Pseudo-differential Operators

Notation :

Let $\Omega \subset \mathbb{R}^n$ be open. Let $k \in \mathbb{N} \cup \{0, \infty\}$.
$C^k(\Omega)$: Complex valued functions on Ω that are k-times continuously differentiable.
$C^k_0(\Omega)$: Function s in $C^k(\Omega)$ which vanish everywhere outside a compact subset of Ω. We set $D(\Omega) = C^\infty_0(\Omega)$
We will use multi-indices to denote partial derivatives. As a reminder, a multi-index is an an element $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$ such that $|\alpha| \equiv \alpha_1 + \cdots + \alpha_n$, and $\alpha! = \alpha_1 \cdots \alpha_n$. We will sometimes denote $\frac{\partial}{\partial x_j}$ by ∂_{x_j} or ∂_j. We set $D_j \equiv -i\frac{\partial}{\partial x_j}$ where i is the imaginary unit. Then we set $\partial^\alpha \equiv \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$, and $D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}$. For $x \in \mathbb{R}^n$, we also set $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.
From now on, for the sake of brevity we will assume implicitly that all definititions, and anything that must be done on some domain, takes place on Ω

A differential operator on is a finite linear combination of derivatives arbitrary orders with smooth coefficients. The order of the operator is the highest order derivative included in the linear combination. Explicitly, a differential operator of order n is

$$P = \sum_{|\alpha| \leqslant n} a_\alpha(x) D^\alpha$$

where $a_\alpha \in C^\infty$ are the coefficients. The symbol P is the polynomial function of ξ defined on $\Omega \times \mathbb{R}^n$ by

$$p(x, \xi) = \sum_{|\alpha| \leqslant m} a_\alpha(x) \xi^\alpha.$$

and its principle symbol is

$$p_n(x, \xi) = \sum_{|\alpha| = m} a_\alpha(x) \xi^\alpha$$

where n is the order of the highest derivative.

Distributions :

A distribution is a linear functional f on $D(\Omega)$ such that for any compact subset $K \subset \Omega$, there exists an integer n and a constant C such that for all $\varphi \in D(\Omega)$ which vanish everywhere outside of K, we have

$$|\langle f, \varphi \rangle| \leqslant C \sup_{x \in K} \sup_{|\alpha| \leqslant n} |\partial^\alpha \varphi(x)|$$

where $\langle f, \varphi \rangle$ is to be defined. As usual, the space of distributions on $D(\Omega)$ is denoted by $D'(\Omega)$. If $f \in L^1_{loc}(\Omega)$, the space of locally integrable functions on Ω, then we set

$$\langle f, \varphi \rangle = \int_\Omega f(x) \varphi(x) \, dx \qquad (1)$$

for al $\varphi \in D(\Omega)$, so that $L^1_{loc}(\Omega) \subset D'(\Omega)$. Motivated by integration by parts, the derivative f' of a distribution f is defined by

$$\langle f', \varphi \rangle = -\langle f, \varphi' \rangle,$$

and this coincides with the derivative of f if f is a differentiable function, as can be seen by using integration by parts on on (1) and the fact that φ is zero everywhere outside some subset of Ω. Thus any differetiable operator P can be extended to a linear mapping from D' to D' since

$$\langle Pf, \varphi \rangle = \langle f,^{t} P\varphi \rangle,$$

where ${}^{t}P\varphi \sum_{|\alpha| \leqslant n} (-1)^{|\alpha|} D^{\alpha}(a_{\alpha}\varphi)$.

Convolutions :

Let $f,\ g \in D(\Omega)$. Then we represent the convolution of f and g by $f * g$, defined as

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

where the last equality just follows from a simple change of variable. Intuitively, if you imagine g as a bump function, then the convolution of f with g is a weighted average of f around x. That the convolution of f is smoother than f itself is an important property of the convolution, and can be understood intuitively by the fact that convoluting is a kind of averaging, and so any bad behaviours of the function (ie. sudden changes in value) tend to be eliminated due to this sort of averaging. The convolution has the following algebraic properties:

1. $f * g = g * f$ (commutativity)
2.$(f * g) * h = f * (g * h)$ (associativity)
3.$f * (g + h) = f * g + f * g$ (distributivity)
4. For any $a \in \mathbb{C}, a(f * g) = (af) * g = f * (ag)$ (associativity with scalar multiplication)

5. There is no identity element

It is also true that $D(\Omega)$ is closed under convolutions, and so $D(\Omega)$ with the convolution forms a commutative algebra. Although there is no identity element, we can approximate the identity by choosing an appropriate function (called a mollifier, see Figure 3 on page 8) such as a normalized Gaussian (or any appropriate function that approximates the Dirac delta function). Actually, there is a standard methodology for constructing functions which approximate identities: Take an absolutely integrable function ν on $\mathbb{R}^n$, and define

$$\nu_\epsilon(x) = \frac{\nu(\frac{x}{\epsilon})}{\epsilon^n}$$

then

$$\lim_{\epsilon \to 0^+} \int_{\mathbb{R}^n} \nu_\epsilon f = f(0)$$

for all smooth (actually continuous is sufficient) compactly supported functions f, hence $\nu_\epsilon \to \delta$ as $\epsilon \to 0^+$ in $D'(\mathbb{R}^n)$. We can define the convolution for less restrictive spaces of functions, such as $L^1(\Omega)$, but for our purposes we will define it for the space of functionals on $D(\Omega)$: $D'(\Omega)$. Let $u \in D'(\Omega)$, $v \in S'$, then we set

$$u * v = \langle u, v_x \rangle$$

where $v_x(y) = v(x - y)$. It easily follows that $\partial^\alpha(u * v) = \partial^\alpha u * v = u * \partial^\alpha v$, and also supp $(u * v) \subset$ supp $u+$supp v.

Something very important is that there is a regularization procedure: Let $\varphi \in D(\mathbb{R}^n)$ be nonnegative with integral equal to 1, and let $\epsilon > 0$. Set $\varphi_\epsilon = \frac{\varphi(\frac{x}{\epsilon})}{\epsilon^n}$. Then for $u \in D'(\mathbb{R}^n)$, set $u_\epsilon = u * \varphi_\epsilon$, then for all $v \in D(\mathbb{R}^n)$, we have that

$$\int u_\epsilon v \to \langle u, v \rangle \text{ as } \epsilon \to 0.$$

So we can approximate distributions by regular functions.

Finally we define the convolution of distributions. Let $u \in D(\mathbb{R}^n)$, v, $\varphi \in$

$D(\mathbb{R}^n)$, then

$$\int (u * v)\varphi = \langle u, \tilde{v} * \varphi \rangle,$$

where $\tilde{v}(x) = v(-x)$. So we set $\langle u * v, \varphi \rangle = \langle u, \tilde{v} * \varphi \rangle$. The differentiation and support properties previously which were previously stated for $u,\ v \in D(\mathbb{R}^n)$ still hold, along with sing supp $(u * v) \subset$ sing supp u + sing supp v, where sing supp means singular support, which is the complement of the largest open set on which a distribution is smooth function, ie. the closed set where the distribution is not a smooth function.

Example 1 : Let δ denote the Dirac delta function and let $f \in D(\mathbb{R})$. Then we have that

$$(\delta' * f)(x) = (\delta * f')(x) = f'(x),$$

so that differentiation is equivalent to convolution with the derivative of the Dirac delta function.

Example 2 : Consider the function

$$\varphi(x) = \frac{1}{\sqrt{\pi}} e^{-x^2},\ \varphi_\epsilon(x) = \frac{1}{\epsilon\sqrt{\pi}} e^{-(\frac{x}{\epsilon})^2}$$

and consider $\frac{\sin x}{x}$.

$$\int_{\mathbb{R}} \frac{\sin x}{x} \frac{1}{\sqrt{\pi}} e^{-x^2}\, dx \approx 0.923,\ \int_{\mathbb{R}} \frac{\sin x}{x} \frac{2}{\sqrt{\pi}} e^{-(2x)^2}\, dx \approx 0.98$$

$$\int_{\mathbb{R}} \frac{\sin x}{x} \frac{10}{\sqrt{\pi}} e^{-(10x)^2}\, dx \approx 0.999,\ \int_{\mathbb{R}} \frac{\sin x}{x} \frac{100}{\sqrt{\pi}} e^{-(100x)^2}\, dx \approx 0.99999$$

and so on. So as $\epsilon \to 0^+$, we see that the integral converges to 1, as expected since $\frac{sinx}{x} \to 1$ as $x \to 0$. Note that these functions don't even meet the conditions that were imposed! Evidently this works for certain more general functions.

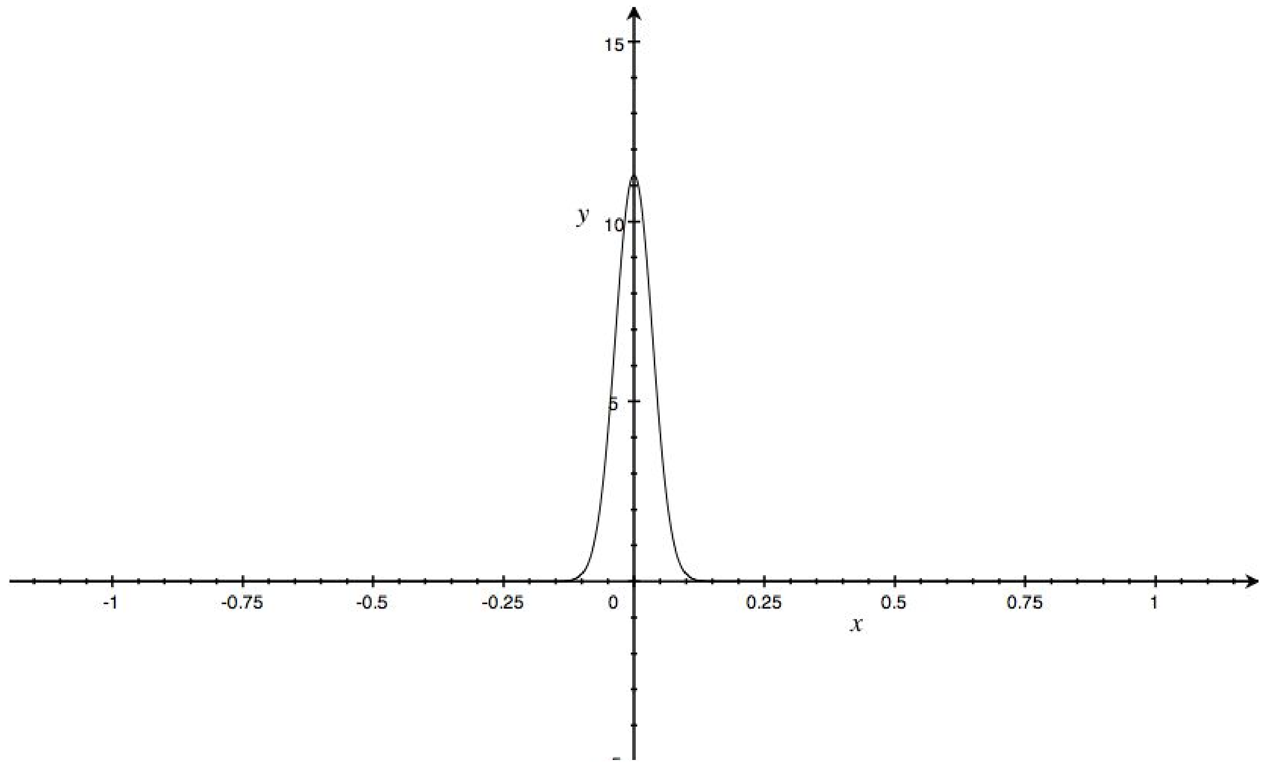

Figure 1: A graph of $\frac{20}{\sqrt{\pi}}e^{-400x^2}$, an approximation to the Dirac delta function.

Example 3 : $\frac{sin}{x}$ is even, and the derivative of $\varphi(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}$ is odd, so the integral of their products is trivially 0, which we would expect from example 1 since the derivative of $\frac{sin}{x}$ at $x = 0$ is 0. So lets consider something more interesting: Consider $e^{-(1-x)^2}$, $\varphi'_\epsilon(x) = -\frac{2x}{\epsilon^3\sqrt{\pi}}e^{-(\frac{x}{\epsilon})^2}$. Let's take the convolution at $x = 0$:

$$(\delta' * \varphi)(0) = \int_{\mathbb{R}} -e^{-(1+x)^2}\frac{2x}{\epsilon^3\sqrt{\pi}}e^{-(\frac{x}{\epsilon})^2}\,dx = \frac{2}{(\epsilon^2+1)^{\frac{3}{2}}e^{\frac{1}{1+\epsilon^2}}},$$

and as $\epsilon \to 0^+$, this goes to $\frac{2}{e}$. Now $\frac{d}{dx}e^{-(1-x)^2}|_{x=0} = \frac{2}{e}$, as expected (although again, this function doesn't meet the imposed conditions).

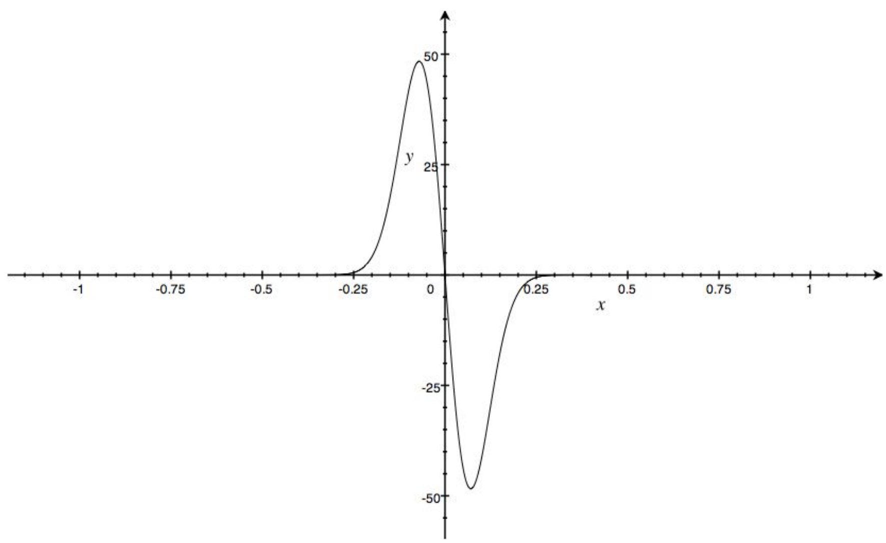

Figure 2: A graph of $-\dfrac{2000x}{\sqrt{\pi}}e^{-100x^2}$, an approximation to the derivative of the Dirac delta function.

The intution behind how δ' works (I will use δ informally here, imagine it is some very localized bump function if you like) is that for $0 < \epsilon \ll 1$, $\delta'(x) \approx \frac{\delta(x+\epsilon)-\delta(x-\epsilon)}{2\epsilon}$, so that

$$\int_{\mathbb{R}} \delta'(x)f(x)\,dx \approx \int_{-\epsilon}^{\epsilon} \frac{\delta(x+\epsilon)-\delta(x-\epsilon)}{2\epsilon} f(x)\,dx = \frac{f(-\epsilon)}{2\epsilon} - \frac{f(\epsilon)}{2\epsilon}$$

$$= -\left(\frac{f(\epsilon)-f(-\epsilon)}{2\epsilon}\right) \approx -f'(0)$$

for a sufficiently well behaved function f.

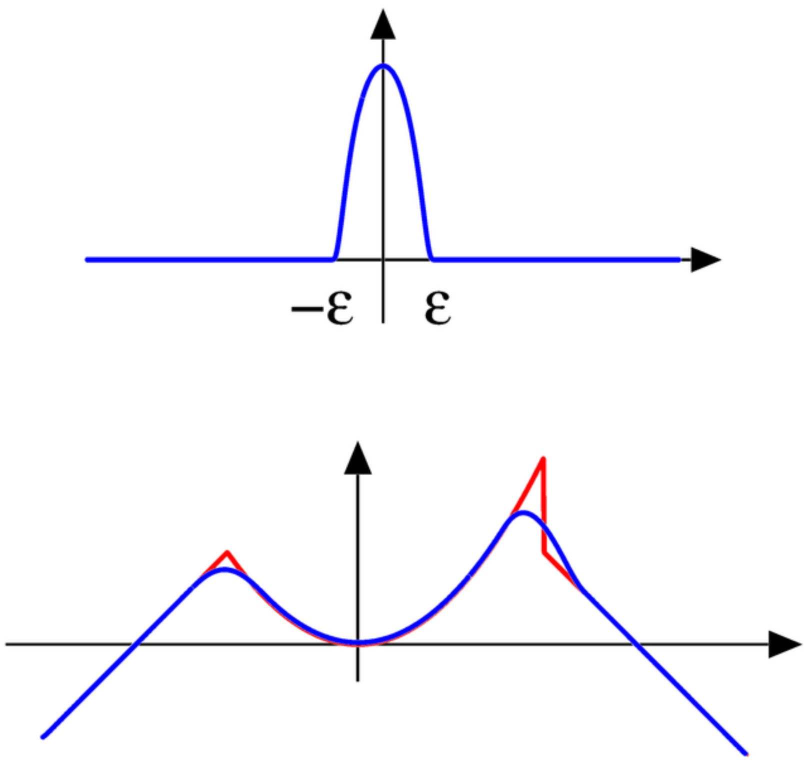

Figure 3: On top, the mollifier; on bottom, a jagged function (red) being mollified by the mollifier on top, and the smoothed out function (blue) after mollification (picture from http://en.wikipedia.org/wiki/Mollifier).

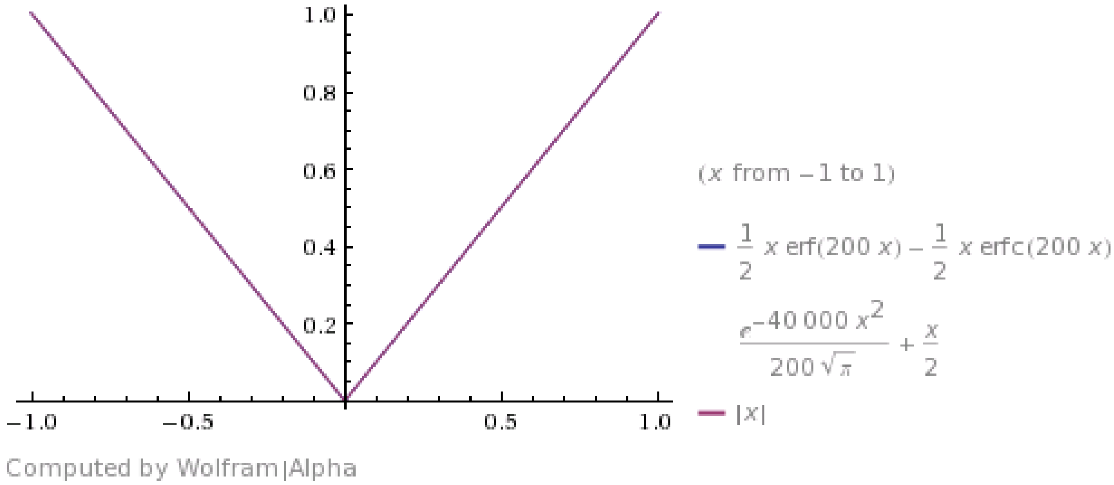

Figure 4: A graph of the convolution of $|x|$ and $\frac{200}{\sqrt{\pi}}e^{-(200x)^2}$ (blue), superimposed with the graph of $|x|$ (red).

You can't even see the difference! However, the convoluted function is smooth at the bottom.

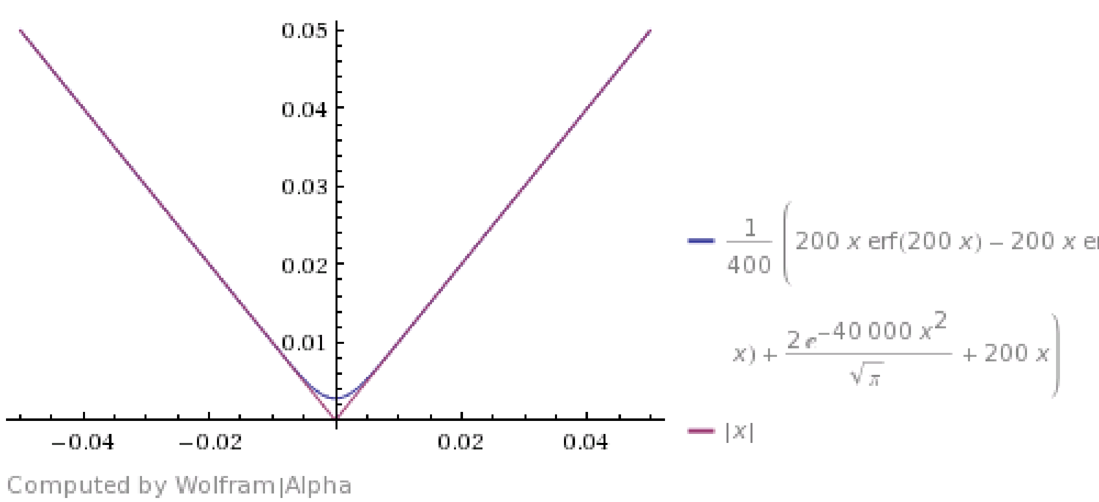

Figure 5: A zoomed in graph of Figure 4. We now see that the graphs agree almost exactly except for very near 0, where one is smooth. Also note that $\epsilon = \frac{1}{200}$ here, which isn't even that small. We can get a much better approximation by making ϵ much smaller.

Fourier Analysis

We define the Schwartz space $S \subset C^{\infty}$ as the set of functins $f \in C^{\infty}(\mathbb{R}^n)$ which satisfy

$$\|f\| \equiv \sup_{x\in\mathbb{R}^n} |x^{\alpha}\partial^{\beta} f(x)| < \infty$$

for all $\alpha, \beta \in \mathbb{N}^n$. Note that $\|\cdot\|$ defines a seminorm. If $x = (x_1, ..., x_n) \in \mathbb{R}^n$, we set $|x| = \|x\|_{l^2} = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}}$. As an example, the function $f(x) = e^{-\frac{|x|^2}{2}}$ belongs to S, as f(x) and its derivatives go to zero faster than any polynomial. Now we define a continuous linear mapping called the Fourier transform, $\mathscr{F}\colon S \to S$,

$$\hat{u}(\xi) \equiv \mathscr{F}(u(x)) = \int e^{-ix\xi} u(x)\, dx, \qquad (1)$$

where $x\xi$ is understood to be the dot product $x \cdot \xi$. The Fourier transform $\mathscr{F}\colon S \to S$ has the following easy to verify properties:
$\widehat{D_j u}(\xi) = \xi_j \hat{u}(\xi)$,
$\widehat{\tau_y u}(\xi) = e^{iy\xi}\hat{u}(\xi)$ where $\tau_y u(x) = u(x+y)$,
$\widehat{x_j u}(\xi) = -D_j \hat{u}(\xi)$
$(\widehat{e^{-ix\nu} u})(\xi) = \tau_{\nu}\hat{u}(\xi)$.

A linear operator on S which is continuous with respect to the semi-norm is called a tempered distribution in $\mathbb{R}^n$, and is denoted S'. By defining $\langle u, \cdot\rangle : S \to \mathbb{R}$ for $u \in S$ by

$$\langle u, v\rangle = \int u(x)v(x)\, dx,$$

we have that $S \subset S'$ (meaning S is isomorphic to a subset of S'), and in fact it is dense.

For $u, v \in S$, we have that

$$\langle \hat{u}, v\rangle = \int \hat{u}(\xi) v(\xi)\, d\xi = \int \left(\int e^{ix\xi} u(x)\, dx \right) v(\xi)\, d\xi = \int u(x) \left(\int e^{ix\xi} v(\xi)\, d\xi \right) dx$$

$$= \int u(x)\hat{v}(x)\, dx = \langle u, \hat{v}\rangle,$$

where Fubini's theorem was used in the third equality. So we see that $\langle \hat{u}, v\rangle = \langle u, \hat{v}\rangle$ for all $u, v \in S$. Thus for $u \in S'$, $v \in S$, we see that the formula

$$\langle \hat{u}, v\rangle = \langle u, \hat{v}\rangle$$

defines a mapping $\mathscr{F}\colon S' \to S'$ and is the unique continuous extention of $\mathscr{F}\colon S \to S'$, and it satisfies the properties given on the previous page. Note that if we restrict $\mathscr{F}$ to $L^1(\mathbb{R}^n)$, then for $u \in L^1(\mathbb{R}^n)$, $\hat{u}$ is given by (1). Now we will derive an inversion result:
From the property that $\widehat{D_j u}(\xi) = \xi_j \hat{u}(\xi)$, we see that

$$0 = \xi_j \hat{1}(\xi) \implies \hat{1}(\xi) = c\delta(\xi)$$

from some $c \in \mathbb{C}$. Using this and the fact that $\hat{\delta} = 1$ (easy to see from definitions), we can see that for $u \in S$,

$$\hat{\hat{u}}(0) = \langle \delta, \hat{\hat{u}}\rangle = \langle 1, \hat{u}\rangle$$

$$= \langle \hat{1}, u\rangle = c\hat{u}(0),$$

so that we just just need to choose some u to find out the constant. It turns out $c = (2\pi)^n$. Now

$$\hat{\hat{u}}(0) = cu(0) \implies \tau_{-y}\hat{\hat{u}}(0) = \tau_{-y} cu(0) = cu(-y)$$

and by the fourth propery this means

$$\widehat{e^{i\xi y}\hat{u}}(0) = cu(-y)$$

and by the second property this means

$$\widehat{\widehat{\tau_y u}}(0) = cu(-y)$$

$$\implies \hat{\hat{u}}(y) = cu(-y)$$

so plugging in our value for c and taking $y \to x$ and rearranging, we see that

$$u(-x) = \frac{1}{(2\pi)^n}\hat{\hat{u}}(x).$$

We can rewrite this expression using the explicit formula for the Fourier transform:

$$u(-x) = \frac{1}{(2\pi)^n}\int e^{-ix\xi}\hat{u}(\xi)\, d\xi \implies u(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi}\hat{u}(\xi)\, d\xi\,.$$

This is known as the Fourier inversion theorem, it is the formula for the inverse Fourier transform, denoted by $\mathscr{F}^{-1}$, and it maps $\hat{u}$ to u.

Now let $u, v \in S$. From the top of page 10 we know that $\langle \hat{u}, \hat{v}\rangle = \langle u, \hat{\hat{v}}\rangle$, so using the inner product $(\cdot,\cdot)$ associated with $L^2(\mathbb{R}^n)$ combined with the Fourier inversion formula, we see that $(\hat{u}, \hat{v}) = (2\pi)^n(u, v)$. Evidently if we extend the domain of the Fourier transform to the square integrable functions, then $\mathscr{F}\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, ie. is an automorphism (since it is also an isomorphism), and that $(2\pi)^{\frac{-n}{2}}\mathscr{F}$ is unitary. This is known as Plancherel's theorem.

Pseudo – differential Operators

Since $P(D)u(x) = \frac{1}{(2\pi)^n}\int P(\xi)_j e^{ix\xi}\hat{u}(\xi)\,d\xi$, we can define a pseudo-differential operator $a(D) \in S'(\mathbb{R}^n)$ by $\widehat{a(D)u(\xi)} \equiv a(\xi)\hat{u}(\xi)$ ($\widehat{a(D)u(\xi)}$ is smoothand slowly increasing). Then we have that

$$a(D)u(x) = \frac{1}{(2\pi)^n}\int a(\xi)e^{ix\xi}\hat{u}(\xi)\,d\xi\,.$$

For instance, letting $a(\xi) = i\xi$, we get that $a(D)$ is just the usual differentiation operator. However letting $a(\xi) = i\sqrt{\xi}$, we get that $a(D)$ is a half-differentiation operator.
We have the basic property that a(D)b(D)=(ab)(D).

Consider the Laplacian operator $\Delta = \partial_1^2 + \cdots + \partial_n^2$. Its symbol is

$$a(\xi) = -|\xi|^2.$$

Let $\omega \in S$ (ω is called a parametrix), δ be the dirac delta at 0, then we can solve the distribution equation

$$\Delta E = \delta + \omega$$

by using Fourier transforms: Let $\hat{E}(\xi) = -\frac{1-\chi(\xi)}{|\xi|^2}$, then

$$\widehat{\Delta E}(\xi) = -|\xi|^2\hat{E}(\xi) = 1 - \chi(\xi),$$

and this distribution is smooth away from 0. Now if $f \in S'$

$$\Delta(E * f) = f + \omega * f,$$

and so the distribution $v = E * f$ is an approximate solution to the equation

$$\Delta v = f.$$

Also, sing supp v= sing supp f, since $\Delta v = f + \omega * f$, where $\omega * f \in C^\infty$, so f is smooth where v is, and if f is smooth near some point x_0, then $v = E * f = E * (\chi f) + E * (1 - \chi)f$, ($\chi$ is equal to 1 near x_0, and χf is smooth there). So $E * \chi f \in C^\infty$, and

$$(E * (1 - \chi)f)(x) = \int E(x - y)(1 - \chi(y))f(y)\, dy$$

only has $x - y$ away from 0 if x is sufficiently close to x_0. Thus $u(x) \in C^\infty$ for x sufficiently close to x_0. As a matter of fact, we can conclude that any solution of $\Delta v = f$ has the property that sing supp v=sing supp f. If f is smooth near x_0, and χ is smooth and equals 1 near x_0, then $\Delta \chi v = f$ near x_0, and so is smooth near x_0. So since χv and $\Delta \chi v$ are in S', we see that

$$E * \Delta(\chi v) = \chi v + \omega * \chi v = \chi v + \text{something in } C^\infty,$$

and so from before we see that $\chi v \in C^\infty$ near x_0.

Non – Constant Coefficient Operators :

For $P = \sum a_\alpha D_x^\alpha$, $a_\alpha \in S$, we have the formula

$$Pu(x) = (2\pi)^{-n} \int e^{ix\xi} p(x, \xi)\hat{u}(\xi)\, d\xi$$

$$p(x, \xi) = \sum a_\alpha(x)\xi^\alpha.$$

Symbols

Definition : Let m $\in \mathbb{R}$. Let $\mathbb{S}^m = \mathbb{S}^m(\mathbb{R}^n \times \mathbb{R}^n)$ be the set of all $a \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ with the property that for all α, β,

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leqslant C_{\alpha,\beta}(1+|\xi|)^{m-|\beta|}.$$

We denote $\mathbb{S}^{-\infty} = \bigcap_m \mathbb{S}^m$. Elements of $\mathbb{S}^m$ are called symbols of order m.

Example 1 : The funcion $a(x,\xi) = e^{ix\xi}$ is not a symbol.

Example 2 : For $f \in S$, $f(\xi)$ is a symbol of order $-\infty$.

Propertes :
1) $a \in \mathbb{S}^m \implies \partial_x^\alpha \partial_\xi^\beta a \in \mathbb{S}^{m-|\beta|}$,
2) $a \in \mathbb{S}^m$ and $b \in \mathbb{S}^k \implies ab \in \mathbb{S}^{m+k}$,
3) $a \in \mathbb{S}^m \implies a \in S'(\mathbb{R}^{2n})$.
Lemma 1 : If $a_1, ..., a_k \in \mathbb{S}^0$, and $F \in C^\infty(\mathbb{C}^k)$, then $F(a_1, ...a_k) \in S^0$.

Proof. We may assume without loss of generality that a_i are real and that $F \in C^\infty(\mathbb{R}^k)$ since the real and imaginary parts of a_i are in $\mathbb{S}^0$. Now

$$\frac{\partial}{\partial x_j} F(a) = \frac{\partial F}{\partial a_i} \frac{\partial a_i}{\partial x_j} \quad (1)$$

$$\frac{\partial}{\partial \xi_j} F(a) = \frac{\partial F}{\partial a_i} \frac{\partial a_i}{\partial \xi_j} \quad (2)$$

we Einstein summation notation is been emplored. We proceed by induction.

If $|\alpha| + |\beta| = 0$, it is clear that the estimate holds. Now suppose it is true for $|\alpha| + |\beta| \leqslant 0, 1, ...p$, and consider the case $|\alpha| + |\beta| \leqslant p + 1$. By an application of the Leibniz differentiation formula to (1) and (2), and the induction hypothesis applied to the derivatives of $\frac{\partial F}{\partial a_i}(a)$, we get the desired result. $\square$

Semi – norm:

We define the semi-norm on $\mathbb{S}^m$ by

$$|a|^m_{\alpha,\beta} = \sup_{(x,\xi)\in\mathbb{R}^n\times\mathbb{R}^n} \big\{(1 + |\xi|)^{-(m-|\beta|)}|\partial^\alpha_x \partial^\beta_\xi a(x, \xi)|\big\}.$$

Convergence $a_n \to a$ means that for all α, β, $|a_n - a|^m_{\alpha,\beta} \to 0$ as $n \to \infty$. With this semi-norm, we have a complete space (a Frechet space).

Approximation Lemma: Let $a \in \mathbb{S}^0(\mathbb{R}^n \times \mathbb{R}^n)$ and set $a_\epsilon(x, \xi) = a(x, \epsilon\xi)$. Then a_ϵ is bounded in $\mathbb{S}^0$, and $a_\epsilon \to a_0$ as $\epsilon \to 0$ in $\mathbb{S}^m$ for all $m > 0$.

Proof. Let $0 \leqslant \epsilon, m \leqslant 1$, and α, β be abritrary. For $\beta = 0$,

$$\partial^\alpha_x(a_\epsilon - a_0) = \int_0^1 \partial_t\partial^\alpha_x a(x, t\epsilon\xi)\, dt = \int_0^{\epsilon\xi} \partial_s\partial^\alpha_x a(x, s)\, ds,$$

with $s = \epsilon\xi t$. Thus

$$|\partial^\alpha_x(a_\epsilon - a_0)| \leqslant \int_0^{\epsilon\xi} |\partial_s\partial^\alpha_x a(x, s)|\, ds \leqslant \int_0^{\epsilon\xi} C\frac{ds}{1 + |s|} = C \log(1 + \epsilon|\xi|).$$

So we get that

$$|\partial^\alpha_x(a_\epsilon - a_0)| \leqslant C \log(1 + \epsilon|\xi|),$$

and since $\log(1+x)|_{x=0} \leqslant (1+x)^m|_{x=0}$, and $\frac{1}{1+x} \leqslant C_m m(1+x)^{m-1}$ for $.x \geqslant 0$, we see that $\log(1 + x) \leqslant C_m(1 + x)^m$, and this gives the desired result.

Now for $\beta \neq 0$, $\partial_x^\alpha \partial_\xi^\beta a_0 = 0$, and

$$|\partial_x^\alpha \partial_\xi^\beta a_\epsilon| \leqslant C_{\alpha,\beta}\epsilon^{|\beta|}(1+\epsilon|\xi|)^{|\beta|}$$

then since

$$C_{\alpha,\beta}\epsilon^{|\beta|}(1+\epsilon|\xi|)^{|\beta|} \leqslant C_{\alpha,\beta}(1+|\xi|)^{|\beta|},$$

we have the result. □

Asymptotic Sums:
Let $a_j \in \mathbb{S}^{m_j}$ for a decreasing sequence $m_j \to -\infty$. Generally $\sum_{j=0}^N a_j$ does not converge as $N \to \infty$, but we can still give meaning to the series. We will write

$$a \sim \sum a_j$$

if for all $N \geqslant 0$,

$$a - \sum_{j=0}^{N} \in \mathbb{S}^{m_{N+1}}.$$

Borel Lemma: Let (b_j) be a sequence of complex numbers. There exists a function $f \in C^\infty(\mathbb{R})$ such that for all j, $f^{(j)}(0) = b_j$, so that $f(x) \sim \sum_j b_j \frac{x^j}{j!}$ when $x \to 0$.

Proof. Let χ be a C^∞ function equal to 1 for $|x| \leqslant 1$ and 0 for $|x| \geqslant 2$. Let (λ_j) be a sequence of positive numbers tending to ∞. We will show that pick (λ_j) so that the function defined by

$$f(x) = \sum_j b_j \frac{x^j}{j!}\chi(\lambda_j x)$$

has the desired properties. First off, the series converges pointwise. Let

$N \in \mathbb{N}$. If $j \geqslant N$, then the N^{th} derivative of the j^{th} term is equal to

$$f_j^N(x) = \sum_{0 \leqslant i \leqslant N} \binom{N}{i} b_j \frac{x^{j-i}}{(j-i)!} x^{N-i}(\lambda_j x)\lambda_j^{N-i}.$$

Now remember that the support of χ is contained in $|x| < 2$, so that $\lambda_j x$ is bounded in the supports of χ and its derivatives. Thus there is a constant C_N such that

$$|f_j^{(N)}(x)| \leqslant \frac{C_N |b_j| \lambda_j^{N-j}}{(j-N)!}.$$

Thus if we pick $\lambda_j \leqslant 1 + |b_j|$, then the series $\sum_j |f_j^(N)(x)|$ is uniformly convergent for $x \in \mathbb{R}$, so that $f \in C^\infty$, and that its derivatives are obtained from term by term differentiation, and that

$$f^N(0) = b_N.$$

$\square$

Proposition: There exists an $a \in \mathbb{S}^{m_0}$ such that $a \sim \sum_j a_j$, and supp $a \subset \bigcup_j$supp a_j (proof omitted, see reference (1)).

Definition : A symbol $a \in \mathbb{S}^m$ is said to be classical if $a \sim \sum_j a_j$, where a_j are homogeneous functions of degree $m - j$ for $|\xi| \geqslant 1$, ie. $a_j(x, \lambda\xi) = \lambda^{m-j} a_j(x, \xi)$ for $|\xi|,\ \lambda \geqslant 1$.

Pseudo – differential Operators in Schwartz Space

Proposition : If $a \in \mathbb{S}^m$ and $u \in S$, then the formula

$$\mathrm{Op}(a)u(x) = \frac{1}{(2\pi)^n} \int e^{ix\xi} a(x, \xi) \hat{u}(\xi)\, d\xi$$

defines a function on S, and the mapping $(a, u) \to$Op$(a)u$ is continuous. This operator Op from $\mathbb{S}^m$ to the linear operators on S is injective and satisfies the comutation relations

$$[\mathrm{Op}(a), D_j] = i\mathrm{Op}(\partial_{x_j}, a),$$

$$[\mathrm{Op}(a), x_j] = -i\mathrm{Op}(\partial_{x_j}, a).$$

Proof. First off, since $\hat{u} \in S$ and $a \in \mathbb{S}^m$, we have that

$$|\mathrm{Op}(a)u(x)| \leqslant \frac{1}{(2\pi)^n}\int |a(x,\xi)||\hat{u}(\xi)|\, d\xi\big| = \frac{1}{(2\pi)^n}\int |a(x,\xi)|(1+|\xi|)^{-m}(1+|\xi|)^m|\hat{u}(\xi)|\, d\xi$$

$$\leqslant \frac{1}{(2\pi)^n}\sup_{\xi}\{|a(x,\xi)|(1+|\xi|)^{-m}\}\int (1+|\xi|)^m|\hat{u}(\xi)|\, d\xi,$$

and so $\mathrm{Op}(a)u$ is bounded..

Now for the commutation relations:

$$\mathrm{Op}(a)D_j u(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)\widehat{D_j u}(\xi)\, d\xi$$

$$= \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)\xi_j\hat{u}(\xi)\, d\xi,$$

where in the second equality we have used a property of Fourier transforms earlier discussed. Now

$$D_j(\mathrm{Op}(a))(x) = -i\frac{1}{(2\pi)^n}\int e^{ix\xi}i\xi_j a(x,\xi)\hat{u}(\xi)\, d\xi - i\mathrm{Op}(\partial_{x_j}a)u(x),$$

and so from the these last two formulas we see the first commutation relation. For the second commutation relation,

$$\mathrm{Op}(a)x_j u(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)\widehat{x_j u}(\xi)\, d\xi = \frac{1}{(2\pi)^n}\int -e^{ix\xi}a(x,\xi)D_{\xi_j}\hat{u}(\xi)\, d\xi$$

where in the second equality we have again used a property of the Fourier transform discussed earlier. Now

$$x_j(\mathrm{Op}(a)u)(x) = x_j\frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)\hat{u}(\xi)\,d\xi$$

$$= \frac{1}{(2\pi)^n}\int (D_{\xi_j}e^{ix\xi})a(x,\xi)\hat{u}(\xi)\,d\xi$$

$$= \frac{1}{(2\pi)^n}\Big(\int D_{\xi_j}(e^{ix\xi}a(x,\xi)\hat{u}(\xi))\,d\xi - \int e^{ix\xi}D_{\xi_j}a(x,\xi)\hat{u}(\xi)\,d\xi - \int e^{ix\xi}a(x,\xi)D_{\xi_j}\hat{u}(\xi)\,d\xi\Big).$$

Now with the fundamental theorem of calculus, we see that in the above expression, the integral on the left is 0 since $\hat{u} \in L^2(\mathbb{R}^n)$ (since $u \in S \subset L^2$, and the Fourier transform sends L^2 funtions to L^2 functions, and so $\hat{u}(\xi)$ goes to 0 at infinity. Remember that the integrals are over $\mathbb{R}^n$), so that we are left with

$$x_j(\mathrm{Op}(a)u)(x) = -\frac{1}{(2\pi)^n}\int e^{ix\xi}D_{\xi_j}a(x,\xi)\hat{u}(\xi)\,d\xi - \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)D_{\xi_j}\hat{u}(\xi)\,d\xi,$$

$$= i\mathrm{Op}(\partial_{x_j}a)u(x) - \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)D_{\xi_j}\hat{u}(\xi)\,d\xi,$$

and so we can see the second commutation relation. The commutation relations imply that $x^\alpha D^\beta(Op(a)u)$ is a linear combination of the terms

$$\mathrm{Op}(\partial_\xi^{\alpha'}\partial_x^{\beta'})(x_{\alpha''}D^{\beta''}u), \mathrm{with} \alpha' + \alpha'' = \alpha,\ \beta' + \beta'' = \beta.$$

Thus $x^\alpha D^\beta(Op(a)u)$ is bounded by the product of a semi-norm of $u \in S$ and by a semi-norm of $a \in \mathbb{S}^m$, hence is continuous. All that's left is to prove injectivity. Suppose that for all $u \in S$ and for all $x \in \mathbb{R}^n$, we have

$$\int e^{ix\xi}a(x,\xi)\hat{u}(\xi)\,d\xi = 0.$$

Fix x. then the function b defined as

$$b(\xi) = \frac{a(x,\xi)}{(1+|\xi|^2)^{\frac{m}{2}+\frac{n}{4}+\frac{1}{2}}}$$

is in $L^2(\mathbb{R}^n)$ and is orthogonal to all functions of the form

$$\varphi(\xi) = e^{-ix\xi}(1+|\xi|^2)^{\frac{m}{2}+\frac{n}{4}+\frac{1}{2}}\overline{\hat{u}(\xi)},$$

and if u is in S then so is φ, and so $b = 0$ by the density of S in L^2. $\square$

Kernel: Let $a \in \mathbb{S}^{-\infty}$. Then for $u \in S$, we have

$$\operatorname{Op}(a)u(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)\hat{u}(\xi)\,d\xi$$

$$= \frac{1}{(2\pi)^n}\int e^{ix\xi}a(x,\xi)\,d\xi \int e^{-iy\xi}u(y)\,dy$$

$$= \frac{1}{(2\pi)^n}\int u(y)\,dy \int e^{i(x-y)\xi}a(x,\xi)\,d\xi$$

where we have used Fubini's theorem in the third equality. So we see that the kernel K of $\operatorname{Op}(a)$ is

$$K(x,y) = \frac{1}{(2\pi)^n}\int e^{i(x-y)\xi}a(x,\xi)\,d\xi = \frac{1}{(2\pi)^n}(\mathscr{F}_\xi a)(x, y-x),$$

where $\mathscr{F}_\xi$ means the Fourier transform with respect to ξ.

Adjoints

For an arbitrary operator $A : S \to S$, we want an operator $A^* : S \to S$ such that for all$u, v \in S$,

$$(Au, v) = (u, A^*v).$$

By a density argument if A^* exists then it is unique, and it is called the adjoint of A. Should A^* exist, then we can define $A : S' \to S'$ by the formula

$$(Au, v) = (u, A^* v)$$

for all $u \in S'$, $v \in S$, where $(u, v) = \langle u, \bar{v}\rangle$. This means that we can rewrite the definition of A as

$$\langle Au, v\rangle = \langle u, \overline{A^* \bar{v}}\rangle.$$

Example 1 : Let $P = \sum_{|\alpha| \leqslant m} a_\alpha(x) D^\alpha$ be a differential operator with slowly increasing smooth coefficients. Then

$$(D_j u, v) = \int D_j u\, \bar{v} = -i \int \partial_j u\, \bar{v}$$

$$= i \int u \partial_j \bar{v} = \int u \overline{D_j v} = (u, D_j v),$$

where we have used integration by parts in the third equality, and the fact that $u, v \in S$ to conclude that the boundry term is zero. Since the coefficients are slowly increasing, we conclude that $P^* v = \sum_{|\alpha| \leqslant m} D^\alpha (\bar{a}_\alpha v)$. The fact that $(Dju, v) = (u, Djv)$ is extremely important in quantum mechanics, where all observables (quantities that can be measured) are represented by hermitian operators O (and hence satisfy $O = O^*$), and where $\hbar Dj$ represents the momentum operator for the j^{th} coordinate, and $\hbar$ is the reduced Planck's constant.

Example 2 : Let $a(D)$ be a pseudo-differential operator with constant coefficients. Then for $u, v \in S$, we have

$$(a(D)u, v) = \frac{1}{(2\pi)^n}(a\hat{u}, \hat{v}) = \frac{1}{(2\pi)^n}(\hat{u}, \bar{a}\hat{v}) = (u, \bar{a}(D)v),$$

so that $a(D)^* = \bar{a}(D)$.

Now let's show that if A^* exists, we can wrtie K^* using the kernel K of A:

$$\langle K(x,y), u(y)v(x)\rangle = \langle Au, v\rangle = \langle u, \overline{A^*\bar{v}})$$

$$\overline{\langle \bar{u}, A^*\bar{v}\rangle} = \overline{\langle K^*(y,x), \bar{v}(x)\bar{u}(y)\rangle}$$

$$\implies K^*(y,x) = \overline{K(x,y)}\,.$$

Now in general we would like to find the adjoint of pseudo-differential operators. To do this it is enough to check if K is the kernal of the symbol a, then the operator with kernel K^* sends Shwartz functions to Schwartz functions. Now we will assume that the symbol a (thus a^* as well) is in $S(\mathbb{R}^{2n})$ and then extend it to $S'(\mathbb{R}^{2n})$ by continuity. We have

$$K^*(x,y) = \overline{K(y,x)} = \frac{1}{(2\pi)^n}\int e^{i(x-y)\xi}\overline{a(y,\xi)}\,d\xi$$

and

$$a^*(x,\xi) = \int K^*(x, x-y)e^{-iy\xi}\,dy = \frac{1}{(2\pi)^n}\int e^{iy(\nu-\xi)}\bar{a}(x-y,\nu)\,dyd\nu$$

$$= \frac{1}{(2\pi)^n}\int e^{-iy\nu}\bar{a}(x-y,\xi-\nu)\,dyd\nu,$$

so we have found our formula for a^*.

The following two theorems are fundamental to symbolic calculus, and will be stated without proof (see reference (1)).

Theorem 1 : If $a, \in S^m$, then $a^* \in S^m$ and

$$a^*(x,\xi) \sim \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha D_x^\alpha \bar{a}(x,\xi)\,.$$

Particularly, if $A = Op(a)$ is a pseudo-differential operator of order m, then $A^* = Op(a^*)$ is a pseudo-differential operator of order m, and thus A extends to an operator from $S'(\mathbb{R}^n)$ to $S'(\mathbb{R}^n)$.

Theorem 2 : If $a_1 \in S^{m_1}$, $a_2 \in S^{m_2}$, then $Op(a_1)Op(a_2) = Op(b)$ where

$$b(x,\xi) = \frac{1}{(2\pi)^n} \int e^{-i(x-y)(\xi-\nu} a_x(x\nu) a_2(y,\xi)\, dyd\nu\,,$$

and we write $b = a_1 \# a_2 \in S^{m_1+m_2}$ (# is just notation representing the symbol that results from multipliying two operators, ie. $a(x,D)b(x,D) = (a\#b)(x,D))$, and $b \sim \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha a_1 D_x^\alpha a_2$.

Fun with Pseudo – differential Operators

For the sake of brevity, the functions in this section will be assumed to be sufficiently nice for whatever is written to make sense.

Example 1: Consider the Laplacian, and some function u. We have that

$$\widehat{\Delta u}(\xi) = -|\xi|^2 \hat{u}(\xi),$$

so we can define the square root of the Laplacian by the property that it satisfies

$$\widehat{\sqrt{\Delta} u}(\xi) = i|\xi|\hat{u}(\xi). \tag{1}$$

Taking inverse Fourier transforms, we have that

$$\sqrt{\Delta}u(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi} i|\xi|\hat{u}(\xi)\, d\xi. \quad (2)$$

We would hope that by defining $\sqrt{\Delta}$ this way, that $\sqrt{\Delta}\sqrt{\Delta}{=}\Delta$ (otherwise what is the point?), let's double check: looking at (2), we see that

$$\sqrt{\Delta}(\sqrt{\Delta}u)(x) = \frac{1}{(2\pi)^n}\int e^{ix\xi} i|\xi|\widehat{\sqrt{\Delta}u}(\xi)\, d\xi,$$

and using (1), we get that

$$\sqrt{\Delta}\sqrt{\Delta}u(x) = -\frac{1}{(2\pi)^n}\int e^{ix\xi}|\xi|^2\, d\xi,$$

which is the correct equation for Δu.
In fact we can define derivatives of arbitrary order this way, consider the differential operator in one dimension $\frac{d^n}{dx^n}$ for $n \in \mathbb{N}$:

$$\frac{d^n}{dx^n}u = \frac{1}{(2\pi)}\int\int e^{i(x-y)\xi}(i\xi)^n u(y)\, dy\, d\xi,$$

so that for $s \in \mathbb{C}$, we can define the fractional differential operator $\frac{d^s}{dx^s}$ by

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

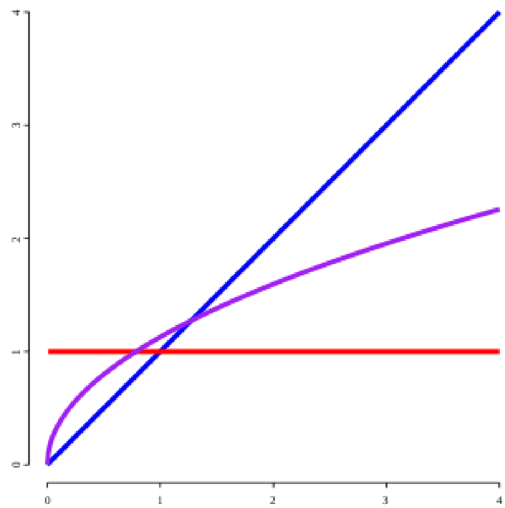

Figure 6: A graph of $f(x) = x$ (blue), its first derivative (red), and its half-derivative $\left(\frac{2}{\sqrt{\pi}}\sqrt{x}\text{ - in purple}\right)$ (picture from http://en.wikipedia.org/wiki/Fractional_calculus)).

Pseudo-differential operators are very important in relativistic quantum mechanics, where Dirac found his equation (Dirac equation) describing relativistic quantum mechanics by factoring the Laplacian: for massless particles, $E^2 = p^2c^2$, where E is energy, p is momentum, and c is the speed of light (if you don't know quantum mechanics, just take this at face value). Writing these as operators, $p = -i\hbar\nabla$, so that $E^2 = -c^2\hbar^2\Delta$, and so

$$E = \hbar c\sqrt{-\Delta}\,.$$

In $\mathbb{R}^2$, the Dirac operator D, is defined by

$$D = -i\sigma_x\partial_x - i\sigma_y\partial_y,$$

where

$$\sigma_x = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \ \sigma_y = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$$

are known as the Pauli matrices. All of these operator act on wavefunctions, $\psi : \mathbb{R}^2 \to \mathbb{C}^2$,

$$\psi(x, y) = \begin{pmatrix} \chi(x, y) \\ \varphi(x, y) \end{pmatrix},$$

which describe the spin of electrons (top row is the probability amplitude that an electron will be found to be spin up when measured, and the bottom row is the probability amplitude that the electron will be found to be spin down when measured). Using the matrix form it is easy to verify that $D^2\psi = -\Delta\psi$, so that

$$D = \sqrt{-\Delta}\,.$$

References

(1) Alinhac, Serge. Gerard, Patrick. Pseudo-differential Operators and the Nash-Moser Theorem. AMS.

(2) Friedrich, Thomas (2000), Dirac Operators in Riemannian Geometry, American Mathematical Society, ISBN 978-0-8218-2055-1

(3) http://www.quora.com/Quantum-Field-Theory/How-and-when-do-physicists-use-the-Dirac-equation