given by a Schwartz function ϕ (or one with enough decay at infinity). It is also easy to see that the Fourier transform, when thus extended to a map $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$, still has the standard properties, e.g. $\mathcal{F}(D_{x_j}u) = \xi_j \mathcal{F}u$. Indeed, by definition, for all $\psi \in \mathcal{S}$,

$$(\mathcal{F}(D_{x_j}u))(\psi) = (D_{x_j}u)(\mathcal{F}\psi) = -u(D_{x_j}\mathcal{F}\psi) = u(\mathcal{F}(\xi_j\psi)) = (\mathcal{F}u)(\xi_j\psi) = (\xi_j\mathcal{F}u)(\psi),$$

finishing the proof.

The inverse Fourier transform of a tempered distribution is defined analogously, and it satisfies $\mathcal{F}^{-1}\mathcal{F} = \mathrm{Id} = \mathcal{F}\mathcal{F}^{-1}$ on tempered distributions as well.

As an example, we find the Fourier transform of the distribution $u = u_1$ given by the constant function 1. Namely, for all $\psi \in \mathcal{S}$,

$$\hat{u}(\psi) = u(\hat{\psi}) = \int_{\mathbb{R}^n} \hat{\psi}(x)\, dx = (2\pi)^n \mathcal{F}^{-1}(\hat{\psi})(0) = (2\pi)^n \psi(0) = (2\pi)^n \delta_0(\psi).$$

Here the first equality is from the definition of the Fourier transform of a tempered distribution, the second from the definition of u, the third by realizing that the integral of any function ϕ (in this case $\phi = \hat{\psi}$) is just $(2\pi)^n$ times its inverse Fourier transform evaluated at the origin (directly from the definition of $\mathcal{F}^{-1}$ as an integral), the fourth from $\mathcal{F}^{-1}\mathcal{F} = \mathrm{Id}$ on Schwartz functions, and the last from the definition of the delta distribution. Thus, $\mathcal{F}u = (2\pi)^n\delta_0$, which is often written as $\mathcal{F}1 = (2\pi)^n\delta_0$. Similarly, the Fourier transform of the tempered distribution u given by the function $f(x) = e^{ix\cdot a}$, where $a \in \mathbb{R}^n$ is a fixed constant, is given by $(2\pi)^n\delta_a$ since

$$\hat{u}(\psi) = u(\hat{\psi}) = \int_{\mathbb{R}^n} e^{ix\cdot a}\hat{\psi}(x)\, dx = (2\pi)^n \mathcal{F}^{-1}(\hat{\psi})(a) = (2\pi)^n \psi(a) = (2\pi)^n \delta_a(\psi),$$

while its inverse Fourier transform is given by δ_{-a} since

$$\mathcal{F}^{-1}u(\psi) = u(\mathcal{F}^{-1}\psi) = \int_{\mathbb{R}^n} e^{ix\cdot a}\mathcal{F}^{-1}\psi(x)\, dx = \mathcal{F}(\mathcal{F}^{-1}\psi)(-a) = \psi(-a) = \delta_{-a}(\psi).$$

We can also perform analogous calculations on δ_b, $b \in \mathbb{R}^n$:

$$\mathcal{F}\delta_b(\psi) = \delta_b(\mathcal{F}\psi) = (\mathcal{F}\psi)(b) = \int e^{-ix\cdot b}\psi(x)\, dx,$$

i.e. the Fourier transform of δ_b is the tempered distribution given by the function $f(x) = e^{-ix\cdot b}$. With $b = -a$, the previous calculations confirm what we knew anyway namely that $\mathcal{F}\mathcal{F}^{-1}f = f$ (for this particular f).

Note that the Fourier transform of a compactly supported distribution can be calculated directly. Indeed, $g_\xi(x) = e^{-ix\cdot\xi}$ is a $\mathcal{C}^\infty$ function (of x), and compactly supported distributions can be evaluated on these. Thus, we can define $\mathcal{F}u$ as the tempered distribution given by the function $\xi \mapsto u(g_\xi)$. For example, if $u = \delta_b$, then $\mathcal{F}u$ is given by the function $\delta_b(g_\xi) = g_\xi(b) = e^{i\xi\cdot b}$ in accordance with our previous calculation. Of course, if u is given by a continuous function f of compact support, then $u(g_\xi) = \int f(x)g_\xi(x)\, dx = \int e^{-ix\cdot\xi}f(x)\, dx = (\mathcal{F}f)(\xi)$ – indeed, this motivated the definition of $\mathcal{F}u$. This definition is also consistent with the general one for tempered distributions, as we have seen on the particular example of delta distributions. The fact that for compactly supported distributions u, $\mathcal{F}u$ is given by $\xi \mapsto u(g_\xi)$ shows directly that for such u, $\mathcal{F}u$ is given by a $\mathcal{C}^\infty$ function: $u(g_\xi) = u(e^{-ix\cdot\xi})$, and differentiating this with respect to ξ simply differentiates g_ξ, i.e. simply gives another exponential (times a linear function), which is still $\mathcal{C}^\infty$.

MATH 152: THE FOURIER TRANSFORM – THE INVERSION FORMULA

Recall that $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ is the space of Schwartz functions, i.e. the functions $\phi \in \mathcal{C}^\infty(\mathbb{R}^n)$ with the property that for any multiindices $\alpha, \beta \in \mathbb{N}^n$, $x^\alpha \partial^\beta \phi$ is bounded. Here we wrote $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$, and $\partial^\beta = \partial_{x_1}^{\beta_1} \dots \partial_{x_n}^{\beta_n}$; with $\partial_{x_j} = \frac{\partial}{\partial x_j}$. (This notation with α, β, is called the multiindex notation.)

We defined the Fourier transform on $\mathcal{S}$ as

$$(\mathcal{F}\phi)(\xi) = \hat{\phi}(\xi) = \int_{\mathbb{R}^n} e^{-ix\cdot\xi}\, \phi(x)\, dx, \tag{1}$$

and the inverse Fourier transform as

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

We showed by integration by parts that $\mathcal{F}, \mathcal{F}^{-1}$ satisfy

$$\mathcal{F} D_{x_j}\phi = \xi_j \mathcal{F}\phi,\ -D_{\xi_j}\mathcal{F}\phi = \mathcal{F}(x_j\phi),\ D_{x_j} = i^{-1}\partial_j, \tag{3}$$

with similar formulae for the inverse Fourier transform:

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

We used this to show that $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ and similarly for $\mathcal{F}^{-1}$; indeed, if $\phi \in \mathcal{S}$, then $x^\alpha \partial^\beta \phi$ is bounded for all multiindices α, β. But the Fourier transform of this a constant multiple of $\partial^\alpha \xi^\beta \hat{\phi}$. But we in fact have that $(1+|x|^2)^{(n+1)/2} x^\alpha \partial^\beta \phi$ is also bounded (the first factor in effect simply increases α), so $|x^\alpha \partial^\beta \phi| \leq C(1+|x|^2)^{-(n+1)/2}$ for some $C > 0$. Thus,

$$\begin{aligned}|\partial^\alpha \xi^\beta \hat{\phi}(\xi)| &= |\int_{\mathbb{R}^n} e^{-ix\cdot\xi}\, (x^\alpha \partial^\beta \phi)(x)\, dx| \\ &\leq \int_{\mathbb{R}^n} |e^{-ix\cdot\xi}\, (x^\alpha \partial^\beta \phi)(x)|\, dx \leq \int_{\mathbb{R}^n} C(1+|x|^2)^{-(n+1)/2} = M < +\infty,\end{aligned}$$

so $\sup |\partial^\alpha \xi^\beta \hat{\phi}| \leq M$, i.e. $\partial^\alpha \xi^\beta \phi$ is bounded indeed. Although the derivatives and the multiplications are in the opposite order as in the definition of $\mathcal{S}$, using Leibniz' rule (i.e. the product rule) for differentiation, we get other terms of the same form, so we conclude that $\hat{\phi} \in \mathcal{S}$ indeed. The proof for the inverse Foruier transform is of course very similar.

We also calculated 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

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

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

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

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.

Now we can show that T is the identity map on all Schwartz functions using the following lemma.

Lemma 0.1. *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. 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.