Distribution Theory (and its Applications)

Lecturer: Anthony Ashton

Scribe: Paul Minter

Michaelmas Term 2018

These notes are produced entirely from the course I took, and my subsequent thoughts. They are not necessarily an accurate representation of what was presented, and may have in places been substantially edited. Please send any corrections to pdtwm2@cam.ac.uk

Recommended book: Hörmander, *Analysis of linear partial differential operators*, Volume 1

Contents

0. Motivation

We start with some examples of motivate the things we will be talking about in the course.

Example 0.1 (Derivative of Dirac Delta?).

We often define the "Dirac Delta", via requring

$$\int_{-\infty}^{+\infty} \delta(x-x_0)f(x)\,\mathrm{d}x = f(x_0) \qquad \textit{for all suitable } f.$$

No such function exists however in terms of Lebesgue/Riemann integrability. But let us suppose it did. Then what would δ', its derivative, be? Well, we could differentiate the above under the integral sign, i.e.

$$\begin{aligned}\int_{-\infty}^{\infty} \delta'(x-x_0)f(x)\,\mathrm{d}x \text{ “ = ” } & \lim_{h\to 0}\int_{-\infty}^{\infty} \frac{\delta(x-x_0+h)-\delta(x-x_0)}{h}\cdot f(x)\,\mathrm{d}x \\ &= \lim_{h\to 0}\frac{1}{h}\int_{-\infty}^{\infty} \delta(x-(x_0-h))f(x)-\delta(x-x_0)f(x)\,\mathrm{d}x \\ &= \lim_{h\to 0}\frac{1}{h}[f(x_0-h)-f(x_0)] \\ &= -f'(x_0)\end{aligned}$$

or in other words,

$$\int_{-\infty}^{\infty} \delta'(x-x_0)f(x)\,\mathrm{d}x = -\int_{-\infty}^{\infty} \delta(x-x_0)f'(x)\,\mathrm{d}x.$$

This suggests the integration by parts formula holds.

Example 0.2 (Fourier Trasnforms of Polynomials?). *The standard definition of the Fourier transform is*

$$\hat{f}(\lambda) := \int_{-\infty}^{\infty} e^{-i\lambda x} f(x)\,\mathrm{d}x \qquad \textit{for } \lambda\in\mathbb{R}.$$

If f is absolutely integrable (i.e. $f\in L_1(\mathbb{R})$), then

$$|\hat{f}(\lambda)| \le \int_{-\infty}^{\infty} \left|e^{-i\lambda x}f(x)\right|\,\mathrm{d}x = \int_{-\infty}^{\infty} |f(x)|\,\mathrm{d}x = \|f\|_{L_1(\mathbb{R})}.$$

But what if $f \not\to 0$ as $|x|\to\infty$? (e.g. if not in $L_1(\mathbb{R})$). You might have come across the 'equality'

$$\delta(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\lambda x}\,\mathrm{d}x$$

before, although this integral does not make sense. However if 'true', this would imply that

$$\hat{1}(\lambda) = 2\pi\delta(\lambda)$$

i.e. the Fourier transform of 1 is related to this mysterious 'function' δ. But then using this, what would $\hat{f}(\lambda)$ be, when $f(x)=x^n$? Well, we have

$$\int_{-\infty}^{\infty} x^n e^{-i\lambda x}\,\mathrm{d}x = \int_{-\infty}^{\infty}\left(i\frac{\mathrm{d}}{\mathrm{d}\lambda}\right)^n e^{-i\lambda x}\,\mathrm{d}x = \left(i\frac{\mathrm{d}}{\mathrm{d}\lambda}\right)^n[2\pi\delta(\lambda)] = 2\pi(i)^n\delta^{(n)}(\lambda).$$

So if we could understand the derivatives of δ, we could potentially make sense of Fourier transforms of polynomials.

Alternatively we could invoke Parseval's theorem, which says for all sufficiently nice f, g,

$$\int_{-\infty}^{\infty} f(\lambda)\hat{g}(\lambda) = \int_{-\infty}^{\infty} \hat{f}(\lambda)g(\lambda)\ \mathrm{d}\lambda.$$

We could use this to extend the Fourier transform to more general functions by requiring this to hold. E.g. if $f(x) = x$, then we could define $\hat{f}$ by requiring that, for all "nice" g, we have

$$\int_{-\infty}^{\infty} \hat{f} g(\lambda) = \int_{-\infty}^{\infty} \underbrace{\lambda}_{=f(\lambda)} \hat{g}(\lambda)\ \mathrm{d}\lambda.$$

So to compute $\hat{f}$ we would need to find the function such that this holds.

Example 0.3 (Discontinuous Solutions to PDE). *We often want solutions to certain PDEs to have discontinuities. For example, an explosion creates a shockwave. The pressure will be discontinuous over the shock wave, and we want the solution to reflect this.*

With one spatial dimension, we would want the pressure $p = p(x,t)$ to satisfy the wave equation,

$$(\star) \qquad \Box p := \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} - \frac{\partial^2 p}{\partial x^2} = 0.$$

[Wlog set $c = 1$.] So assume that $p \in C^2(\mathbb{R}^2)$. Then we want to make sense of the PDE $\Box p = 0$ for a larger family of functions, and not just C^2 ones. By integration the above we get,

$$0 = \int_{\mathbb{R}^2} f(x,t)\Box p\ \mathrm{d}x\mathrm{d}t \qquad \forall f \in C_c^2(\mathbb{R}^2).$$

If we integrate by parts we then get

$$0 = \int_{\mathbb{R}^2} p(x,t)\Box f\ \mathrm{d}x\mathrm{d}t$$

*where all the boundary terms vanish due to the compact support of f. But now this equations makes sense if p is just integrable, e.g. if p was just continuous. This gives the notion of a weak solution: We say $p = p(x,t)$ is a **weak solution** to $(\star)$ if:*

$$0 = \int_{\mathbb{R}^2} p\Box f\ \mathrm{d}x\mathrm{d}t \qquad \forall f \in C_c^2(\mathbb{R}^2).$$

Note: In each of the above examples, we have to introduce a space of auxiliary ("nice") functions so we could extend the range of some classical definitions. This is the essence of distribution theory.

In distribution theory functions are replaced by distributions, which are defined as linear maps from some auxiliary space of test functions to $\mathbb{C}$. So start with some topological vector space V of test functions (so we have a notion of convergence from the topology). Then we say $u \in V^*$ (i.e. in the topological dual) if $u : V \to \mathbb{C}$ is linear and respects convergence in V (i.e. is continuous), i.e. if the pairing between V^* and V is denoted by $\langle \cdot, \cdot \rangle$[(i)], then $u \in V^*$ implies $\langle u, f_n \rangle \to \langle u, f \rangle$ if $f_n \to f$ in V.

Example 0.4. *Let $V = C^\infty(\mathbb{R})$, with the topology of local uniform convergence. In this topology, we have $f_n \to f$ if and only if $f_n|_K \to f|_K$ uniformly for all $K \subset \mathbb{R}$ compact.*

Then define $\delta_{x_0} : V \to \mathbb{R}$ by: $\langle \delta_{x_0}, f \rangle := f(x_0)$. Then δ_{x_0} is a distribution (i.e. $\delta_{x_0} \in V^*$), and is exactly the Dirac delta from before.

[(i)] By the pairing we mean $\langle \cdot, \cdot \rangle : V^* \times V \to \mathbb{C}$ defined by $\langle u, f \rangle := u(f)$

1. Distributions

1.1. **Test Functions and Distributions.**

Let $X \subset \mathbb{R}^n$ be open. Then the first space of test functions we look at is:

Definition 1.1. *Define* $\mathscr{D}(\boldsymbol{X})$ *to be the topological vector space* $C_c^\infty(X)$ *with the topology defined by:*

$$\varphi_m \to 0 \ \text{ in } \mathscr{D}(X) \quad \text{if} \quad \exists \text{ compact } K \subset X \text{ with } \mathrm{supp}(\varphi_m) \subset K \ \forall m \text{ and}$$
$$\partial^\alpha \varphi_m \to 0 \text{ uniformly in } X \text{ for every multi-index } \alpha.$$

Note: The compact set K in the above definition must be the same for the <u>whole</u> sequence.

This is a nice set of functions. Compact support ensures that φ and its derivatives vanish at the boundary of X, which means that integration by parts is easy, i.e. if $\varphi, \psi \in \mathscr{D}(X)$ then

$$\int_X \varphi \cdot \partial^\alpha \psi \ \mathrm{d}x = (-1)^{|\alpha|} \int_X \psi \cdot \partial^\alpha \varphi \ \mathrm{d}x$$

i.e. we have no boundary terms.

We also have Taylor's theorem to arbitrary order, i.e.

$$\varphi(x+h) = \sum_{|\alpha| \leq N} \frac{h^\alpha}{\alpha!} \partial^\alpha \varphi(x) + R_N(x,h)$$

where $R_N(x,h) = o\left(|h|^{N+1}\right)$ uniformly in x [See Example Sheet 1].

As a general philosophy for distribution theory, 'nicer' test functions will have more assumptions on them, and will give us more distributions.

Definition 1.2. *A linear map* $u : D(X) \to \mathbb{C}$ *is called a* ***distribution****, written* $u \in \mathscr{D}'(X)$*, if for each compact* $K \subset X$*,* $\exists$ *constants* $C = C(K), N = N(K)$*, such that*

$$|\langle u, \varphi \rangle| \leq C \sum_{|\alpha| \leq N} \sup_X |\partial^\alpha \varphi| \tag{1.1}$$

for all $\varphi \in \mathscr{D}(X)$ *with* $\mathrm{supp}(\varphi) \subset K$.

If N *can be chosen independent of* K*, then the smallest such* N *is called the* ***order*** *of* u*, and is denoted* $\mathrm{ord}(u)$.

We refer to an inequality like (1.1) as a *semi-norm estimate* (semi-norm since it is only for a given compact set K). Note that the supremum can be over K instead of X, since all the φ have support in K.

Example 1.1. *Recall the Dirac delta, δ_{x_0}, for $x_0 \in X$. This was defined by*

$$\langle \delta_{x_0}, \varphi \rangle := \varphi(x_0)$$

and so thus

$$|\langle \delta_{x_0}, \varphi \rangle| = |\varphi(x_0)| \leq \sup |\varphi|$$

and thus we have (1.1) with $C \equiv 1$ and $N \equiv 0$, independent of supp(φ). So hence $\delta_{x_0} \in \mathscr{D}'(X)$ and ord$(\delta_{x_0}) = 0$.

Example 1.2. *A more interesting example is to define the linear map $T : \mathscr{D}(X) \to \mathbb{C}$ by:*

$$\langle T, \varphi \rangle := \sum_{|\alpha| \leq M} \int_X f_\alpha(x) \cdot \partial^\alpha \varphi(x) \, \mathrm{d}x$$

where $f_\alpha \in C(X)$. Note that this is well-defined since the φ have compact support, and so the integrals exist since then $f_\alpha \cdot \partial^\alpha \varphi$ is continuous with compact support.

We can see that this is a distribution. Indeed, fix a compact set $K \subset X$. Then if $\varphi \in D(X)$ with supp$(\varphi) \subset K$, we have:

$$\begin{aligned}
|\langle T, \varphi \rangle| &\leq \sum_{|\alpha| \leq M} \int_X |f_\alpha(x)| \cdot |\partial^\alpha \varphi(x)| \, \mathrm{d}x \\
&= \sum_{|\alpha| \leq M} \int_K |f_\alpha(x)| \cdot |\partial^\alpha \varphi(x)| \, \mathrm{d}x \\
&\leq \sum_{|\alpha| \leq M} \sup_K |\partial^\alpha \varphi| \cdot \int_K |f_\alpha(x)| \, \mathrm{d}x \\
&\leq \left(\max_{|\alpha| \leq M} \int_K |f_\alpha(x)| \, \mathrm{d}x \right) \cdot \sum_{|\alpha| \leq M} \sup_K |\partial^\alpha \varphi|
\end{aligned}$$

which is exactly (1.1) with $C = \max_{|\alpha| \leq M} \int_K |f_\alpha(x)| \, \mathrm{d}x$, and $N = M$.

So hence this shows that $T \in \mathscr{D}'(X)$ and ord$(T) \leq M$ (as some of the f_α could be 0).

Note that we still have $T \in \mathscr{D}'(X)$ is the f_α are just locally integrable (i.e. $f_\alpha \in L^1_{loc}(X)$, i.e. $f_\alpha \in L^1(K)$ for all compact $K \subset X$).

In general, if $g \in L^1_{\text{loc}}(X)$ we can define $T_g \in \mathscr{D}'(X)$ via:

$$\langle T_g, \varphi \rangle := \int_X g \cdot \varphi \, \mathrm{d}x.$$

Example 1.2 above is essentially a linear combination of such distributions. We often abuse notation and write $T_g \equiv g$. For example, on $\mathscr{D}'(\mathbb{R})$, we might say "consider the distribution $u(x) = x$". Clearly this is a function, not a distribution. But what we mean is the distribution

$$\langle u, \varphi \rangle := \int_{\mathbb{R}} x \varphi(x) \, \mathrm{d}x.$$

Example 1.3 (Distribution with 'infinite' order)**.** *Consider on $\mathscr{D}(\mathbb{R})$, the linear map*

$$\langle u, \varphi \rangle := \sum_{m=0}^{\infty} \varphi^{(m)}(m)$$

where $\varphi^{(m)}$ denotes m-th derivative. Note that for each $\varphi \in \mathscr{D}(\mathbb{R})$, this is a finite sum since φ has compact support. So in particular we get

$$|\langle u, \varphi \rangle| \leq \sum_{m=0}^{\infty} |\varphi^{(m)}(m)| \leq \sum_{m=0}^{\infty} \sup |\varphi^{(m)}|$$

and so if supp$(\varphi) \subset [-K, K]$, we see that $|\langle u, \varphi \rangle| \leq \sum_{m=0}^{K} \sup |\varphi^{(m)}|$. So hence if we take $K \to \infty$ (which we can do, since we can always find some function which has support contained in $[-K, K]$ for each K), this shows that there is no universal N value as in (1.1) which works for all $K \subset \mathbb{R}$ compact. Hence the order of this u is not defined (it is 'infinite' if you like).

So how does this definition of distribution relate to continuity? The key result from analysis is that for a linear map (between normed vector spaces), continuity is equivalent to boundedness. This gives the so-called **sequential continuity definition** of $\mathscr{D}'(X)$:

Lemma 1.1 (Sequential Continuity Definition of $\mathscr{D}'(X)$)**.**

Let $u : \mathscr{D}(X) \to \mathbb{C}$ be a linear map. Then:

(1.2) $$u \in \mathscr{D}'(X) \iff \lim_{n\to\infty} \langle u, \varphi_m \rangle = 0 \text{ for all sequences } (\varphi_m)_m \text{ with } \varphi_m \to 0 \text{ in } \mathscr{D}(X).$$

Proof. $\underline{(\Rightarrow)}$: Suppose $u \in \mathscr{D}'(X)$, and $\varphi_m \to 0$ in $\mathscr{D}(X)$. In particular, we know $\exists K \subset X$ compact with supp$(\varphi_m) \subset K$ for all m. Then from (1.1), we have

$$|\langle u, \varphi_m \rangle| \leq C \sum_{|\alpha| \leq N} \sup |\partial^\alpha \varphi_m| \to 0$$

since $\sup |\partial^\alpha \varphi_m| \to 0$ (from $\varphi_m \to 0$ in $\mathscr{D}(X)$).

$\underline{(\Leftarrow)}$: Suppose the implication is not true. Then we could have the RHS of (1.2) holding, but no estimate of the form of (1.1) holds. Hence $\exists K \subset X$ compact such that (1.1) fails for every choice of C, N. So take $C = N = m \in \mathbb{N}_{>0}$. Then since (1.1) fails with these choices of C, N we know $\exists \varphi_m \in \mathscr{D}(X)$ with supp$_m(\varphi) \subset K$ and

$$|\langle u, \varphi_m \rangle| > m \sum_{|\alpha| \leq m} \sup |\partial^\alpha \varphi_m|.$$

In particular, since the RHS here is ≥ 0 this tells us that $|\langle u, \varphi_m \rangle| > 0$. So define

$$\tilde{\varphi}_m := \varphi_m / \langle u, \varphi_m \rangle.$$

Then we have

$$1 = |\langle u, \tilde{\varphi}_m \rangle| > m \sum_{|\alpha| \leq m} \sup |\partial^\alpha \tilde{\varphi}_m|$$

i.e.
$$\sum_{|\alpha|\leq m}\sup|\partial^\alpha\tilde{\varphi}_m|<\frac{1}{m}\quad\Longrightarrow\quad\sup|\partial^\alpha\tilde{\varphi}_m|<\frac{1}{m}\quad\forall|\alpha|\leq m.$$
But then this shows that $\tilde{\varphi}_m\to 0$ in $\mathscr{D}(X)$. But then we can apply (1.2) to this sequence to see that we require
$$\langle u,\varphi_m\rangle\to 0\ \text{ as }\ m\to\infty.$$
But we know $\langle u,\varphi_m\rangle=1$ for all m, and so we have a contradiction. So the implication must be true.

□

1.2. **Limits in $\mathscr{D}'(X)$.**

Often we have sequences of distributions, $(u_m)_m\subset\mathscr{D}'(X)$.

Definition 1.3. *We say $u_m\to 0$ in $\mathscr{D}'(X)$ if $\lim_{m\to\infty}\langle u_m,\varphi\rangle=0$ for all $\varphi\in\mathscr{D}(X)$, i.e. w^*-convergence.*

Remark: So far we have only defined convergence to zero in the spaces $\mathscr{D}(X)$ and $\mathscr{D}'(X)$. However since these are topological vector spaces we get natural definitions of convergence to other limits:

- $\varphi_m\to\varphi$ in $\mathscr{D}(X)$ means $\varphi_m-\varphi\to 0$ in $\mathscr{D}(X)$
- $u_m\to u$ in $\mathscr{D}'(X)$ means $u_m-u\to 0$ in $\mathscr{D}'(X)$.

An important result is that $\mathscr{D}'(X)$ is that pointwise limits of distributions are also distributions.

Theorem 1.1 (Non-Examinable). *Suppose $(u_m)_m\subset\mathscr{D}'(X)$, and suppose that $\langle u,\varphi\rangle:=\lim_{m\to\infty}\langle u_m,\varphi\rangle$ exists for all $\varphi\in\mathscr{D}(X)$. Then $u\in\mathscr{D}'(X)$.*

Proof. This is just an application of the uniform-boundedness principle/Banach-Steinhaus. See Hörmander, Vol 1 for more details.

□

Limits in $\mathscr{D}'(X)$ can look strange however.

Example 1.4. *Consider $\mathscr{D}'(\mathbb{R})$, and $u_m(x):=\sin(mx)$. Then:*
$$\begin{aligned}\langle u_m,\varphi\rangle&=\int_{\mathbb{R}}\sin(mx)\varphi(x)\,\mathrm{d}x\\&=\frac{1}{m}\int_{\mathbb{R}}\cos(mx)\varphi'(x)\,\mathrm{d}x\end{aligned}$$

via integrating by parts, which thus shows

$$|\langle u_m, \varphi\rangle| \leq \frac{1}{m}\int_{\mathbb{R}} |\varphi'(x)|\, \mathrm{d}x \to 0 \ \ \textit{as} \ \ m \to \infty$$

for any $\varphi \in \mathscr{D}(\mathbb{R})$, which shows that $u_m \to 0$ in $\mathscr{D}'(\mathbb{R})$. So hence

$$\text{“}\lim_{m\to\infty} \sin(mx) = 0 \ \ \textit{in} \ \ \mathscr{D}'(\mathbb{R})\text{”}$$

if you will (just remember we are viewing these as distributions, <u>not</u> functions).

1.3. **Basic Operations.**

1.3.1. *Differentiation and Multiplication by Smooth Functions.*

To motivate the definition of derivatives of distributions, we look at the **duality** with the smooth function case. If $u \in C^\infty(X)$ then we can define $\partial^\alpha u \in \mathscr{D}'(X)$ via (for $\varphi \in \mathscr{D}(X)$)

$$\begin{aligned}\langle \partial^\alpha u, \varphi\rangle &= \int_X \partial^\alpha u \cdot \varphi \, \mathrm{d}x \\ &= (-1)^{|\alpha|}\int_X u \cdot \partial^\alpha \varphi \, \mathrm{d}x \\ &= (-1)^{|\alpha|}\langle u, D^\alpha \varphi\rangle\end{aligned}$$

where we have integrated by parts, which we can do since everything is smooth. Thus we define more generally:

Definition 1.4. *For $u \in \mathscr{D}'(X)$ and $f \in C^\infty(X)$, we define the* ***derivative*** *of the distribution fu by:*

$$\langle \partial^\alpha (fu), \varphi\rangle := (-1)^{|\alpha|}\langle u, f\partial^\alpha \varphi\rangle$$

for $f \in \mathscr{D}(X)$.

We call $\partial^\alpha u$ the ***distributional/weak derivative*** *of u.*

Note: In the case of $u \in C^\infty(X)$, then these distributional derivatives are exactly the derivatives of u. However more generally we can define them for distributions, and we see that for any distribution <u>all</u> the distributional derivatives are defined.

Remark: Taking $\alpha = 0$ in the definition, this also tells us that we define multiplication by a smooth function of a distribution by

$$\langle fu, \varphi\rangle := \langle u, f\varphi\rangle.$$

Example 1.5. *For δ_x the Dirac Delta we have:*

$$\langle \partial^\alpha \delta_x, \varphi\rangle = (-1)^{|\alpha|}\langle \delta_x, \partial^\alpha \varphi\rangle = (-1)^{|\alpha|}\partial^\alpha \varphi(x).$$

Compare this with our heuristic definition of $\partial^\alpha \delta_x$ from §0.

Example 1.6. *Define the Heaviside function by*

$$H(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0. \end{cases}$$

Then since $H \in L^1_{\text{loc}}(\mathbb{R})$, it defines a distribution $H \in \mathscr{D}'(\mathbb{R})$. Then we have:

$$\begin{aligned} \langle H', \varphi \rangle := -\langle H, \varphi' \rangle &= -\int_{\mathbb{R}} H(x)\varphi'(x)\, \mathrm{d}x \\ &= -\int_0^\infty \varphi'(x) \\ &= -[\varphi(x)]_{x=0}^{x=\infty} \\ &= \varphi(0) \\ &= \langle \delta_0, \varphi \rangle \end{aligned}$$

where we have used the fact the φ has compact support and so it is zero near ∞. Hence this shows that $H' = \delta_0$ in $\mathscr{D}'(\mathbb{R})$.

Remark: In general, clearly if $u, v \in \mathscr{D}'(X)$ and we have $\langle u, \varphi \rangle = \langle v, \varphi \rangle$ for all $\varphi \in \mathscr{D}(X)$, then we say that $u = v$ in $\mathscr{D}'(X)$.

With this knowledge, we are now able to solve the simplest of distributional equations: $u' = 0$.

Lemma 1.2. *Suppose $u' = 0$ in $\mathscr{D}'(\mathbb{R})$. Then $u =$ constant in $\mathscr{D}'(\mathbb{R})$.*

Key point. Fix $\varphi \in \mathscr{D}(\mathbb{R})$. We want to consider $\langle u, \varphi \rangle$. Now if we could find ψ such that $\psi' = \varphi$, then we would get

$$\langle u, \varphi \rangle = \langle u, \psi' \rangle = -\langle u', \psi \rangle = \langle 0, \psi \rangle.$$

The only issue with such an argument is that ψ would be defined via an integral of φ, and thus does not necessarily have compact support. But we need $\psi \in \mathscr{D}'(\mathbb{R})$ for this to work. We cannot just subtract a constant to make it have compact support, since it would mess up either end. Thus we need to subtract a constant off in a 'weighted' way, as the proof demonstrates.

Proof. Fix $\theta_0 \in \mathscr{D}(\mathbb{R})$ with

$$1 = \int_{\mathbb{R}} \theta_0 \, \mathrm{d}x = \langle 1, \theta_0 \rangle.$$

Take $\varphi \in \mathscr{D}(\mathbb{R})$ arbitrary. Then write:

$$\varphi = \underbrace{(\varphi - \langle 1, \varphi \rangle \theta_0)}_{=:\varphi_A} + \underbrace{\langle 1, \varphi \rangle \theta_0}_{=:\varphi_B}.$$

Clearly we have $\langle 1, \varphi_A \rangle = \int_{\mathbb{R}} \varphi_A \, \mathrm{d}x = 0$. Now set:

$$\psi_A(x) := \int_{-\infty}^{x} \varphi_A(y) \, \mathrm{d}y,$$

which is smooth since it is the integral of a smooth function.

Claim: $\psi_A \in \mathscr{D}(\mathbb{R})$.

Proof of Claim. Since θ_0, φ have compact support, so does φ_A. Hence for x sufficiently negative (outside the support of φ_A), this integral is just the integral of 0, and so we see $\psi_A(x) = 0$ for all $x \leq -C$, for some $C > 0$.

Also since $\varphi_A = 0$ for all $x \geq C'$, we see that ψ_A is constant for $x \geq C'$. But we know that $\psi_A(x) \to 0$ as $x \to \infty$ since $\int_{\mathbb{R}} \varphi_A \, \mathrm{d}x = 0$, and so this constant must be 0, i.e. $\psi_A(x) = 0$ for all $x \geq \tilde{C}$.

Hence this shows ψ_A has compact support, and so $\psi_A \in \mathscr{D}(\mathbb{R})$.

□

So $\psi_A \in \mathscr{D}(\mathbb{R})$ and $\psi_A' = \varphi_A$. So we have:

$$\begin{aligned}
\langle u, \varphi \rangle &= \langle u, \varphi_A \rangle + \langle u, \varphi_B \rangle \\
&= \langle u, \psi_A' \rangle + \langle 1, \varphi \rangle \cdot \underbrace{\langle u, \theta_0 \rangle}_{=c,\text{ a constant independent of } \varphi} \\
&= -\langle \underbrace{u'}_{=0 \text{ by assumption}}, \psi_A \rangle + \langle c, \varphi \rangle \\
&= \langle c, \varphi \rangle
\end{aligned}$$

i.e. $\langle u, \varphi \rangle = \langle c, \varphi \rangle$ for all $\varphi \in \mathscr{D}(\mathbb{R})$, and thus $u = c$ in $\mathscr{D}'(\mathbb{R})$.

□

Remark: We can use similar arugments to solve other distributional equations. We can also use what we know about the distributional derivatives of other distributions to help us.

1.3.2. *Reflection and Translation.*

For $\varphi \in \mathscr{D}(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$, we can define the **translate** of φ by:

$$(\tau_h \varphi)(x) := \varphi(x - h)$$

and the **reflection** of φ by:

$$\check{\varphi}(x) := \varphi(-x).$$

τ_h is called the **translation operator**, and $\vee$ the **reflection operator**. We can extend these definitions to $\mathscr{D}'(\mathbb{R}^n)$ by duality[(ii)], as we have been doing.

[(ii)] By duality, we just mean see the result first for distributions defined by smooth functions, and then define the general case obey the same relation, just like we did for, e.g. derivatives.

Definition 1.5. *For $u \in \mathscr{D}'(\mathbb{R}^n)$ and $h \in \mathbb{R}^n$ we define the* ***distributional reflection*** *and* ***reflections*** *by:*
$$\langle \tau_h u, \varphi \rangle := \langle u, \tau_{-h}\varphi\rangle \quad \text{and} \quad \langle \check{u}, \varphi\rangle := \langle u, \check{\varphi}\rangle.$$

Remark: If u we smooth, then one can check that these definitions agree with what we would expect, e.g.
$$\langle \check{u}, \varphi\rangle = \int_{\mathbb{R}^n} u(-x)\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} u(x)\varphi(-x)\,\mathrm{d}x = \langle u, \check{\varphi}\rangle$$
etc.

So do the definitions of translation and distributional derivative coincide in the usual way? It turns out that they do, and the following lemma shows this for directional derivatives.

Lemma 1.3 (Directional Distributional Derivatives)**.** *Let $u \in \mathscr{D}'(\mathbb{R}^n)$ and $h \in \mathbb{R}^n\backslash\{0\}$. Then define:*
$$v_h := \frac{\tau_{-h}(u) - u}{|h|}.$$
Then if $h/|h| \to m \in S^{n-1}$, then we have $v_h \to m \cdot \partial u$ in $\mathscr{D}'(\mathbb{R}^n)$ as $h \to 0$.

Remark: If u were smooth, then we would have
$$v_h(y) = \frac{u(y+h) - u(y)}{|h|}.$$

Proof. By definition we have
$$\langle v_h, \varphi\rangle = \frac{1}{|h|}\langle u, \tau_h\varphi - \varphi\rangle.$$
Then by Taylor's theorem, we know
$$(\tau_h\varphi)(x) - \varphi(x) = \varphi(x-h) - \varphi(x) = -\sum_i h_i \frac{\partial\varphi}{\partial x_i}(x) + R_1(x,h)$$
where $R_1(x,h) = o(|h|)$ in $\mathscr{D}(\mathbb{R}^n)$. [See Example Sheet 1, Question 2.] Then by the sequentially continuity definition of $\mathscr{D}'(\mathbb{R}^n)$ (this is to pull the $R_1(x,h)$ term out) and the convergence $h/|h| \to m$ we have
$$\begin{aligned}
\langle v_h, \varphi\rangle &= -\left\langle u, \sum_i \frac{h_i}{|h|}\cdot\frac{\partial\varphi}{\partial x_i}\right\rangle + o(1)\\
&= -\sum_i \frac{h_i}{|h|}\left\langle u, \frac{\partial\varphi}{\partial x_i}\right\rangle + o(1)\\
&\longrightarrow \sum_i m_i \left\langle u, \frac{\partial \varphi}{\partial x_i}\right\rangle = \left\langle \sum_i m_i \frac{\partial u}{\partial x_i}, \varphi\right\rangle = \langle m\cdot\partial u, \varphi\rangle
\end{aligned}$$
i.e. $v_h \to m\cdot\partial u$ in $\mathscr{D}'(\mathbb{R}^n)$ as $h \to 0$.

□

1.3.3. *Convolution between $\mathscr{D}'(\mathbb{R}^n)$ and $\mathscr{D}(\mathbb{R}^n)$.*

By combining reflection and translation, we see that for $\varphi \in \mathscr{D}(\mathbb{R}^n)$ we have

$$(\tau_x \check{\varphi})(y) = \check{\varphi}(y-x) = \varphi(x-y).$$

So if $u \in C^\infty(\mathbb{R}^n)$, we have the classical notion of **convolution** via:

$$\begin{aligned}(u*\varphi)(x) &= \int_{\mathbb{R}^n} u(x-y)\varphi(y)\,\mathrm{d}y\\ &= \int_{\mathbb{R}^n} u(y)\varphi(x-y)\,\mathrm{d}y\\ &= \int_{\mathbb{R}^n} u(y)(\tau_x\check{\varphi})(y)\,\mathrm{d}y\\ &= \langle u, \tau_x\check{\varphi}\rangle\end{aligned}$$

and this last equality makes sense for any distribution on $\mathbb{R}^n$. Hence we define (by duality, if you will):

Definition 1.6. *If $u \in \mathscr{D}'(\mathbb{R}^n)$, then the **(distributional) convolution** of u with $\varphi \in \mathscr{D}(\mathbb{R}^n)$ is defined by:*

$$(u*\varphi)(x) := \langle u, \tau_x\check{\varphi}\rangle$$

and thus is a function $(u*\varphi): \mathbb{R}^n \to \mathbb{R}$.

So the convolution of a distribution with a test function gives another function. What can be said about this function? In the case where the distribution is defined by a smooth function we know that convolutions are usually smooth as well. Can the same be said here? With an extra condition, it turns out that it can be, and the usual derivative formula holds.

Lemma 1.4 (Derivatives of (Distributional) Convolution)**.** *Let $\varphi \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ and define $\Phi_x(y) := \varphi(x,y)$. Suppose that for each $x \in \mathbb{R}^n$, $\exists$ a neighbourhood N_x of x on which for all $x' \in N_x$ we have supp$(\Phi_{x'}) \subset K$, for some K compact (independent of $x' \in N_x$). Then,*

$$\partial_x^\alpha \langle u, \Phi_x\rangle = \langle u, \partial_x^\alpha \Phi_x\rangle.$$

Proof. The assumption on Φ_x means that for each $x \in \mathbb{R}^n$, we can find an open neighbourhood N_x of x and K compact such that $\mathrm{supp}\big(\varphi(\cdot,\cdot)|_{N_x\times\mathbb{R}^n}\big) \subset N_x \times K$. [Shown in Figure 1.]

By Taylor's theorem, for fixed $x \in \mathbb{R}^n$ we have:

$$\Phi_{x+h}(y) - \Phi_x(y) = \sum_i h_i \frac{\partial\varphi}{\partial x_i}(x,y) + R_1(x,y,h)$$

where $R_1 = o(|h|)$ and $\mathrm{supp}(R_1(x,\cdot,h)) \subset K$ for some compact K for $|h|$ sufficiently small [this is by our assumption on φ, as for $|h|$ sufficiently small we will have $x+h \in N_x$ and so $\mathrm{supp}(\Phi_{x+h}) \subset K$.]

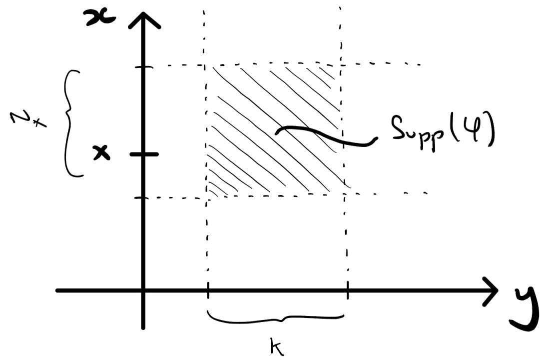

FIGURE 1. An illustration of the assumption on φ in Lemma 1.4.

So hence we see that for each such x, we have $R_1(x,\cdot,h) = o(|h|)$ as an equality in $\mathscr{D}(\mathbb{R}^n)$. So using this and sequential continuity (Lemma 1.1) we have (using the above):

$$\begin{aligned}\langle u, \Phi_{x+h}\rangle - \langle u, \Phi_x\rangle &= \left\langle u, \sum_i h_i \frac{\partial \Phi_x}{\partial x_i}\right\rangle + o(|h|)\\ &= \sum_i h_i \left\langle u, \frac{\partial \Phi_x}{\partial x_i}\right\rangle + o(|h|).\end{aligned}$$

So hence this must be the Taylor expansion of $\langle u, \Phi_x\rangle$ w.r.t x, and thus we see (combinging $o(h)$ terms)

$$\frac{\partial}{\partial x_i}\langle u, \Phi_x\rangle = \left\langle u, \frac{\partial \Phi_x}{\partial x_i}\right\rangle.$$

Thus this proves the result for first order derivatives. The general result then follows from repeating the above by induction.

□

Corollary 1.1 (Important). *Suppose $u \in \mathscr{D}'(\mathbb{R}^n)$ and $\varphi \in \mathscr{D}(\mathbb{R}^n)$. Then:*

$$u * \varphi \in C^\infty(\mathbb{R}^n) \quad \textit{and} \quad \partial^\alpha(u*\varphi) = u * \partial^\alpha \varphi.$$

Proof. Fix such a u, φ. Then define $\tilde{\varphi} \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ by

$$\tilde{\varphi}(x,y) := \tau_x \check{\varphi}(y) = \varphi(x-y).$$

Then for $x \in \mathbb{R}^n$, we can clearly find a neighbourhood N_x of x and $K \subset \mathbb{R}^n$ compact as in Lemma 1.4, since φ has compact support (e.g. just take a ball about x and the relevant translate of the support of φ).

Thus we can apply Lemma 1.4 to this $\tilde{\varphi}$, which proves the result since $\langle u, \Phi_x\rangle = (u*\varphi)(x)$ here.

□

Remark: Heuristically, we can "write" $\delta(x) = \lim_{n\to\infty} \delta_n(x)$. Then since we can write

$$u(x) = (u * \delta)(x) = \lim_{n\to\infty} (u * \delta_n)(x)$$

which is a limit of smooth functions. So heuristically, this says that distributions can be "viewed" as a limit of smooth functions, which makes our lives very simple if we can make sense if this/if it is true, because we know how to understand limits of smooth functions!

1.4. Density of $\mathscr{D}(\mathbb{R}^n)$ in $\mathscr{D}'(\mathbb{R}^n)$.

We have now seen that for $u \in \mathscr{D}'(\mathbb{R}^n)$ and $\varphi \in \mathscr{D}(\mathbb{R}^n)$ that $u * \varphi \in C^\infty(\mathbb{R}^n)$, no matter how "weird" u is. We often call $u * \varphi$ a **regularisation** of u.

Lemma 1.5. *Suppose $u \in \mathscr{D}'(\mathbb{R}^n)$ and $\varphi, \psi \in \mathscr{D}(\mathbb{R}^n)$. Then:*

$$(u * \varphi) * \psi = u * (\varphi * \psi)$$

i.e. $*$ is associative (when we have two test functions, 1 distribution)[(iii)].

Note: On the LHS, the second ('outer') convolution is actually a classical convolution of smooth functions, since $u * \varphi$ is smooth, whilst on the RHS the outer convolution is the one defined for distributions.

Proof. For fixed $x \in \mathbb{R}^n$ we have

$$\begin{aligned} [(u * \varphi) * \psi](x) &:= \int (u * \varphi)(x - y)\psi(y)\, \mathrm{d}y \\ &= \int \langle u, \tau_{x-y}\check{\varphi}\rangle \psi(y)\, \mathrm{d}y \\ &= \int \langle u(z), (\tau_{x-y}\check{\varphi})(z)\rangle \psi(y)\, \mathrm{d}y \\ &= \int \langle u(z), \varphi(x - y - z)\psi(y)\rangle\, \mathrm{d}y \end{aligned}$$

where we have brought in the factor of $\psi(y)$ into the distribution by linearity, since for each y, $\psi(y)$ is just a constant.

Now to see that this is what we are after, write the final integral as a Riemann sum, and so we get

$$\begin{aligned} &= \lim_{h\to 0} \sum_{m\in\mathbb{Z}^n} \langle u(z), \varphi(x - z - hm)\psi(hm)h^n\rangle \\ &= \lim_{h\to 0} \left\langle u(z), \sum_{m\in\mathbb{Z}^n} \varphi(x - z - hm)\psi(hm)h^n \right\rangle \end{aligned} \qquad (\star)$$

where we are able to bring in the 'infinite sum' into the brackets by linearity of distributions, since for each $h > 0$ the sum is actually a finite sum since ψ has compact support (so for $|m|$ large enough $\psi(hm) = 0$).

(iii) Although this is the only way we can make sense of this, since we can't convolute a distribution with another distribution.

Now set

$$F_h(z) := \sum_{m\in\mathbb{Z}^n} \varphi(x-z-hm)\psi(hm)h^m.$$

We can then show that $\text{supp}(F_h)$ is contained inside a fixed compact $K \subset \mathbb{R}^n$ for $|h| \leq 1$, and that we have

$$F_h(z) \to (\varphi * \psi)(x-z) \qquad \text{pointwise}$$

and we also have $\partial^\alpha F_h \to \partial^\alpha(\varphi * \psi)$ pointwise. But for convergence in $\mathscr{D}(\mathbb{R}^n)$, we want uniform convergence of these.

But note that for each multi-index α we have $|\partial^\alpha F_h| \leq M_\alpha$, and thus we have equicontinuity of all derivatives. Hence by Arzelá-Ascoli, by passing to a subsequence if necessary, we get that $F_h(z) \to \varphi * \psi(x-z)$ uniformly. Hence

$$\sum_{m\in\mathbb{Z}^n} \varphi(x-z-hm)\psi(hm)h^n \to (\varphi*\psi)(x-z) \qquad \text{in } \mathscr{D}(\mathbb{R}^n) \quad \text{as } h \to 0.$$

[We may need to take a diagonal subsequence when invoking Arzelá-Ascoli to get all derivatives converge uniformly as well.]

Thus by sequential continuity we can exchange the limit in $(\star)$ with $\langle\cdot,\cdot\rangle$, and thus we get

$$\begin{aligned}(\star) &= \left\langle u(z), \lim_{h\to 0}\sum_{m\in\mathbb{Z}^n} \varphi(x-z-hm)\psi(hm)h^n \right\rangle \\ &= \langle u(z), (\varphi*\psi)(x-z)\rangle \\ &= \left\langle u, \tau_x(\varphi \check{*} \psi)\right\rangle^{\text{(iv)}} \\ &= = [u*(\varphi*\psi)](x)\end{aligned}$$

which completes the proof. □

Theorem 1.2 (Density Theorem). *Let $u \in \mathscr{D}'(\mathbb{R}^n)$. Then $\exists$ a sequence $(\varphi_m)_m \subset \mathscr{D}(\mathbb{R}^n)$ such that $\varphi_m \to u$ in $\mathscr{D}'(\mathbb{R}^n)$, i.e.*

$$\langle \varphi_m, \theta\rangle \to \langle u, \theta\rangle \quad \forall \theta \in \mathscr{D}(\mathbb{R}^n).$$

Proof. First fix $\psi \in \mathscr{D}(\mathbb{R}^n)$ with $\int \psi \, dx = 1$. Then define:

$$\psi_m(x) := m^n\psi(mx)^{\text{(v)}}$$

Also fix $\chi \in \mathscr{D}(\mathbb{R}^n)$ with

$$\chi(x) = \begin{cases} 1 & \text{on } |x| < 1 \\ 0 & \text{on } |x| > 2. \end{cases}$$

Then define

$$\chi_m(x) := \chi(x/m)$$

This χ will function as a cut off function to ensure compact support of our φ_m.

Finally define

$$\varphi_m := \chi_m(u*\psi).$$

(iv) Apologies, I can't seem to get a wider 'check' symbol...

(v) This is a bit like an approximation to a δ-function - the function gets 'squished'and more 'spiked' as $m \to \infty$.

From the above we know that $\varphi_m \in \mathcal{D}(\mathbb{R}^n)$. Then for $\theta \in \mathcal{D}(\mathbb{R}^n)$ we have

$$\begin{aligned}\langle \varphi_m, \theta\rangle &= \langle u * \psi_m, \chi_m\theta\rangle \\ &= \left[(u*\psi_m)*(\chi_m\check{\theta})\right](0) \qquad (\star)\\ &= u*\left[\psi_m*(\chi_m\check{\theta})\right](0)\end{aligned}$$

where in the second line we have used the fact that in general, $\langle u, \varphi\rangle = (u*\check{\varphi})(0)$, and then we have used Lemma 1.5 (associativity of convolution).

Now,

$$\begin{aligned}\left[\psi*(\chi\check{\theta})\right](x) &= \int \psi_m(x-y)\chi_m(-y)\theta(-y)\ \mathrm{d}y\\ &= \int m^n\psi(m(x-y))\chi(-y/m)\theta(-y)\ \mathrm{d}y\\ &= \int \psi(y)\chi\left(\frac{y}{m^2}-\frac{x}{m}\right)\theta\left(\frac{y}{m}-x\right)\ \mathrm{d}y \qquad [\text{change variables } y' = m(x-y) \text{ to get this}]\\ &= \theta(-x) + \int \psi(y)\left[\chi\left(\frac{y}{m^2}-\frac{x}{m}\right)\theta\left(\frac{y}{m}-x\right)-\theta(-x)\right]\ \mathrm{d}y\\ &=: \theta(-x)+R_m(-x)\\ &= \check{\theta}(x)+\check{R}_m(x)\end{aligned}$$

where we have used that $\int \psi(y)\ \mathrm{d}y = 1$ and where

$$R(x) := \int \psi(y)\left[\chi\left(\frac{y}{m^2}+\frac{x}{m}\right)\theta\left(\frac{y}{m}+x\right)-\theta(x)\right]\ \mathrm{d}y.$$

Now (†) gives

$$\langle \varphi_m, \theta\rangle = (u*\check{\theta})(0)+(u*\check{R}_m)(0) = \langle u,\theta\rangle + \langle u, R_m\rangle.$$

Then it is straightforward to show that $R_m \to 0$ in $\mathcal{D}(\mathbb{R}^n)$ [**Exercise** to check], and thus by the sequential continuity of $\mathcal{D}'(\mathbb{R}^n)$ we get

$$\langle \varphi_m, \theta\rangle \to \langle u, \theta\rangle$$

and so since $\theta \in \mathcal{D}(\mathbb{R}^n)$ was arbitrary, we deduce that $\varphi_m \to u$ in $\mathcal{D}'(\mathbb{R}^n)$.

□

Remark: So we see that test functions (or more precisely, their associated distributions) are dense in $\mathcal{D}'(\mathbb{R}^n)$. It is a viewpoint of some mathematicians (e.g. Terry Tao) that this is the best way to do everything with distributions: first prove the result for the test function case and then use density/continuity to get the result for all distributions. [Terry Tao wrote about this on his blog.]

2. Distributions with Compact Support

Definition 2.1. *For $u \in \mathscr{D}'(X)$ we say that u* ***vanishes*** *on $Y \subset X$ if $\langle u, \varphi \rangle = 0$ for all $\varphi \in \mathscr{D}(X)$ with supp$(\varphi) \subset Y$.*

We then define the ***support*** *of u, supp(u), as the complement of the largest open set on which u vanishes, i.e.*

$$\mathrm{supp}(u) = X \setminus \bigcup \{Y \subset X : u \text{ vanishes on } Y\}.$$

In practice, usually it is easy to see what the support of a distribution is.

Example 2.1. *Consider the Dirac Delta, δ_x. Then since $\langle \delta_x, \varphi \rangle := \varphi(x)$ consider $Y = X \setminus \{x\}$. Then Y is open, and any $\varphi \in \mathscr{D}(X)$ with supp$(\varphi) \subset Y$ will have $\langle \delta_x, \varphi \rangle = \varphi(y) = 0$. Thus u vanishes on Y, and clearly u does not vanish on X. Thus we have supp$(\delta_x) = \{x\}$.*

2.1. **Test Functions and Distributions.**

We have seen that if we have more restrictions on the test functions we are considering (so there are less of them), then we have more distributions and so there is a seesaw effect. For example, we have so far only considered test functions which were smooth and had compact support, and so the associated distributions did not care about the boundary of our domain. However if we drop the compact support assumption, then we expect the distributions to care about what happens at the boundary, and so they change.

So let us define a new space of test functions.

Definition 2.2. *Define a space of test functions by $\mathscr{E}(X)$, which is the space of smooth functions $\varphi : X \to \mathbb{C}$ with the topology:*

$$\varphi_m \to 0 \text{ in } \mathscr{E}(X) \quad \text{if} \quad \partial^\alpha \varphi_m \to 0 \text{ uniformly on each compact } K \subset X, \text{ for each } \alpha.$$

i.e. $\mathscr{E}(X) := C^\infty(X)$ with the topology of local uniform convergence.

Remark: This makes $\mathscr{E}(X)$ into a Frechét space.

Note that clearly $\mathscr{D}(X) \subset \mathscr{E}(X)$. So intuitively since the distributions are the dual spaces of these, we expect $\mathscr{E}'(X) \subset \mathscr{D}'(X)$.

Remark: How are the topologies on $\mathscr{D}(X)$, $\mathscr{E}(X)$ related? Are they compatible, i.e. is the topology on $\mathscr{D}(X)$ just the induced subspace topology?

Definition 2.3. *We say that a linear map $u : \mathscr{E}(X) \to \mathbb{C}$ defines an element of $\boldsymbol{\mathscr{E}'(X)}$, the space of **distributions on $\boldsymbol{\mathscr{E}(X)}$**, if $\exists$ a compact set $K \subset X$ and constants C, N such that*

$$|\langle u, \varphi \rangle| \leq C \sum_{|\alpha| \leq N} \sup_K |\partial^\alpha \varphi| \qquad \forall \varphi \in \mathscr{E}(X).$$

Note: These distributions necessarily have finite order due to the uniformality of N.

As before we then have:

Lemma 2.1 (Sequential Continuity Definition of $\mathscr{E}'(X)$)**.** *Let $u : \mathscr{E}(X) \to \mathbb{C}$ be a linear map. Then:*

$$u \in \mathscr{E}'(X) \quad \Leftrightarrow \quad \lim_{n\to\infty} \langle u, \varphi_m \rangle = 0 \text{ for each } (\varphi_m)_m \subset \mathscr{E}(X) \text{ with } \varphi_m \to 0 \text{ in } \mathscr{E}(X).$$

Proof. Almost identical as the $\mathscr{D}'(X)$ version - **Exercise** to prove [need to use a compact exhaustion of X].

□

Note: If $u \in \mathscr{E}'(X)$ then it is true that $\text{supp}(u) \subset K$, where K is as in the definition of $\mathscr{E}'(X)$ (this is immediate from the definitions). But we do not always have $\text{supp}(u) = K$!

So how do $\mathscr{E}'(X)$ and $\mathscr{D}'(X)$ relate to one another? It turns out that we have $\mathscr{E}'(X) \subset \mathscr{D}'(X)$ (interpreted suitably) just as we expected. Also, elements of $\mathscr{D}'(X)$ with compact support actually define elements of $\mathscr{E}'(X)$ and thus the two spaces the compatible in the nicest way. This essential tells us that elements of $\mathscr{E}'(X)$ are just elements of $\mathscr{D}'(X)$ with compact support.

Lemma 2.2. *Let $u \in \mathscr{E}'(X)$. Then $u|_{\mathscr{D}(X)}$ defines an element of $\mathscr{D}'(X)$ that has compact support and finite order.*

Conversely, if $u \in \mathscr{D}'(X)$ has compact support, then $\exists!$ $\tilde{u} \in \mathscr{E}'(X)$ such that $\tilde{u}|_{\mathscr{D}(X)} = u$ in $\mathscr{D}'(X)$ and supp$(\tilde{u}) =$ supp(u).

Remark: In the converse statement, we do not need any assumption on the order of u (since we know elements of $\mathscr{E}'(X)$ have finite order, we might expect this). This is because finite order is implied by u having compact support. Indeed, suppose $u \in \mathscr{D}'(X)$ has compact support. Then set $\rho \in C_c^\infty(X)$ be 1 on $\text{supp}(u)$. Then we have $u = \rho u$ and if $K \subset X$ is compact, for all $\varphi \in \mathscr{D}(X)$ with $\text{supp}(\varphi) \subset K$ we have

$$|\langle u, \varphi \rangle| = |\langle \rho u, \varphi \rangle| = |\langle u, \rho \varphi \rangle|.$$

Now $\rho\varphi$ is smooth of compact support in supp(u). Thus applying the definition of $\mathscr{D}'(X)$ with $K = \text{supp}(\varphi)$ we get

$$|\langle u,\varphi\rangle| = |\langle u,\rho\varphi\rangle| \leq C\sum_{|\alpha|\leq N}\sup_{\text{supp}(u)\cap K}|\partial^\alpha(\rho\varphi)| = C\sum_{|\alpha|\leq N}\sup_K|\partial^\alpha(\rho\varphi)| \leq \tilde{C}\sum_{|\alpha|\leq N}\sup_K\partial^\alpha\varphi$$

where in the last inequality we have used the Leibniz rule for derivatives and then pulled all supremum's involving ρ out. So finally note that this N used was the one for the compact set supp(u) (and so is independent of K), and the above holds for any $\varphi \in \mathscr{D}(X)$ with supp$(\varphi) \subset K$. Hence this shows that u has order $\leq N < \infty$.

Now onto proving the lemma.

Proof. Since $\mathscr{D}(X) \subset \mathscr{E}(X)$, for $u \in \mathscr{E}'(X)$ $u|_{\mathscr{D}(X)}$ is well-defined in $\mathscr{D}'(X)$. Moreover by definition of $\mathscr{E}'(X)$, $\exists$ a compact $K \subset X$ and constants C, N such that

$$|\langle u,\varphi\rangle| \leq C\sum_{|\alpha|\leq N}\sup_K|\partial^\alpha\varphi|$$

for all $\varphi \in \mathscr{D}(X)$ (since $\mathscr{D}(X) \subset \mathscr{E}(X)$).

Thus clearly $u|_{\mathscr{D}(X)} \in \mathscr{D}'(X)$ and supp$(u) \subset K$ and ord$(u) \leq N$.

For the converse, suppose that $u \in \mathscr{D}'(X)$ has compact support, say K. We will first show existence and then uniqueness. Then define $\tilde{u} : \mathscr{E}(X) \to \mathbb{C}$ via:

$$\langle\tilde{u},\varphi\rangle := \langle u,\rho\varphi\rangle$$

for all $\varphi \in \mathscr{E}(X)$, where $\rho \in \mathscr{D}(X)$ is chosen so that $\rho \equiv 1$ on supp(u) (i.e. $\tilde{u} := \rho u$).[vi] Then we have

$$|\langle\tilde{u},\varphi\rangle| = |\langle u,\rho\varphi\rangle| \leq C\sum_{|\alpha|\leq N}\sup_K|\partial^\alpha(\rho\varphi)|$$

since supp$(\rho\varphi) \subset K$ is compact. So by Leibniz, we can expand $\partial^\alpha(\rho\varphi)$ and as $N < \infty$ we can pull out the (finite) factor $\max_{|\alpha|\leq N}\left\{\sup_K|\partial^\alpha\rho|\right\}$ to get:

$$|\langle\tilde{u},\varphi\rangle| \leq \tilde{C}\sum_{|\alpha|\leq N}\sup_K|\partial^\alpha\varphi|$$

and so hence this shows $\tilde{u} \in \mathscr{E}'(X)$. So we have existence of such a $\tilde{u}$.

For uniqueness, suppose $\exists\tilde{v} \in \mathscr{E}'(X)$ such that $\tilde{u}|_{\mathscr{D}(X)} = \tilde{v}|_{\mathscr{D}(X)}$ and supp$(\tilde{u}) = \text{supp}(u) = \text{supp}(\tilde{v})$. Then for any $\varphi \in \mathscr{E}(X)$, write

$$\varphi = \underbrace{\rho\varphi}_{=:\varphi_0} + \underbrace{(1-\rho)\varphi}_{=:\varphi_1}$$

where ρ is as before. Then clearly $\varphi_0 \in \mathscr{D}(X)$ and $\varphi_1 \equiv 0$ on supp$(\tilde{u}) = \text{supp}(\tilde{v})$. So by the definition of support we have

$$\langle\tilde{u},\varphi\rangle = \langle\tilde{u},\varphi_0\rangle + \underbrace{\langle\tilde{u},\varphi_1\rangle}_{=0} = \langle\tilde{v},\varphi_0\rangle + 0 = \langle\tilde{v},\varphi_0\rangle + \underbrace{\langle\tilde{v},\varphi_1\rangle}_{=0} = \langle\tilde{v},\varphi\rangle$$

since $\tilde{u}|_{\mathscr{D}(X)} = \tilde{v}|_{\mathscr{D}(X)}$. Hence $\tilde{u} = \tilde{v}$ and we are done.

[vi] The idea is that u only sees things in supp$(u) = K$. So if we kill off things outside supp(u) our new $\tilde{u}$ should work - we use ρ to do this killing off.

□

Note: This means that we can treat $\mathscr{E}'(X)$ as the subspace of $\mathscr{D}'(X)$ containing the distributions of compact support. Thus as $\mathscr{E}'(X) \subset \mathscr{D}'(X)$ is a subspace, all previous definitions of differentiation, convolutions, etc, for $\mathscr{D}'(X)$ carry over to $\mathscr{E}'(X)$.

[So essentially we have shifted the compact support assumption from the test functions to the distributions.]

2.2. **Convolution between $\mathscr{E}'(\mathbb{R}^n)$ and $\mathscr{D}'(\mathbb{R}^n)$.**

Definition 2.4. *For $\varphi \in \mathscr{E}(\mathbb{R}^n)$ and $u \in \mathscr{E}'(\mathbb{R}^n)$ define (as usual) the* ***convolution*** *by*

$$(u * \varphi)(x) := \langle u, \tau_x \check{\varphi} \rangle.$$

As before we have $u * \varphi \in C^\infty(\mathbb{R}^n)$ and $\partial^\alpha(u * \varphi) = u * \partial^\alpha \varphi$.

So we have defined convolutions for $\mathscr{D}(X)$ with $\mathscr{D}'(X)$ and $\mathscr{E}(X)$ with $\mathscr{E}'(X)$. Since $\mathscr{D}(X) \subset \mathscr{E}(X)$ we can thus define the convolution of $\mathscr{D}(X)$ with $\mathscr{E}'(X)$. But both of these have compact support, and so we might expect the convolution to have compact support.

Indeed, if we have $\varphi \in \mathscr{D}(\mathbb{R}^n)$ and $u \in \mathscr{E}'(\mathbb{R}^n)$, then

$$0 = (u * \varphi)(x) = \langle u, \tau_x \check{\varphi} \rangle \text{ “ = ” } \langle u(y), \varphi(x - y) \rangle$$

and thus this is zero unless $(x - y) \in \text{supp}(\varphi)$ for some $y \in \text{supp}(u)$. So we see that

$$\text{supp}(u * \varphi) \subset \text{supp}(u) + \text{supp}(\varphi)$$

and so $\text{supp}(u * \varphi)$ is compact as both sets on the RHS are comapct. So hence in this case we see that $u * \varphi \in \mathscr{D}(\mathbb{R}^n)$, i.e.

$$* : \mathscr{E}(\mathbb{R}^n) \times \mathscr{E}'(\mathbb{R}^n) \to \mathscr{E}(\mathbb{R}^n) \quad \text{and} \quad * : \mathscr{D}(\mathbb{R}^n) \times \mathscr{E}'(\mathbb{R}^n) \to \mathscr{D}(\mathbb{R}^n)$$

and from before we know

$$* : \mathscr{D}(\mathbb{R}^n) \times \mathscr{D}'(\mathbb{R}^n) \to \mathscr{E}(\mathbb{R}^n).$$

Thus as the result gives another function in $\mathscr{D}(\mathbb{R}^n)$ in this case, we can convolute it with any element of $\mathscr{D}'(X)$. Thus we can define:

Definition 2.5 (Convolution of Distributions). *Suppose $u_1, u_2 \in \mathscr{D}'(R^n)$ with at least one of u_1, u_2 having compact support (i.e. in $\mathscr{E}'(\mathbb{R}^n)$). Then we can define $u_1 * u_2$ by:*

$$(u_1 * u_2)(\varphi) := u_1 * (u_2 * \varphi) \qquad \forall \varphi \in \mathscr{D}(\mathbb{R}^n).$$

Note: As above, this is well-defined. Indeed, if $u_1 \in \mathscr{E}'(\mathbb{R}^n)$, then $u_2 * \varphi \in \mathscr{E}(\mathbb{R}^n)$, and so $u_1 * (u_2 * \varphi)$ is well-defined via Definition 2.4. However if $u_2 \in \mathscr{E}'(\mathbb{R}^n)$ then $u_2 * \varphi \in \mathscr{D}(\mathbb{R}^n)$, and so $u_1 * (u_2 * \varphi)$ makes sense in the way of Definition 1.6.

Remark: One can show that $u_1 * u_2 \in \mathscr{D}'(R^n)$ - see Example Sheet 2.

Lemma 2.3. *For $u_1, u_2 \in \mathscr{D}'(\mathbb{R}^n)$ with at least one of u_1, u_2 in $\mathscr{E}'(\mathbb{R}^n)$, we have*

$$u_1 * u_2 = u_2 * u_1$$

i.e. $*$ is commutative.

Proof. Let $\varphi, \psi \in \mathscr{D}(\mathbb{R}^n)$. Then by repeated application of Lemma 1.5 we have

$$\begin{aligned}(u_1 * u_2) * (\varphi * \psi) &:= u_1 * (u_2 * (\varphi * \psi))\\ &= u_1 * ((u_2 * \varphi) * \psi) \quad \text{by Lemma 1.5}\\ &= u_1 * (\psi * (u_2 * \varphi)) \quad \text{since } f * g = g * f \text{ for smooth functions}\\ &= (u_1 * \psi) * (u_2 * \varphi) \quad \text{by Lemma 1.5.}\end{aligned}$$

Now by symmetry in the above we have

$$(u_2 * u_1) * (\varphi * \psi) = (u_2 * u_1) * (\psi * \varphi) = (u_2 * \varphi) * (u_1 * \psi) = (u_1 * \psi) * (u_2 * \varphi)$$

where in the first and third equalities we have used the fact that $f * g = g * f$ for smooth functions. Hence we see

$$(u_1 * u_2) * (\varphi * \psi) = (u_2 * u_1) * (\varphi * \psi)$$

i.e. setting $E = u_1 * u_2 - u_2 * u_1$,

$$E * (\varphi * \psi) = 0 \quad \forall \varphi, \psi \in \mathscr{D}(\mathbb{R}^n).$$

But then using the trick that $\langle u, \varphi \rangle = (u * \check{\varphi})(0)$ we have

$$\langle E * \varphi, \psi \rangle = ((E * \varphi) * \check{\psi})(0) = 0 \quad \forall \varphi, \psi \quad \implies \quad E * \varphi = 0 \quad \forall \varphi$$

and so

$$\langle E, \varphi \rangle = (E * \check{\varphi})(0) = 0 \quad \forall \varphi \quad \implies \quad E = 0 \quad \text{in } \mathscr{D}'(\mathbb{R}^n)$$

and so we are done.

$\square$

Now if $u \in \mathscr{D}'(\mathbb{R}^n)$ we can make sense of $u * \delta_0$. Indeed we have:

$$(u * \delta_0) * \psi := u * (\delta_0 * \psi) = u * \psi$$

for all $\psi \in \mathscr{D}(\mathbb{R}^n)$, where we have used that

$$(\delta_0 * \psi)(x) := \langle \delta_0, \tau_x \check{\tau} \rangle = (\tau_x \check{\psi})(0) = \psi(0) \quad \text{i.e.} \quad \delta_0 * \psi = \psi \ \text{ as functions.}$$

So hence we see that $u * \delta_0 = u$ as distributions, for all $u \in \mathscr{D}'(\mathbb{R}^n)$ [compare this with the heuristic formula, "$u(x) = \int \delta_0(x-y) u(y)\, \mathrm{d}y$".]

3. Tempered Distributions and Fourier Analysis

3.1. Functions of Rapid Decay.

We have another new set of test functions:

Definition 3.1. *A smooth function $\varphi : \mathbb{R}^n \to \mathbb{C}$ is called a* ***Schwartz function*** *if:*

$$\|\varphi\|_{\alpha,\beta} := \sup_{\mathbb{R}^n} \left|x^\alpha \partial^\beta \varphi\right| < \infty$$

for each multi-index α, β (i.e. each semi-norm is finite).

We write $\mathbf{S}(\mathbb{R}^\mathbf{n})$ for the space of all Schwartz functions on $\mathbb{R}^n$, with the topology induced by the semi-norms, i.e.

$$\varphi_n \to 0 \ \text{ in } S(\mathbb{R}^n) \quad \text{if} \quad \|\varphi_m\|_{\alpha,\beta} \to 0 \quad \forall \alpha, \beta.$$

Note: This makes $S(\mathbb{R}^n)$ into a locally convex space.

The intuition is that Schwartz functions are those smooth functions whose derivatives tends to 0 faster than any polynomial. The standard example of a Schwartz function is the Gaussian, $\varphi(x) = e^{-|x|^2}$.

Note: Clearly we have $C_c^\infty(\mathbb{R}^n) \subset S(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n)$.

Definition 3.2. *A linear map $u : S(\mathbb{R}^n) \to \mathbb{C}$ is called a* ***tempered distribution*** *if $\exists$ constants C, N such that:*

$$|\langle u, \varphi\rangle| \leq C \sum_{|\alpha|,|\beta|\leq N} \|\varphi\|_{\alpha,\beta}$$

i.e. they are distributions on $S(\mathbb{R}^n)$. We write $\mathbf{S}'(\mathbb{R}^\mathbf{n})$ for the space of tempered distributions.

Again we can give an equivalent definition of $S'(\mathbb{R}^n)$ in terms of sequential continuity [**Exercise** to check].

Remark: We can show that:

$$\mathscr{E}'(\mathbb{R}^n) \hookrightarrow S'(\mathbb{R}^n) \hookrightarrow \mathscr{D}'(\mathbb{R}^n)$$

and those embeddings are in fact continuous embeddings[(vii)]. This is essentially a consequence of $\mathscr{D}(\mathbb{R}^n) \hookrightarrow S(\mathbb{R}^n) \hookrightarrow \mathscr{E}(\mathbb{R}^n)$.

We can define differentiation on $S'(\mathbb{R}^n)$ as before, as well as convolution between $S(\mathbb{R}^n)$ and $S'(\mathbb{R}^n)$. Note that we cannot however multiply a Schwartz function by a smooth function and hope to obtain another Schwartz function: consider

$$\exp(-|x|^2) \cdot \exp(|x|^2) = 1,$$

[(vii)] We say that X is **continuously embedded** in Y, written $X \hookrightarrow Y$, if $X \subset Y$ (or at least X is isomorphic to a subspace of Y) and whenever $x_n \to 0$ in X we have $x_n \to 0$ in Y. Since the subspace topology on X is the smallest topology on X for which the inclusion is continuous, this means the topology on X contains the subspace topology.

which is the product of a Schwartz function with a smooth function which doesn't give a Schwartz function. This means that we can't define multiplication by arbitrary smooth functions on $S'(\mathbb{R}^n)$, but only by those which preserve Schwartz functions.

3.2. **Fourier Transform on $S(\mathbb{R}^n)$.**

Definition 3.3. *We define the **Fourier transform** on $L^1(\mathbb{R}^n)$, $\mathscr{F} : f \mapsto \hat{f}$, by:*

$$\hat{f}(\lambda) := \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} f(x)\, \mathrm{d}x \qquad \text{for } \lambda \in \mathbb{R}^n.$$

Note that we have $S(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, since for $\varphi \in S(\mathbb{R}^n)$:

$$\begin{aligned}\int_{\mathbb{R}^n} |\varphi|\, \mathrm{d}x &= \int_{\mathbb{R}^n} (1+|x|)^{-N} \cdot (1+|x|)^n |\varphi|\, \mathrm{d}x \\ &\leq C \sum_{|\alpha|\leq N} \|\varphi\|_{\alpha,0} \int_{\mathbb{R}^n} \frac{1}{(1+|x|)^N}\, \mathrm{d}x \\ &< \infty\end{aligned}$$

for N sufficiently large.

So hence we can take the Fourier transform of Schwartz functions.

Lemma 3.1. *We have $\mathscr{F} : L_1(\mathbb{R}^n) \to C(\mathbb{R}^n)$, i.e. if $f \in L^1(\mathbb{R}^n)$, then $\hat{f} \in C(\mathbb{R}^n)$.*

Proof. This is just dominated convergence, since if $\lambda_m \to \lambda$ in $\mathbb{R}^n$ we have

$$\hat{f}(\lambda) = \int_{\mathbb{R}^n} e^{-i\lambda_m\cdot x} f(x) \to \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} f(x)\, \mathrm{d}x = \hat{f}(\lambda)$$

where the integrals convergence by dominated convergence, since $|e^{-i\lambda_m\cdot x} f(x)| \leq |f(x)| \in L_1(\mathbb{R}^n)$. Hence this shows that $\hat{f} \in C(\mathbb{R}^n)$.

□

This lemma tells us that if f decays fast enough at ∞ (e.g. so is in L_1), then the Fourier transform of f is more regular (e.g. continuous). So what happens if f were to decay even faster at ∞? Would $\mathscr{F}(f)$ be more regular? This turns out to be the key point of the Fourier transform: it exchanges decay (which tends to be easy to check in practice) with regularity (which tends to be hard to check).

[So when solving a Fourier transformed PDE, just by checking the decay rate of the Fourier transformed solution can tell us about the regularity of the original solution.]

Lemma 3.2. *If $\varphi \in S(\mathbb{R}^n)$, then:*

$$\widehat{(D^\alpha \varphi)}(\lambda) = \lambda^\alpha \hat{\varphi}(\lambda) \quad \text{and} \quad \widehat{\left(x^\beta \varphi\right)}(\lambda) = (-1)^{|\beta|} D^\beta \hat{\varphi}(\lambda)$$

where we write $D \equiv -i\partial$, for ∂ the usual derivative[viii].

Proof. On integrating by parts:

$$\begin{aligned}
\widehat{(D_x^\alpha \varphi)}(\lambda) &= \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} D_x^\alpha \varphi \, \mathrm{d}x \\
&= (-1)^{|\alpha|} \int_{\mathbb{R}^n} \varphi \cdot \underbrace{D_x^\alpha\left[e^{-i\lambda\cdot x}\right]}_{(-1)^{|\alpha|}\lambda^\alpha e^{-i\lambda\cdot x}} \, \mathrm{d}x \\
&= \lambda^\alpha \int_{\mathbb{R}^n} \varphi e^{-i\lambda\cdot x} \, \mathrm{d}x \\
&= \lambda^\alpha \hat{\varphi}(\lambda).
\end{aligned}$$

And by differentiating under the integral,

$$\begin{aligned}
D_\lambda^\beta \hat{\varphi}(\lambda) &= D_\lambda^\beta \left(\int_{\mathbb{R}^n} e^{-i\lambda\cdot x} \varphi(x) \, \mathrm{d}x \right) \\
&= \int_{\mathbb{R}^n} (-1)^{|\beta|} x^\beta e^{-i\lambda\cdot x} \varphi(x) \, \mathrm{d}x \\
&= (-1)^{|\beta|} \widehat{\left(x^\beta \varphi\right)}(\lambda)
\end{aligned}$$

where we have used dominated convergence to justify differentiating under the integral sign.

□

Important: The above lemma tells us exactly how the Fourier transform interchanges local regularity with decay at ∞, and vice versa. For example, $D^\alpha \varphi$ is 'less regular' than φ, and so we would expect that $\widehat{(D^\alpha \varphi)}$ should 'decay slower' than $\hat{\varphi}$, which is what the above shows since it gets multiplied by a polynomial term λ^α.

We may have seen the "Fourier inversion formula" before, which gives:

$$\varphi(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\lambda\cdot x} \hat{\varphi}(\lambda) \, \mathrm{d}\lambda.$$

We will prove this in the following result:

Theorem 3.1. *The Fourier transform maps $S(\mathbb{R}^n)$ to itself, and is a continuous isomorphism here, i.e.*

$$\mathscr{F} : S(\mathbb{R}^n) \to S(\mathbb{R}^n) \quad \text{is a } \underline{\text{continuous isomorphism}}$$

[viii] This is convention for Fourier transforms as it removes factors of i from the derivatives.

[so in particular it is bounded (as $\mathscr{F}$ is linear), and if $\varphi_m \to 0$ in $S(\mathbb{R}^n)$, then $\hat{\varphi}_m \to 0$ in $S(\mathbb{R}^n)$].

Proof. First we need to show that $\mathscr{F}$ maps $S(\mathbb{R}^n)$ to $S(\mathbb{R}^n)$. So if $\varphi \in S(\mathbb{R}^n)$, we have:

$$|\lambda^\alpha D^\beta \hat{\varphi}| = \left|\int_{\mathbb{R}^n} e^{-i\lambda\cdot x} D^\alpha(x^\beta \varphi)\, \mathrm{d}x\right| < \int_{\mathbb{R}^n} |D^\alpha(x^\beta\varphi)|\, \mathrm{d}x < \infty$$

and thus we see that $\|\hat{\varphi}\|_{\alpha,\beta} < \infty$ for each α, β. Also, $\hat{\varphi}$ is smooth from Lemmas 3.1 and 3.2. Thus $\hat{\varphi} \in S(\mathbb{R}^n)$.

We also see that $\hat{\varphi}_m \to 0$ in $S(\mathbb{R}^n)$ if $\varphi_m \to 0$ in $S(\mathbb{R}^n)$.

So we now just need infectivity and surjectivity, which will come from Fourier inversion. So consider the quantity:

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\lambda\cdot x}\hat{\varphi}(\lambda)\, \mathrm{d}\lambda = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\lambda\cdot x}\left[\int_{\mathbb{R}^n} e^{-i\lambda\cdot y}\varphi(y)\, \mathrm{d}y\right] \mathrm{d}x.$$

Note that the function $(\lambda, y) \mapsto e^{i\lambda\cdot(x-y)}\varphi(y)$ is <u>not</u> necessarily absolutely integrable, and so we cannot apply Fubini to interchange the order of integration.[(ix)] So to get around this, we use dominated convergence and insert a regulariser. Indeed, note by dominated convergence that

$$\int_{\mathbb{R}^n} e^{i\lambda\cdot x}\hat{\varphi}(\lambda)\, \mathrm{d}\lambda = \lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^n} e^{i\lambda\cdot x - \varepsilon|\lambda|^2}\hat{\varphi}(\lambda)\, \mathrm{d}\lambda$$

where we can use dominated convergence since $|e^{i\lambda\cdot x}e^{-\varepsilon|\lambda|^2}\hat{\varphi}| \leq |\hat{\varphi}| \in L_1(\mathbb{R}^n)$.

Now for each $\varepsilon > 0$, we can use Fubini and so:

$$\begin{aligned}
\int_{\mathbb{R}^n} e^{i\lambda\cdot x-\varepsilon|\lambda|^2}\hat{\varphi}(\lambda)\, \mathrm{d}\lambda &= \int_{\mathbb{R}^n}\varphi(y)\left[\int_{\mathbb{R}^n} e^{i\lambda\cdot(x-y)-\varepsilon|\lambda|^2}\, \mathrm{d}\lambda\right]\mathrm{d}y\\
&= \int_{\mathbb{R}^n}\varphi(y)\left[\prod_{j=1}^n\int_{\mathbb{R}} e^{i\lambda_j(x_j-y_j)-\varepsilon\lambda_j^2}\, \mathrm{d}\lambda_j\right]\mathrm{d}y\\
(\star)\quad &= \int_{\mathbb{R}^n}\varphi(y)\left[\prod_{j=1}^n\left(\frac{\pi}{\varepsilon}\right)^{1/2} e^{-\frac{(x_j-y_j)^2}{4\varepsilon}}\right]\mathrm{d}y\\
&= \left(\frac{\pi}{\varepsilon}\right)^{n/2}\int_{\mathbb{R}^n}\varphi(y)e^{-\frac{|x-y|^2}{4\varepsilon}}\, \mathrm{d}y\\
(\dagger)\quad &= \left(\frac{\pi}{\varepsilon}\right)^{n/2} 2^n\cdot\varepsilon^{n/2}\int_{\mathbb{R}^n}\varphi(x-2\sqrt{\varepsilon}y')e^{-|y'|^2}\, \mathrm{d}y'\\
&= \pi^{n/2}\cdot 2^n\cdot\int_{\mathbb{R}^n}\varphi(x-2\sqrt{\varepsilon}y)e^{-|y|^2}\, \mathrm{d}y
\end{aligned}$$

(ix) Recall that Fubini's theorem says us that if $\int\int |f(x,y)|\, \mathrm{d}x\mathrm{d}y < \infty$, i.e. f is absolutely integrable, then we can interchange the order of integration.

where in (†) we changed variables $y' = (x-y)/2\sqrt{\varepsilon}$, and we shall explain ($\star$) momentarily. We can then use dominated convergence in this last line since pointwise $\varphi(x-2\sqrt{\varepsilon}y) \to \varphi(x)$ as $\varepsilon \to 0$, to get that this

$$\begin{aligned}
&\to \pi^{n/2}2^n \int_{\mathbb{R}^n} \varphi(x)e^{-|y|^2}\, \mathrm{d}y\\
&= \pi^{n/2}2^n\varphi(x)\int_{\mathbb{R}^n} e^{-|y|^2}\, \mathrm{d}y\\
&= (2\pi)^n\varphi(x)
\end{aligned}$$

since the last integral is a product of n-Gaussian integrals and so equals $\pi^{n/2}$. So combining all of this we see:

$$\varphi(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\lambda\cdot x}\hat{\varphi}(\lambda)\, \mathrm{d}\lambda.$$

So hence $\mathscr{F}$ is injective since this shows if $\hat{\varphi}$, then we must have $\varphi \equiv 0$. Moreover this shows that $\mathscr{F}$ is surjective since:

$$\varphi(x) := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-i\lambda\cdot x}\hat{\varphi}(-\lambda)\, \mathrm{d}y = \int_{\mathbb{R}^n} e^{-i\lambda\cdot x}\left[\frac{1}{(2\pi)^n}\hat{\varphi}(-\lambda)\right] \mathrm{d}\lambda = \mathscr{F}\,[\underbrace{\left[\frac{1}{(2\pi)^n}\hat{\varphi}(-\lambda)\right]}_{\in S(\mathbb{R}^n)}$$

and thus we are done.

$\square$

Aside: Fokas and Gelfand provided a beautiful proof of this theorem via the Riemann-Hilbert problem.

Note: At ($\star$) in the above proof we have used that:

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{iz\alpha-\varepsilon z^2}\, \mathrm{d}z &= \int_{-\infty}^{\infty} e^{-\varepsilon\left(z-\frac{i\alpha}{2\varepsilon}\right)^2}\cdot e^{-\frac{\alpha^2}{4\varepsilon}}\, \mathrm{d}z\\
&= e^{-\alpha^2/4\varepsilon}\int_{-\infty}^{\infty} e^{-\varepsilon\left(z-\frac{i\alpha}{2\varepsilon}\right)^2}\, \mathrm{d}z\\
&= e^{-\alpha^2/4\varepsilon}\int_{\mathrm{Im}(z)=\alpha/2\varepsilon} e^{-\varepsilon\left(z-\frac{i\alpha}{2\varepsilon}\right)^2}\, \mathrm{d}z \qquad \text{by Cauchy's theorem from Complex Analysis}\\
&= e^{-\alpha^2/4\varepsilon}\underbrace{\int_{-\infty}^{\infty} e^{-\varepsilon z^2}\, \mathrm{d}z}_{=(\pi/\varepsilon)^{1/2}}\\
&= \left(\frac{\pi}{\varepsilon}\right)^2 e^{-\alpha^2/4\varepsilon}.
\end{aligned}$$

We apply Cauchy's theorem on the boundary of a box of the form

$$\{z \in \mathbb{C} : \mathfrak{R}(z) \in [-N,N], \mathfrak{I}(z) \in [0,\alpha/2\varepsilon]\}$$

where the integrals on the side parts (of constant real part) vanish as we take $N \to \infty$.

We will now use all of this to define the Fourier transform of a tempered distribution.

3.3. **Fourier Transform of $S'(\mathbb{R}^n)$.**

We need the following lemma to motivative the definition of $\mathscr{F}$ on $S'(\mathbb{R}^n)$.

Lemma 3.3 (Parseval's Theorem). *For $\varphi, \psi \in S(\mathbb{R}^n)$ we have*

$$\int_{\mathbb{R}^n} \varphi(x)\hat{\psi}(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \hat{\varphi}(x)\psi(x)\,\mathrm{d}x$$

or alternatively, if $(\cdot,\cdot)$ is the L_2-inner product, this says $(\varphi, \mathscr{F}(\psi)) = (\mathscr{F}(\varphi), \psi)$, i.e. $\mathscr{F}$ is self-adjoint here.

Proof. By Fubini's theorem we have

$$\begin{aligned}\int_{\mathbb{R}^n} \varphi(x)\hat{\psi}(x)\,\mathrm{d}x &= \int_{\mathbb{R}^n} \varphi(x)\left[\int_{\mathbb{R}^n} e^{-i\lambda\cdot x}\psi(\lambda)\,\mathrm{d}\lambda\right]\mathrm{d}x\\ &= \int_{\mathbb{R}^n} \psi(\lambda)\left[\int_{\mathbb{R}^n} e^{-i\lambda\cdot x}\varphi(x)\,\mathrm{d}x\right]\mathrm{d}\lambda\\ &= \int_{\mathbb{R}^n} \psi(\lambda)\hat{\varphi}(\lambda)\,\mathrm{d}\lambda\end{aligned}$$

where we have been able to apply Fubini since the integral is absolutely convergent, i.e. since $(x,\lambda) \mapsto e^{-i\lambda\cdot x}\varphi(x)\psi(\lambda)$ is in $L_1(\mathbb{R}^n \times \mathbb{R}^n)$. Hence we are done.

$\square$

Treating $u \in S(\mathbb{R}^n)$ as its associated tempered distribution, this reads:

$$\langle \hat{u}, \varphi\rangle = \langle u, \hat{\varphi}\rangle.$$

But the RHS os well-defined for <u>any</u> $u \in S'(\mathbb{R}^n)$, since $\mathscr{F} : S(\mathbb{R}^n) \to S(\mathbb{R}^n)$. Thus we define in general:

Definition 3.4. *For $u \in S'(\mathbb{R}^n)$, we define the* ***Fourier transform*** *of the tempered distribution u by:*

$$\langle \hat{u}, \varphi\rangle := \langle u, \hat{\varphi}\rangle \qquad \forall \varphi \in S(\mathbb{R}^n).$$

So we can always take the Fourier transform of a tempered distribution.

Example 3.1. *Take $u = \delta_0 \in S'(\mathbb{R}^n)$ (noting that all distributions of compact support are tempered distributions). Then:*

$$\langle \hat{\delta}_0, \varphi\rangle := \langle \delta_0, \hat{\varphi}\rangle = \hat{\varphi}(0) = \int_{\mathbb{R}^n} \varphi(x)\,\mathrm{d}x = \langle 1, \varphi\rangle$$

and this is true for all $\varphi \in S(\mathbb{R}^n)$, and thus $\hat{\delta}_0 = 1$.

Also,

$$\langle \hat{1}, \varphi \rangle := \langle 1, \hat{\varphi} \rangle = \int_{\mathbb{R}^n} \hat{\varphi}(\lambda)\, \mathrm{d}\lambda = (2\pi)^n \varphi(0) = \langle (2\pi)^n \delta_0, \varphi \rangle$$

where we have used Fourier inversion to evaluate the integral of $\hat{\varphi}$. Thus we see:

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

or in old, heuristic language,

$$\text{“}\delta_0(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\lambda\cdot x}\, \mathrm{d}\lambda\text{”.}$$

We can extend Lemma 3.2 to $S'(\mathbb{R}^n)$, so that

$$\widehat{D^\alpha u} = \lambda^\alpha \hat{u} \qquad \text{and} \qquad \widehat{x^\beta u} = (-1)^{|\beta|} D^\beta \hat{u}$$

for all $u \in S'(\mathbb{R}^n)$, where the derivatives here are distributional derivatives [see Example Sheet 2 for a proof].

So now we shall see that the Fourier transform of a tempered distribution is another tempered distribution. This will enable us to try to solve PDEs in the space of tempered distributions (which includes a lot of functions when we identify functions by their associated functions) by taking the Fourier transform of the PDE.

Theorem 3.2. *The Fourier transform extends to a continuous isomorphism:*

$$\mathscr{F} : S'(\mathbb{R}^n) \to S'(\mathbb{R}^n).$$

Proof. Suppose that $\varphi_m \to 0$ in $S(\mathbb{R}^n)$. Then we know by Theorem 3.1 that $\hat{\varphi}_m \to 0$ in $S(\mathbb{R}^n)$. So hence:

$$\langle \hat{u}, \varphi_m \rangle := \langle u, \hat{\varphi}_m \rangle \to 0$$

since $u \in S'(\mathbb{R}^n)$. So by the sequential continuity definition of $S'(\mathbb{R}^n)$, we deduce that $\hat{u} \in S'(\mathbb{R}^n)$ (so $\mathscr{F}$ does map $S'(\mathbb{R}^n)$ into $S'(\mathbb{R}^n)$).

We also claim that:

$$\check{u} = (2\pi)^{-n} \widehat{(\hat{u})}.$$

Indeed,

$$\langle \check{u}, \varphi \rangle = \langle u, \check{\varphi} \rangle \overset{(\dagger)}{=} \left\langle u, (2\pi)^{-n} \widehat{(\hat{\varphi})} \right\rangle = \left\langle (2\pi)^{-n} \widehat{(\hat{u})}, \varphi \right\rangle$$

for all $\varphi \in S(\mathbb{R}^n)$, where ($\dagger$) follows from the Fourier inversion theorem since

$$\check{\varphi}(x) = \varphi(-x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} \hat{\varphi}(\lambda)\, \mathrm{d}\lambda = (2\pi)^{-n} \widehat{(\hat{\varphi})}(x)$$

and so we have a Fourier inversion result for tempered distributions, which as before implies that $\mathscr{F}$ is injective and surjective.

Finally for continuity, we have that if $u_m \to 0$ in $S'(\mathbb{R}^n)$, then:

$$\begin{aligned} &\langle u_m, \varphi\rangle \to 0 \qquad \forall \varphi \in S(\mathbb{R}^n) \\ \iff\ &\langle u_m, \hat{\varphi}\rangle \to 0 \qquad \forall \varphi \in S(\mathbb{R}^n) \\ \iff\ &\langle \hat{u}_m, \varphi\rangle \to 0 \qquad \forall \varphi \in S(\mathbb{R}^n) \end{aligned}$$

i.e. $\hat{u}_m \to 0$ in $S'(\mathbb{R}^n)$. So done.

□

3.4. **Sobolev Spaces.**

The Sobolev spaces are a very useful space of functions.

Definition 3.5. *For $s \in \mathbb{R}$, let $\boldsymbol{H^s(\mathbb{R}^n)}$, denote the set of $u \in S'(\mathbb{R}^n)$ for which $\hat{u} \in S'(\mathbb{R}^n)$ can be identified with a locally integrable function $\lambda \mapsto \hat{u}(\lambda)$ for which:*

$$\|u\|_{H^s}^2 := \int_{\mathbb{R}^n} |\hat{u}(\lambda)|^2 \langle\lambda\rangle^{2s}\, d\lambda < \infty$$

*where $\langle\lambda\rangle := (1+|\lambda|^2)^{1/2}$ is called the **Japanese bracket** (for some reason...).*

Note: Since the Fourier transform exchanges differentiability and decay, this is essentially asking for s-differentiability of u.

We can define Sobolev spaces on $X \subset \mathbb{R}^n$ via localisation: note that if $u \in \mathscr{D}'(X)$ and $\varphi \in \mathscr{D}(X)$, then $\varphi u \in \mathscr{E}'(\mathbb{R}^n) \subset S'(\mathbb{R}^n)$.

Definition 3.6. *We say $u \in \boldsymbol{H^s_{loc}(X)}$ if $u \in \mathscr{D}'(X)$ and $\varphi u \in H^s(\mathbb{R}^n)$ for every $\varphi \in \mathscr{D}(X)$. This is the **local Sobolev space**.*

A key first result is the Sobolev embedding, which tells us when we can infer regularity of $u \in H^s(\mathbb{R}^n)$.

Lemma 3.4 (Sobolev Lemma). *Let $u \in H^s(\mathbb{R}^n)$. Then if $s > n/2$, then $u \in C(\mathbb{R}^n)$.*

Proof. By Cauchy-Schwarz,

$$\begin{aligned} \int_{\mathbb{R}^n} |\hat{u}(\lambda)|\, d\lambda &= \int_{\mathbb{R}^n} \langle\lambda\rangle^{-s}\cdot\langle\lambda\rangle^{s}|\hat{u}(\lambda)|\, d\lambda \\ &\leq \left(\int_{\mathbb{R}^n} \langle\lambda\rangle^{-2s}\right)^{1/2}\left(\int_{\mathbb{R}^n} |\hat{u}(\lambda)|^2\langle\lambda\rangle^{2s}\, d\lambda\right)^{1/2} \\ &= \|u\|_{H^s}\cdot\left(\int_{S^{n-1}} d\sigma \int_0^\infty (1+r^2)^{-s} r^{n-1}\, dr\right)^{1/2} \end{aligned}$$

where we have changed to polar coordinates (and so $\mathrm{d}\lambda = \mathrm{d}\sigma \cdot r^{n-1}\mathrm{d}r$, where $\mathrm{d}\sigma$ is the area element on S^{n-1}).

Now note that if $s = \frac{n}{2} + \varepsilon$, for $\varepsilon > 0$, then

$$(1+r^2)^{-s} r^{n-1} \sim r^{-2s+n-1} = r^{-1-2\varepsilon}$$

for $r \gg 0$, which is integrable. Hence this shows that the integral on the RHS is $< \infty$, and so

$$\int_{\mathbb{R}^n} |\hat{u}(\lambda)| \, \mathrm{d}\lambda < \infty \qquad \text{i.e.} \quad \hat{u} \in L^1(\mathbb{R}^n).$$

Now for $\varphi \in S(\mathbb{R}^n)$,

$$\begin{aligned}
\langle u, \hat{\varphi} \rangle = \langle \hat{u}, \varphi \rangle &= \int_{\mathbb{R}^n} \hat{u}(\lambda)\varphi(\lambda) \, \mathrm{d}\lambda \\
&= \int_{\mathbb{R}^n} \hat{u}(\lambda) \left[\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\lambda\cdot x} \hat{\varphi}(x) \, \mathrm{d}x \right] \mathrm{d}\lambda \\
&= \int_{\mathbb{R}^n} \hat{\varphi}(x) \underbrace{\left[\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\lambda\cdot x} \hat{u}(\lambda) \, \mathrm{d}\lambda \right]}_{=:F(x)} \mathrm{d}x \\
&= \langle F, \hat{\varphi} \rangle
\end{aligned}$$

where we have used Fubini's theorem, which we can do since $\hat{u} \in L_1(\mathbb{R}^n)$ and so $(x, \lambda) \mapsto \hat{\varphi}(x)\hat{u}(\lambda)e^{i\lambda\cdot x}$ is absolutely integrable (i.e. in $L_1(\mathbb{R}^n \times \mathbb{R}^n)$).

So hence we can identify $u \in S'(\mathbb{R}^n)$ with the function F above (as the above says $\hat{u} = \hat{F}$), and so we have

$$u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\lambda\cdot x} \hat{u}(\lambda) \, \mathrm{d}\lambda.$$

But we know the Fourier transform of an integrable function is continuous (by dominated convergence - this is Lemma 3.1), and thus this shows $u \in C(\mathbb{R}^n)$.

□

Note: In the above we are trying to deduce a Fourier inversion formula for functions in L^1, which we don't know holds a priori as we have only seen it for $S(\mathbb{R}^n)$. Hence we need to check it directly in the above.

Corollary 3.1. *Suppose $u \in H^s(\mathbb{R}^n)$ for every $s > n/2$. Then $u \in C^\infty(\mathbb{R}^n)$.*

Proof. Follows from Lemma 3.4, since $D^\alpha u$ will be in $H^s(\mathbb{R}^n)$ for all $s > n/2$ (as differentiating reduces the exponent s, but we have this for all $s > n/2$). So hence we can just apply Lemma 3.4 inductively to get all the $D^\alpha u$ can be identified with continuous functions, and thus u can be identified with a smooth function. [**Exercise** to check details more thoroughly.]

□

4. Applications of the Fourier Transform

4.1. **Elliptic Regularity.**

We are now interested in PDE's of the form: $P(D)u = f$, where $u, f \in \mathscr{D}'(X)$ and P is a polynomial (recall that $D = -i\partial$). Note that these are PDEs with constant coefficients. For example, if $P(\lambda) = -(\lambda_1^2 + \cdots + \lambda_n^2)$ then

$$P(D) = -\left(\left(-i\frac{\partial}{\partial x_1}\right)^2 + \cdots + \left(-i\frac{\partial}{\partial x_n}\right)^2\right) = \Delta$$

is the Laplacian.

The main question we are interested in here is one of *regularity*: if $f \in H^s_{\text{loc}}(X)$, then what can be said of the solution u (assuming it exists)?

Definition 4.1. *For a differential operator $P(D) = \sum_{|\alpha|\leq N} c_\alpha D^\alpha$, the **order** of P is N (assuming that not all a_α with $|\alpha| = N$ are 0).*

*The **principal symbol** of P is then:*

$$\sigma_P(\lambda) := \sum_{|\alpha|=N} c_\alpha \lambda^\alpha.$$

*We say that $P(D)$ is **elliptic** if $\sigma_P(\lambda) \neq 0$ for all $\lambda \in \mathbb{R}^n\backslash\{0\}$, i.e. if $\ker(\sigma) = \{0\}$.*

Lemma 4.1. *If $P(D)$ is an N'th order elliptic operator, then $\exists$ constants C, R such that*

$$|P(\lambda)| \geq C\langle\lambda\rangle^N \qquad \text{for all } |\lambda| > R.$$

Remark: Heuristically, if we tried to solve $P(D)u = f$, we could take the Fourier transform of the equation to get $P(\lambda)\hat{u} = \hat{f}$, since the Fourier transform changes D's into λ's. So then we would want to say $\hat{u} = \hat{f}/P(\lambda)$ and take the inverse Fourier transform to find the solution u, but P could be zero and this may not be integrable! The above Lemma 4.1 tells us that this isn't the case if we take λ far enough from 0.

Proof. Since P is elliptic we know $\sigma_P(\lambda) \neq 0$ for all $\lambda \in S^{n-1}$. Thus by compactness of S^{n-1}, σ_P attains its (non-zero) minima on S^{n-1} and thus $\exists c_0 > 0$ such that

$$\min_{\lambda\in S^{n-1}} |\sigma_P(\lambda)| = c_0 > 0.$$

Then for any $\lambda \neq 0$,

$$|\sigma_P(\lambda)| = \left|\sum_{|\alpha|=N} c_\alpha\lambda^\alpha\right| = |\lambda|^N \cdot \underbrace{\left|\sum_{|\alpha|=N} c_\alpha\left(\frac{\lambda}{|\lambda|}\right)^\alpha\right|}_{=\sigma_P(\lambda/|\lambda|)\,\geq\, c_0} \geq c_0|\lambda|^N.$$

But by the triangle inequality,

$$\begin{aligned}|P(\lambda)| &\geq |\sigma_P(\lambda)| - |P(\lambda) - \sigma_P(\lambda)| \\ &\geq c_0\lambda^N - |\lambda|^N \left[\frac{P(\lambda) - \sigma_P(\lambda)}{|\lambda|^N}\right].\end{aligned}$$

Now since $P(\lambda) - \sigma_P(\lambda)$ is a polynomial of degree $N-1$, we can choose $|\lambda|$ sufficiently large so that

$$\left|\frac{P(\lambda) - \sigma_P(\lambda)}{|\lambda|^N}\right| \leq \frac{c_0}{2}.$$

Hence combining we get

$$|P(\lambda)| \geq \frac{1}{2}c_0|\lambda|^N \geq C\langle\lambda\rangle^N$$

for all $|\lambda|$ sufficiently large, and so we are done.

□

Theorem 4.1 (Elliptic Regularity). *If $P(D)$ is an N'th order elliptic operator and $P(D)u = f$ with $f \in H^s_{loc}(X)$, then $u \in H^{s+N}_{loc}(X)$.*

Remark: The order of the Sobolev space is essentially saying how many times differentiable the function is (see Example sheet 2). Indeed, if $\int_{\mathbb{R}^n} |\hat{u}|\langle\lambda\rangle^{2s}\, \mathrm{d}\lambda < \infty$, then we have

$$\int_{\mathbb{R}^n} |D^\alpha u|^2\, \mathrm{d}\lambda = C\int_{\mathbb{R}^n} |\lambda^\alpha|^2 \cdot |\hat{u}|\, \mathrm{d}\lambda \qquad \forall\ |\alpha| \leq s.$$

Proof. Later. □

Corollary 4.1. *Suppose $P(D)$ is elliptic, and that $P(D)u \in C^\infty(X)$. Then in fact $u \in C^\infty(X)$.*

i.e. elliptic operators are ***hypoelliptic****.*

Proof. Apply Theorem 4.1 and Corollary 3.1. □

We first prove an easier version of the elliptic regularity theorem, using a *parametrix*.

Definition 4.2. *We say that $E \in \mathscr{D}'(\mathbb{R}^n)$ is a* ***parametrix*** *for $P(D)$ if:*

$$P(D)E = \delta_0 + w$$

for some $w \in \mathscr{E}(\mathbb{R}^n)$, i.e. E inverts $P(D)$ up to some smooth function.

The intuition for this definition comes from using Green's functions to solve PDEs. Note that due to the presence of the δ_0, we don't expect a parametrix to be smooth at the origin. This is exactly what the next lemma gives for elliptic PDEs.

Lemma 4.2. *Suppose $P(D)$ is elliptic. Then it admits a parametrix E which is smooth away from the origin.*

Proof. If P is elliptic then by Lemma 4.1 we know that $\exists R > 0$ such that

$$|P(\lambda)| \gtrsim^{(x)} \langle\lambda\rangle^N \qquad \text{for } |\lambda| > R.$$

Now fix $\chi \in \mathscr{D}(\mathbb{R}^n)$ with $\chi \equiv 1$ on $|\lambda| < R$ and $\chi \equiv 0$ on $|\lambda| > R+1$. Then set:

$$\hat{E}(\lambda) := \frac{1-\chi(\lambda)}{P(\lambda)}$$

where the 'hat' notation is suggestive - we will take the inverse Fourier transform of this quantity to get the parametrix we want. Note that $\hat{E}$ is smooth and $|\hat{E}(\lambda)| \lesssim \langle\lambda\rangle^{-N}$ for $|\lambda| > R+1$.

So we have $\hat{E} \in S'(\mathbb{R}^n)$, and so by the Fourier transform,

$$P(\lambda)\hat{E} = 1 - \chi(\lambda) \qquad \Rightarrow \qquad P(D)E = \delta_0 - w$$

where $E = \mathscr{F}^{-1}(\hat{E})$, and $\hat{w} = \chi \in \mathscr{D}(\mathbb{R}^n) \subset S(\mathbb{R}^n)$. Hence $w \in S(\mathbb{R}^n)$ and so in particular it smooth. Then note (using properties of the Fourier transform):

$$\begin{aligned}\left|\big(\widehat{D^\alpha(x^\beta E)}\big)(\lambda)\right| &= \left|\lambda^\alpha D^\beta(\hat{E})\right| \\ &= \left|\lambda^\alpha D^\beta\left(\frac{1}{P(\lambda)}\right)\right| \\ &\lesssim |\lambda|^{|\alpha|-|\beta|-N}\end{aligned}$$

for $|\lambda| > R+1$ (this follows just from checking it, as $P(\lambda)$ is a polynomial of degree N). So hence we see

$$\big(\widehat{D^\alpha(x^\beta E)}\big) \in L_1(\mathbb{R}^n) \qquad \text{for } |\beta| \text{ sufficiently large.}$$

So hence by Lemma 3.1 we have that $D^\alpha(x^\beta E)$ is continuous for each α, for $|\beta|$ sufficiently large. So hence E is smooth away from $x = 0$ (since x^β is always differentiable), and so done.

□

We can now prove the easy version of Theorem 4.1 if $u, f \in \mathscr{E}'(\mathbb{R}^n)$, $P(D)u = f$.

Theorem 4.2 (Easy Version of Elliptic Regularity)**.** *Suppose $P(D)$ is an N'th order elliptic operator and we have $P(D)u = f$, with $u, f \in \mathscr{E}'(\mathbb{R}^n)$ and $f \in H^s_{loc}(\mathbb{R}^n)$. Then $u \in H^{s+N}_{loc}(\mathbb{R}^n)$.*

[x] Recall that we write $A \lesssim B$ to mean $\exists C > 0$ such that $A \leq CB$.

Proof. By Lemma 4.2, we know P admits a parametrix E which is smooth away from the origin, and so $P(D)E = \delta_0 + w$ for some w smooth. Then we have:

$$\begin{aligned} u = u * \delta_0 &= u * (P(D)E - w) \\ &= [P(D)u] * E - u * w \\ &= f * E - u * w. \end{aligned}$$

Let us look at each term on the RHS separately. For $u * w$ we have (from example sheet 2):

$$\widehat{(u * w)} = \hat{u}\hat{w}$$

and thus as $\hat{w} \in S(\mathbb{R}^n)$ and $\hat{u}$ is smooth with $|\hat{u}| \leq C\langle\lambda\rangle^m$, we have $\hat{u}\hat{w} \in S(\mathbb{R}^n)$, i.e. $\widehat{u * w} \in S(\mathbb{R}^n)$, and so (from what we know about the Fourier transform) $\Rightarrow u * w \in S(\mathbb{R}^n)$.

Hence $u * w$ is smooth and so will not effect the regularity of u above.

For $f * E$ we have

$$\left|\widehat{f * E}\right| = |\hat{E}(\lambda)| \cdot |\hat{f}(\lambda)| \lesssim \langle\lambda\rangle^{-N} |\hat{f}(\lambda)|.$$

Now since $f \in H^s(\mathbb{R}^n)$, i.e.

$$\int_{\mathbb{R}^n} |\hat{f}(\lambda)|^2 \langle\lambda\rangle^{2s} \, \mathrm{d}\lambda < \infty$$

and so with the above we see

$$\int_{\mathbb{R}^n} \left|\widehat{f * E}\right|^2 \langle\lambda\rangle^{2s+2N} \, \mathrm{d}\lambda < \infty$$

and so $f * E \in H^{s+N}(\mathbb{R}^n)$. Hence

$$u = \underbrace{E * f}_{\in H^{s+N}(\mathbb{R}^n)} - \underbrace{u * w}_{\in S(\mathbb{R}^n)} \in H^{s+N}(\mathbb{R}^n)$$

and so we are done.

$\square$

Remark: The idea for the general result is as follows. For $u \in H^s_{\text{loc}}(X)$ we have $\varphi u \in H^s(\mathbb{R}^n)$ for any $\varphi \in D(X)$. Now, looking at $P(D)u = f$, we have $\varphi \cdot P(D)u = \varphi f$, i.e.

$$P(D)[u\varphi] + (\varphi \cdot P(D)u - P(D)(u\varphi)) = \varphi f$$

and upon expanding the derivatives,

$$P(D)(u\varphi) = \varphi P(D)u + (\text{lower order terms}).$$

So all the highest order derivatives of u as contained in the first term on the LHS of the above, which is a good term we can work with. We will use this idea and a bootstrapping argument to prove Theorem 4.1.

Proof of Theorem 4.1. We have $P(D)$ an N'th order elliptic operator (with constant coefficients), and $P(D)u = f \in H^s_{\text{loc}}(X)$. Recall the following elementary facts about Sobolev spaces (many of these are exercises on Example Sheet 2):

(i) If $u \in \mathscr{E}'(\mathbb{R}^n)$, then $u \in H^t(\mathbb{R}^n)$ for some $t \in \mathbb{R}$.

(ii) If $u \in H^t(\mathbb{R}^n)$, then $D^\alpha u \in H^{t-|\alpha|}(\mathbb{R}^n)$.

(iii) If $s > t$ then $H^s(\mathbb{R}^n) \subset H^t(\mathbb{R}^n)$.

(iv) If $\varphi \in S(\mathbb{R}^n)$ and $u \in H^t(\mathbb{R}^n)$, then $\varphi u \in H^t(\mathbb{R}^n)$.

To show $u \in H^{s+N}_{\text{loc}}(X)$, it suffices to show that $\varphi u \in H^{s+N}(\mathbb{R}^n)$ for $\varphi \in \mathscr{D}(X)$ arbitrary. So fix $\varphi \in \mathscr{D}(X)$.

Now introduce $\{\psi_0, \dots, \psi_M\} \subset S(\mathbb{R}^n)$, where

$$\text{supp}(\varphi) \subset \text{supp}(\psi_M) \subset \cdots \subset \text{supp}(\psi_0)$$

and $\psi_{i-1}|_{\text{supp}(\psi_i)} \equiv 1$. Note that $\psi_0 u \in \mathscr{E}'(\mathbb{R}^n)$. Hence by property (i) above, $\exists t \in \mathbb{R}$ such that $\psi_0 u \in H^t(\mathbb{R}^n)$. Now note that

$$\begin{aligned} P(D)(\psi_1 u) &= \psi_1 P(D)u + [P(D)\psi_1 u - \psi_1 P(D)u] \\ &= \psi_1 f + [P(D), \psi_1](u) \\ &= \psi_1 f + [P(D), \psi_1](\psi_0 u) \end{aligned}$$

where $[g,h] = gh - hg$ is the commutator, and in the last equality we have used the fact that $\psi_0 \equiv 1$ on $\text{supp}(\psi_1)$ (and so this doesn't change anything).

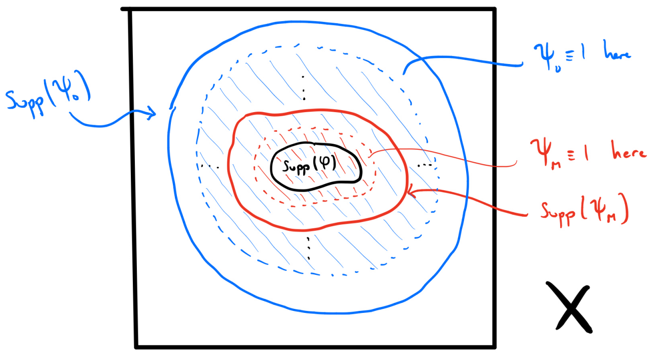

FIGURE 2. An illustration of the ψ_i maps in relation to one another and φ.

Note that (by the Leibniz rule) $[P(D), \psi_1]$ is a differential operator of order $N-1$, with coefficients in $S(\mathbb{R}^n)$ (since $\psi_1 \in S(\mathbb{R}^n)$). Hence as $\psi_0 u \in H^t(\mathbb{R}^n)$ (by property (iv) above) we have $[P(D), \psi_1](\psi_0 u) \in H^{t-(N-1)}(\mathbb{R}^n)$, from property (ii) and (iii). So:

$$P(D)(\psi_1 u) = \underbrace{\psi_1 f}_{\in H^s(\mathbb{R}^n) \text{ as } \psi_1 \text{ is in } S(\mathbb{R}^n)} + \overbrace{[P(D), \psi_1](\psi_0 u)}^{\in H^{t-(N-1)}(\mathbb{R}^n)} \implies P(D)(\psi_1 u) \in H^{\tilde{A}_1}(\mathbb{R}^n),$$

where $\tilde{A}_1 = \min\{s, t - N + 1\}$. This means that (by definition of Sobolev spaces),

$$\int_{\mathbb{R}^n} \left| P(\lambda)\widehat{(\psi_1 u)}(\lambda) \right|^2 \langle \lambda \rangle^{2\tilde{A}_1} \, d\lambda < \infty$$

and hence (since P is of N'th order/Lemma 4.1),

$$\begin{aligned}\int_{\mathbb{R}^n} \left|\widehat{(\psi_1 u)}(\lambda)\right|^2 \langle\lambda\rangle^{2\tilde{A}_1+2N}\, \mathrm{d}\lambda &\lesssim \int_{\mathbb{R}^n} \left|\frac{P(\lambda)}{\langle\lambda\rangle^N}\cdot\widehat{(\psi_1 u)}(\lambda)\right|^2 \cdot \langle\lambda\rangle^{2\tilde{A}_1+2N}\, \mathrm{d}\lambda\\ &= \int_{\mathbb{R}^n} \left|P(\lambda)\widehat{(\psi_1 u)}(\lambda)\right|^2 \langle\lambda\rangle^{2\tilde{A}_1}\, \mathrm{d}\lambda\\ &< \infty\end{aligned}$$

and thus $\psi_1 u \in H^{A_1}(\mathbb{R}^n)$, where $A_1 = \tilde{A}_1 + N = \min\{s+N, t+1\}$.

Now repeating this, we have

$$\begin{aligned}P(D)(\psi_2 u) &= \psi_2 P(D)u + [P(D),\psi_2](u)\\ &= \psi_2 f + [P(D),\psi_2](\psi_1 u)^{\text{(xi)}}\end{aligned}$$

since $\psi_1 = 1$ on $\mathrm{supp}(\psi_2)$. Then continuing as before (just with t replaced by A_1) we find that $\psi_2 u \in H^{A_2}(\mathbb{R}^n)$, where:

$$\begin{aligned}A_2 = \min\{s+N, A_1+1\} &= \min\{s+N, \min\{s+N+1, t+2\}\}\\ &= \min\{s+N, s+N+1, t+2\}\\ &= \min\{s+N, t+2\}.\end{aligned}$$

Thus continuing inductively, we see that $\psi_M u \in H^{A_M}(\mathbb{R}^n)$, where

$$A_M = \min\{s+N, t+M\}.$$

So choosing $M > s+N-t$, so that $A_M = s+N$, then since $\psi_M = 1$ on $\mathrm{supp}(\varphi)$ we get that $\varphi u \in H^{s+N}(\mathbb{R}^n)$.

But then since $\varphi \in \mathscr{D}(X)$ was arbitrary, this gives that $u \in H^{s+N}_{\mathrm{loc}}(X)$ and so we are done.

□

4.2. **Fundamental Solutions.**

Again we study $P(D)u = f$, with $u, f \in \mathscr{D}'(\mathbb{R}^n)$.

Definition 4.3. *We say $E \in \mathscr{D}'(\mathbb{R}^n)$ is a **fundamental solution** for $P(D)$ if*

$$P(D)E = \delta_0.$$

Note: At least formally, if $P(D)E = \delta_0$ and we want to solve $P(D)u = f$, then we have $u = E * f$, since

$$P(D)u = (P(D)E*)f = \delta_0 * f = f.$$

However we have only defined convolutions if one of E, f has compact support, and so this doesn't quite work. But it is the key idea.

(xi) We can do things like Fourier transforms when we have a test function multiplied by u. This is why we are so keen to get ψ's in here.

Example 4.1. *Let $z = x_1 + ix_2 \in \mathbb{C} \cong \mathbb{R}^2$. Then define the* ***Cauchy-Riemann operator*** *by*

$$\frac{\partial}{\partial \bar{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x_1} + i\frac{\partial}{\partial y}\right).$$

We claim that $E = \frac{1}{\pi z}$ is the fundamental solution for $\frac{\partial}{\partial \bar{z}}$.

Proof. Fix $\varphi \in \mathscr{D}(\mathbb{R}^2)$. Note that on $|z| \geq \varepsilon$ (for any $\varepsilon > 0$), we have

$$\frac{\partial}{\partial \bar{z}}E = 0.$$

Also, E is locally integrable and so $E \in \mathscr{D}'(\mathbb{R}^n)$. Now we have:

$$\begin{aligned}\left\langle \frac{\partial}{\partial \bar{z}}E, \varphi\right\rangle &= -\int_{\mathbb{R}^2} E \cdot \frac{\partial \varphi}{\partial \bar{z}}\ \mathrm{d}x_1 \mathrm{d}x_2\\ &= -\lim_{\varepsilon \downarrow 0}\int_{|z|>\varepsilon} E \cdot \frac{\partial \varphi}{\partial \bar{z}}\ \mathrm{d}x_1 \mathrm{d}x_2\\ &= -\lim_{\varepsilon \downarrow 0}\int_{|z|>\varepsilon} \frac{\partial}{\partial \bar{z}}(\varphi E)\ \mathrm{d}x_1 \mathrm{d}x_2 \qquad \text{since } \frac{\partial}{\partial \bar{z}}E = 0 \text{ here.}\end{aligned}$$

Now note that (by definition of $\partial/\partial\bar{z}$):

$$\frac{\partial}{\partial \bar{z}}(\varphi E) = \frac{1}{2}\left[\frac{\partial}{\partial x_1}(\varphi E) + i\frac{\partial}{\partial x_2}(\varphi E)\right]$$

and so by Green's theorem[xii](setting $Q = \varphi E/2$, $P = -i\varphi E/2$) we get

$$\begin{aligned}\left\langle \frac{\partial}{\partial \bar{z}}E, \varphi\right\rangle &= \lim_{\varepsilon \downarrow 0}\frac{1}{2}\int_{|z|=\varepsilon} -i\varphi E \mathrm{d}x_1 + \varphi E \mathrm{d}x_2\\ &= \lim_{\varepsilon \downarrow 0}\int_{|z|=\varepsilon} \frac{\varphi E}{2i}\,[\mathrm{d}x_1 + i\mathrm{d}x_2]\\ &= \lim_{\varepsilon \downarrow 0}\int_{|z|=\varepsilon} \frac{\varphi E}{2i}\ \mathrm{d}z \qquad \text{since } \mathrm{d}z = \mathrm{d}x_1 + i\mathrm{d}x_2\\ &= \lim_{\varepsilon \downarrow 0}\int_0^{2\pi} \frac{\varphi(\varepsilon\cos(\theta), \varepsilon\sin(\theta))}{2i}\cdot\frac{1}{\pi\varepsilon e^{i\theta}}\cdot i\varepsilon e^{i\theta}\ \mathrm{d}\theta \qquad \text{setting } z = \varepsilon e^{i\theta}\ (\text{as } E = 1/\pi z)\\ &= \lim_{\varepsilon \downarrow 0}\frac{1}{2\pi}\int_0^{2\pi} \varphi(\varepsilon\cos(\theta), \varepsilon\sin(\theta))\ \mathrm{d}\theta\\ &= \varphi(0)\\ &= \langle \delta_0, \varphi\rangle.\end{aligned}$$

Hence we see that, as distributions, $\frac{\partial E}{\partial \bar{z}} = \delta_0$ and so we are done.

□

[xii]Recall that Green's theorem says that for suitable P, Q, A:

$$\oint_{\partial A} P\mathrm{d}x + Q\mathrm{d}y = \int\int_A \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ \mathrm{d}x\mathrm{d}y.$$

Example 4.2 (Fundamental Solution of the Heat Equation). *If $L = \frac{\partial}{\partial t} - \Delta_x$ is the heat operator on $\mathbb{R}^n \times \mathbb{R}$, then a fundamental solution is:*

$$E(x,t) = \begin{cases} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x|^2}{4t}} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0. \end{cases}$$

Proof. Note on $t \geq \varepsilon$ (for any $\varepsilon > 0$), we have that

$$\left(\frac{\partial}{\partial t} - \Delta_x\right) E = 0,$$

and that E is locally integrable. So $E \in \mathscr{D}'(\mathbb{R}^n \times \mathbb{R})$.

Then by definition, for any $\varphi \in \mathscr{D}(\mathbb{R}^n \times \mathbb{R})$ we have:

$$\begin{aligned} \langle LE, \varphi\rangle &= \int_0^\infty \mathrm{d}t \int_{\mathbb{R}^n} \mathrm{d}x\ E\left(-\frac{\partial \varphi}{\partial t} - \Delta_x \varphi\right) \\ &= \lim_{\varepsilon \downarrow 0} \int_\varepsilon^\infty \mathrm{d}t \int_{\mathbb{R}^n} \mathrm{d}x\ E\left(-\frac{\partial \varphi}{\partial t} - \Delta_x \varphi\right) \\ &= \lim_{\varepsilon \downarrow 0} \int_\varepsilon^\infty \mathrm{d}t \int_{\mathbb{R}^n} \mathrm{d}x\ -\frac{\partial}{\partial t}(\varphi E) + \lim_{\varepsilon \downarrow 0} \int_\varepsilon^\infty \mathrm{d}t \int_{\mathbb{R}^n} \mathrm{d}x\ \varphi \underbrace{\left(\frac{\partial E}{\partial t} - \Delta_x E\right)}_{=0} \\ &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^n} \mathrm{d}x\ \varphi(x,\varepsilon) E(x,\varepsilon) \\ &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^n} \mathrm{d}x\ \varphi(x,\varepsilon) \cdot \frac{1}{(4\pi\varepsilon)^{n/2}} e^{-\frac{|x|^2}{4\varepsilon}} \qquad \text{and setting } y = \frac{x}{2\sqrt{\varepsilon}}, \\ &= \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}^n} \mathrm{d}y\ \varphi(2\sqrt{\varepsilon} y, \varepsilon) \cdot (4\pi\varepsilon)^{-n/2} e^{-|y|^2} \cdot \varepsilon^{n/2} 2^n \\ &= \lim_{\varepsilon \downarrow 0} \pi^{-n/2} \int_{\mathbb{R}^n} \mathrm{d}y\ \varphi(2\sqrt{\varepsilon} y, \varepsilon) e^{-|y|^2} \\ &= \pi^{-n/2} \varphi(0,0) \int_{\mathbb{R}^n} e^{-|y|^2}\ \mathrm{d}y \\ &= \varphi(0,0) \\ &= \langle \delta_0, \varphi\rangle \end{aligned}$$

(where in the first equality, we get a minus sign on $-\partial\varphi/\partial t$ as this comes from integrating by parts once, whilst the other term we integrate by parts twice and so the sign does not change). Thus we see that, as distributions, $LE = \delta_0$ and so E is the fundamental solution.

□

Remark: So how do we find fundamental solutions in general? Our intuition/first guess might be to consider:

$$E\varphi \longmapsto \langle E, \varphi\rangle := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \frac{\hat{\varphi}(-\lambda)}{P(\lambda)}\,\mathrm{d}\lambda.$$

Then we would (heuristically) have:

$$\langle P(D)E, \varphi\rangle = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \frac{[\widehat{P(-D)\varphi}](-\lambda)}{P(\lambda)}\,\mathrm{d}\lambda = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \frac{P(\lambda)\hat{\varphi}(-\lambda)}{P(\lambda)}\,\mathrm{d}\lambda = \varphi(0) = \langle \delta_0, \varphi\rangle$$

where we have used the Fourier inversion formula. So this would appear to work, but this E may not be well-defined as we could have $P(\lambda) = 0$. We will get around this by constructing **Hörmander's staircase**. The idea is to "wobble" the surface of integration.

Lemma 4.3. *For $x \in \mathbb{R}^n$, write $x = (x', x_n)$, with $x' \in \mathbb{R}^{n-1}$. Then for $\varphi \in \mathscr{D}(\mathbb{R}^n)$, the function $z \mapsto \hat{\varphi}(\lambda', z)$ is complex analytic in $z \in \mathbb{C}$, for each $\lambda' \in \mathbb{R}^{n-1}$. Also,*

$$\left|\hat{\varphi}(\lambda', z)\right| \lesssim_m (1+|z|)^{-m} e^{\delta|\mathrm{Im}(z)|} \qquad \text{for all } m = 0, 1, 2, \ldots$$

for some $\delta > 0$ (i.e. $\hat{\varphi}$ decays rapidly in each horizontal z-direction).

Remark: Here $\lesssim_m$ means $\lesssim$, except the constant can depend on m.

Proof. By definition of the Fourier transform and Fubini's theorem we have:

$$\hat{\varphi}(\lambda', z) = \int_{\mathbb{R}^{n-1}} e^{-i\lambda'\cdot x'}\left(\int_{\mathbb{R}} e^{-izx_n}\varphi(x', x_n)\,\mathrm{d}x_n\right)\mathrm{d}x'.$$

Thus we see that $z \mapsto \hat{\varphi}(\lambda', z)$ is analytic from checking the Cauchy-Riemann equations (we can differentiate under the integral sign by dominated convergence)(xiii). Also, on integrating by parts:

$$\begin{aligned}
|z^m\hat{\varphi}(\lambda', z)| &= \left|\int_{\mathbb{R}^{n-1}} e^{-i\lambda'\cdot x'}\left(\int_{\mathbb{R}}\left[\left(\frac{\partial}{\partial x_n}\right)^m e^{-izx_n}\right]\varphi(x', x_n)\,\mathrm{d}x_n\right)\mathrm{d}x'\right| \\
&= \left|\int_{\mathbb{R}^{n-1}} e^{-i\lambda'\cdot x'}\left(\int_{\mathbb{R}} e^{-izx_n}\cdot\frac{\partial^m\varphi}{\partial x_n^m}(x', x_n)\,\mathrm{d}x_n\right)\mathrm{d}x'\right| \\
&\leq \int_{\mathbb{R}^n} e^{\mathrm{Im}(z)x_n}\left|\frac{\partial^m\varphi}{\partial x_n^m}\right|\mathrm{d}x \\
&\lesssim_m e^{\delta|\mathrm{Im}(z)|}
\end{aligned}$$

where we have used that $\varphi \equiv 0$ for $|x_n| > \delta > 0$, for some $\delta > 0$, so we only need to consider $|x_n| \leq \delta$ in the integral, so we can bound this by the integral of $\left|\frac{\partial^m\varphi}{\partial x_n^m}\right|$ and the maximum the exponential term is, which is this.

□

Note: The above lemma tells us (along with Cauchy's theorem from complex analysis) that

$$\int_{\mathrm{Im}(z)=c} \hat{\varphi}(\lambda', z)\,\mathrm{d}z$$

(xiii) Alternatively we can use Morera's theorem and Fubini's theorem to check this.

is independent of $c \in \mathbb{R}$.

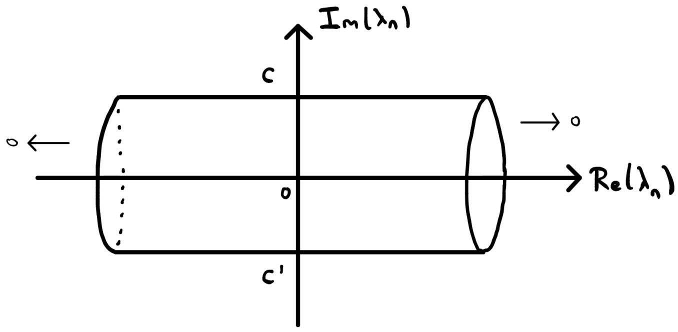

FIGURE 3. Lemma 4.3 tells us the integrals along the vertical sides tend to 0 as they get pushed out. Thus the integral along the horizontal sides is independent of c, by Cauchy's theorem.

Theorem 4.3 (Malgrange-Ehrenpries). *Every non-zero, constant coefficient partial differential operator $P(D)$ admits a fundamental solution.*

Proof. By a rotation of our coordinate axes, we can assume wlog that $P(D)$ is of the form:

$$P(\lambda', \lambda_n) = \lambda_n^M + \sum_{i=0}^{M-1} a_i(\lambda')\lambda_n^i$$

i.e. treat P as a polynomial in λ_n with coefficients depending on λ'. Wlog by scaling the leading coefficient is 1.

Then for fixed $\mu' \in \mathbb{R}^{n-1}$ we can write:

$$P(\mu', \lambda_n) = \prod_{i=1}^{M}(\lambda_n - \tau_i(\mu'))$$

where $\{\tau_i(\mu')\}_{i=1}^M$ are the roots of the M'th order polynomial $P(\mu', \cdot)$.

Now $\exists$ a horizontal line $\mathrm{Im}(\lambda_n) = c(\mu')$ in the complex λ_n—plane, with $|\mathrm{Im}(\lambda_n)| \leq (M+1)(1+\varepsilon)$, for $\varepsilon > 0$ small, such that

$$|\mathrm{Im}(\lambda_n) - \tau_i(\mu')| > 1 \qquad \text{for all } i = 1, \ldots, M.$$

This is seen by considering $2M+2$ strips of width $1+\varepsilon$. Then by the pigeonhole principle, $\exists$ two adjacent strips that don't contain any roots $\tau_i(\mu')$. Then take "$\mathrm{Im}(\lambda_n) = c$", for c the line between these two strips (then clearly $c = c(\mu')$). [See Figure 4.]

So on $\mathrm{Im}(\lambda_n) = c(\mu')$ we have:

$$|P(\mu', \lambda_n)| = \prod_{i=1}^{M} |\mathrm{Im}(\lambda_n) - \tau_i(\mu')| > 1$$

since each term in the product is > 1. Then by continuity (as the roots of a polynomial vary continuously with the polynomials coefficients, which is seen from the formula expression the coefficients

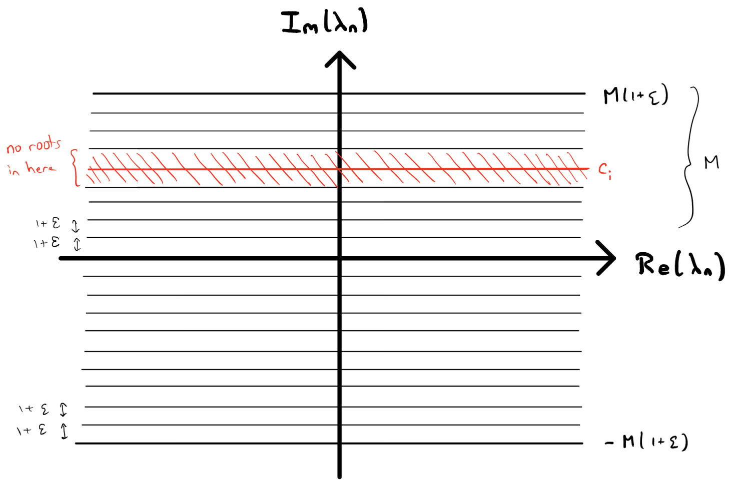

FIGURE 4. An illustration of how the constants c_i are constructed using the pigeon-hole principle. There are M roots but $2M+2$ strips, so two adjacent strips do not contain any roots.

in terms of the roots) the $\tau_i(\mu')$ depend continuously on μ', and thus the same estimate holds on a neighbourhood of μ', $N_{\mu'}$, say.

Then since this argument holds for any $\mu' \in \mathbb{R}^{n-1}$, we can cover $\mathbb{R}^{n-1}$ with such neighbourhoods $\{N_{\mu'}\}_{\mu'\in\mathbb{R}^{n-1}}$. Then by Heine-Borel, we can extract a locally finite subcover, $N_i = N_{\mu_i'}$, so that $|P(\lambda',\lambda_n)| > 1$ for $\lambda' \in N_i$, $\mathrm{Im}(\lambda_n) = c(\mu_i') =: c_i$.

[i.e. get a finite subcover on compact balls of increasing radii - this gives a countable subcover which is finite on every compact subset of $\mathbb{R}^{n-1}$.]

So we have that we can find a horizontal line $\mathrm{Im}(\lambda_n) = c_i$ on each N_i to avoid the roots of P. However we need to be careful on overlaps. Thus therefore define the disjoint sets

$$\Delta_i := N_i \backslash \bigcup_{j=1}^{i-1} N_j$$

for which we see $|P(\lambda',\lambda_n)| > 1$ for $\lambda' \in \Delta_i$ and $\mathrm{Im}(\lambda_n) = c_i$. Moreover we have $\bigcup_{i=1}^{\infty} \overline{\Delta_i} = \mathbb{R}^{n-1}$.

Now define[xiv]:

$$\langle E, \varphi\rangle := \frac{1}{(2\pi)^n}\sum_{i=1}^{\infty}\int_{\Delta_i}\left[\int_{\mathrm{Im}(\lambda_n)=c_i}\frac{\hat{\varphi}(-\lambda',-\lambda_n)}{P(\lambda',\lambda_n)}\,\mathrm{d}\lambda_n\right]\mathrm{d}\lambda'.$$

[xiv]Heuristically we wanted to define

$$\langle E, \varphi\rangle := \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\frac{\hat{\varphi}(-\lambda)}{P(\lambda)}\,\mathrm{d}\lambda = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^{n-1}}\left[\int_{\mathrm{Im}(\lambda_n)=0}\frac{\hat{\varphi}(-\lambda',-\lambda_n)}{P(\lambda)}\,\mathrm{d}\lambda_n\right]\mathrm{d}\lambda'$$

since $\{\mathrm{Im}(\lambda_n)=0\} \equiv \mathbb{R}$. Then here we know from Lemma 4.3 that we can change $\int_{\mathrm{Im}(\lambda_n)=0}$ to $\int_{\mathrm{Im}(\lambda_n)=c_i}$, and we can change the integral $\int_{\mathbb{R}^{n-1}}\mathrm{d}\lambda'$ into $\sum_i\int_{\Delta_i}\mathrm{d}\lambda'$ as the sets Δ_i are disjoint and cover (up to a set of measure zero). So we get the formula given in the proof, which is good because it avoids the roots of P.

By construction we have

$$|P(\lambda', \lambda_n)| > 1 \quad \text{for} \quad (\lambda', \lambda_n) \in \Delta_i \times \{\lambda : \mathrm{Im}(\lambda_n) = c_i\}$$

and it follows [from Example Sheet 3] that $E \in \mathscr{D}'(\mathbb{R}^n)$. Also, we have

$$\begin{aligned}
\langle P(D)E, \varphi \rangle &= \frac{1}{(2\pi)^n} \sum_{i=1}^{\infty} \int_{\Delta_i} \left[\int_{\mathrm{Im}(\lambda_n)=c_i} \frac{\cancel{P(\lambda)}\hat{\varphi}(-\lambda', -\lambda_n)}{\cancel{P(\lambda)}} \, \mathrm{d}\lambda_n \right] \mathrm{d}\lambda' \\
&= \frac{1}{(2\pi)^n} \sum_{i=1}^{\infty} \int_{\Delta_i} \left[\int_{\mathrm{Im}(\lambda_n)=0} \hat{\varphi}(-\lambda', -\lambda_n) \, \mathrm{d}\lambda_n \right] \mathrm{d}\lambda' \qquad \text{by Cauchy's theorem and Lemma 4.3} \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \hat{\varphi}(-\lambda', -\lambda_n) \, \mathrm{d}\lambda_n \mathrm{d}\lambda' \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\varphi}(-\lambda) \, \mathrm{d}\lambda \\
&= \varphi(0) \qquad \text{by Fourier inversion} \\
&= \langle \delta_0, \varphi \rangle
\end{aligned}$$

and so as $\varphi \in \mathscr{D}(\mathbb{R}^n)$ was arbitrary in the above, we get $P(D)E = \delta_0$ and so we are done.

□

Remark: As said before, this method of proof is called **Hörmander's staircase**, since the integration along $\mathbb{R}^{n-1}$ looks like a staircase instead of a plane:

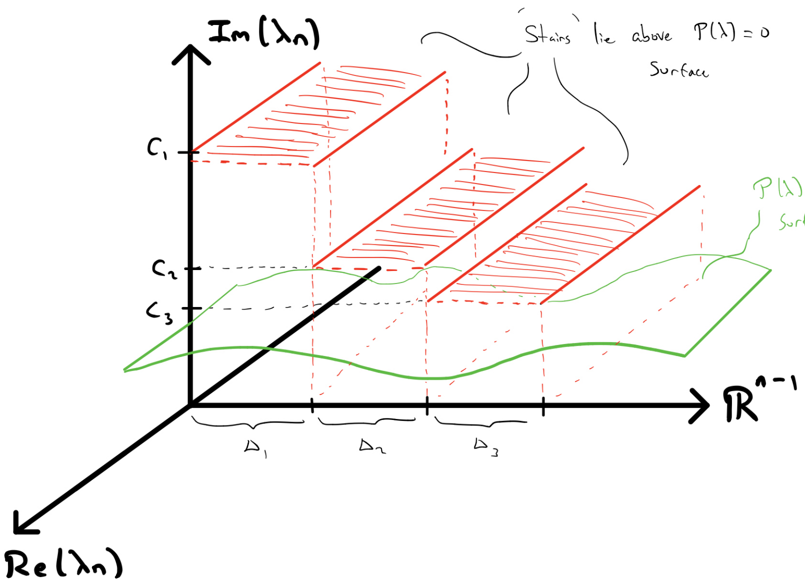

FIGURE 5. An illustration of Hörmander's staircase. We essentially split up the plane $\{\mathrm{Im}(\lambda_n) = 0\} \cong \mathbb{R}$ into disjoint regions and translate them in such a way to avoid the $P(\lambda) = 0$ surface.

This result establishes the existence of a fundamental solution in $\mathscr{D}'(\mathbb{R}^n)$. Using more sophistication, one can show that $\exists$ a fundamental solution in $S'(\mathbb{R}^n)$. [Wagner gave an explicit construction, with a nice formula as above, using complex analysis.]

4.3. **The Structure Theorem for $\mathscr{E}'(X)$.**

We know that if $f \in C(X)$, then we can always define $\partial^\alpha f \in \mathscr{D}'(X)$ (i.e. as a distribution, not a function) via:

$$\langle \partial^\alpha f, \varphi \rangle := (-1)^{|\alpha|} \int_X f \, \partial^\alpha \varphi \, \mathrm{d}x$$

for $\varphi \in \mathscr{D}(X)$. A natural question you might ask then is whether all distributions can be written as a linear combination of such distributions (a bit like a Taylor series), i.e. whether we can find some collection $(f_\alpha)_\alpha \subset \mathbb{C}(X)$ with

$$u = \sum_\alpha \partial^\alpha f_\alpha \qquad \text{in } \mathscr{D}'(X)?$$

We will prove that this is true for $u \in \mathscr{E}(X)$.

So if $u \in \mathscr{E}'(\mathbb{R}^n) \subset S'(\mathbb{R}^n)$, then define the Fourier transform $\hat{u} \in S'(\mathbb{R}^n)$ in the usual way:

$$\langle \hat{u}, \varphi \rangle := \langle u, \varphi \rangle \qquad \forall \varphi \in S(\mathbb{R}^n).$$

Then by definition of the Fourier transform we have:

$$\langle \hat{u}, \varphi \rangle = \langle u(x), \hat{\varphi}(x) \rangle = \left\langle u(x), \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} \varphi(\lambda) \, \mathrm{d}\lambda \right\rangle.$$

Now fix[xv] $\chi \in \mathscr{D}(\mathbb{R}^n)$ with $\chi \equiv 1$ on $|\lambda| < 1$ and $\chi \equiv 0$ on $|\lambda| > 2$. Then define $\varphi_m \in \mathscr{D}(\mathbb{R}^n)$ by:

$$\varphi_m(\lambda) := \chi\left(\frac{\lambda}{m}\right)\varphi(\lambda).$$

Then we can show $\varphi_m \to \varphi$ in $S(\mathbb{R}^n)$ and hence $\hat{\varphi}_m \to \hat{\varphi}$ in $S(\mathbb{R}^n)$ by continuity of the Fourier transform. So,

$$\langle \hat{u}, \varphi \rangle = \lim_{m\to\infty} \langle u, \hat{\varphi}_m \rangle = \lim_{m\to\infty} \left\langle u(x), \int_{\mathbb{R}^m} e^{-i\lambda\cdot x} \varphi_m(\lambda) \, \mathrm{d}\lambda \right\rangle.$$

Now apply the same Riemann sum argument (from Lemma 1.5) to interchange $\langle \cdot, \cdot \rangle$ with $\int_{\mathbb{R}^n} \langle \cdot, \cdot \rangle \, \mathrm{d}\lambda$, and so hence

$$= \lim_{m\to\infty} \int_{\mathbb{R}^n} \langle u(x), e^{-i\lambda\cdot x} \rangle \varphi_m(\lambda) \, \mathrm{d}\lambda = \int_{\mathbb{R}^n} \langle u(x), e^{-i\lambda\cdot x} \rangle \varphi(\lambda) \, \mathrm{d}\lambda$$

where this last equality comes from the dominated convergence theorem and a semi-norm estimate (see below).

So we can identify $\hat{u} \in S'(\mathbb{R}^n)$ with the function:

$$\lambda \longmapsto \hat{u}(\lambda) := \langle u(x), e^{-i\lambda\cdot x} \rangle$$

for $\lambda \in \mathbb{R}^n$. We can then in fact show that $\hat{u} \in C^\infty(\mathbb{R}^n)$, and by the semi-norm definition of $\mathscr{E}'(\mathbb{R}^n)$, $\exists$ constants $C, N \geq 0$ and compact $K \subset \mathbb{R}^n$ such that:

$$\begin{aligned} |\hat{u}(\lambda)| &= \left| \langle u(x), e^{-i\lambda\cdot x} \rangle \right| \\ &\leq C \sum_{|\alpha|\leq N} \sup_K \left| \partial_x^\alpha (e^{-i\lambda\cdot x}) \right| \\ &\lesssim \langle \lambda \rangle^N \end{aligned}$$

and so hence we can apply the dominated convergence theorem in the above.

[xv] The point is we want to do the same Riemann integral/sum argument we did before in Lemma 1.5. However the sums aren't finite here as we do not have compact support. So we need to modify by such a function to get this.

In summary, we see that if $u \in \mathscr{E}'(\mathbb{R}^n) \subset S(\mathbb{R}^n)$, then its Fourier transform $\hat{u} \in S'(\mathbb{R}^n)$ can be identified with the function $\hat{u} := \langle u, f_\lambda \rangle$, where $f_\lambda(x) := e^{-i\lambda\cdot x}$. Moreover we have $\hat{u} \in C^\infty(\mathbb{R}^n)$ and $|\hat{u}(\lambda)| \lesssim \langle \lambda \rangle^N$ for some $N \geq 0$.(xvi)

This is just all preamble to the following theorem:

Theorem 4.4. *For $u \in \mathscr{E}'(X)$, $\exists$ a $\underline{\text{finite}}$ collection $(f_\alpha)_\alpha \subset C(X)$ such that*

$$u = \sum_\alpha \partial^\alpha f_\alpha \qquad \textit{in } \mathscr{E}'(X)$$

and $\text{supp}(f_\alpha) \subset X$ for all α.

Remark: We can use this to prove results: if we can just prove a result for distributions of the form $\partial^\alpha f$ for $f \in C(X)$ then as the above sum is finite we can extend the result to all $u \in \mathscr{E}(X)$.

Proof. Fix $\rho \in \mathscr{D}(X)$ have $\rho \equiv 1$ on supp(u). Then for $\varphi \in \mathscr{E}(X)$ we have

$$\langle u, \varphi \rangle = \langle u, \rho\varphi \rangle + \underbrace{\langle u, (1-\rho)\varphi \rangle}_{=0 \text{ as } 1-\rho\equiv 0 \text{ on supp}(u)}$$

i.e.

$$\langle u, \varphi \rangle = \langle u, \rho\varphi \rangle.$$

Then since $\rho\varphi \in \mathscr{D}(X)$ we can treat both u and $\rho\varphi$ as elements of $\mathscr{E}'(\mathbb{R}^n)$ and $\mathscr{D}(\mathbb{R}^n)$ respectively, by extension by zero.

Then since $\rho\varphi \in \mathscr{D}(\mathbb{R}^n) \subset S(\mathbb{R}^n)$, we can write:

$$\rho\varphi = \hat{\hat{\psi}} \qquad \text{for some } \psi \in S(\mathbb{R}^n)$$

(i.e. double Fourier transform, so $\rho\varphi = (2\pi)^n \check{\psi}$). Hence:

$$\langle u, \varphi \rangle = \langle u, \hat{\hat{\psi}} \rangle = \langle \hat{u}, \hat{\psi} \rangle.$$

Using the Laplacian $\Delta := \sum_{i=1}^n (\partial/\partial x_i)^2$, we can write (as the Fourier transform changes $-\Delta$ into $\lambda_1 + \cdots + \lambda_n^2 = |\lambda|^2$):

$$\hat{\psi} = \langle \lambda \rangle^{-2M} [\widehat{(1-\Delta)^M \psi}]$$

and so:

$$\langle u, \varphi \rangle = \left\langle \langle \lambda \rangle^{-2M} \hat{u}, [\widehat{(1-\Delta)^M \psi}] \right\rangle.$$

Now since $|\hat{u}| \lesssim \langle \lambda \rangle^N$ for some $N \geq 0$ (from the preamble to this theorem), we can choose M such that $\langle \lambda \rangle^{-2M} \hat{u} \in L^1(\mathbb{R}^n)$. Then define:

$$\check{f}(x) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\lambda\cdot x} \langle \lambda \rangle^{-2M} \hat{u}(\lambda)\, \mathrm{d}\lambda.$$

So $f \in C(\mathbb{R}^n)$ by the dominated convergence theorem (i.e. Lemma 3.1). Hence since

$$\check{f} = \frac{1}{(2\pi)^n} \left[\widehat{\langle \lambda \rangle^{-2M} \ \hat{u}} \ \right]$$

(xvi)This demonstrates a general rule: taking the Fourier transform of a distribution of compact support gives a very nice function. So to solve a problem we can take the Fourier transform, solve the problem with this nice function, and then take the inverse Fourier transform to solve the original problem.

we see

$$\begin{aligned}\langle u, \varphi\rangle &= \left\langle \left[\, \widehat{\langle\lambda\rangle^{-2M}\ \hat{u}}\ \right], (1-\Delta)^M\psi \right\rangle \\ &= \left\langle (2\pi)^n \check{f}, (1-\Delta)^M\psi \right\rangle \\ &= \left\langle \check{f}, (1-\Delta)^M(\rho\check{\varphi}) \right\rangle \\ &= \left\langle f, (1-\Delta)^M(\rho\varphi) \right\rangle\end{aligned}$$

which is good since we know f is continuous. By the Leibniz rule we have

$$(1-\Delta)^M(\rho\varphi) = \sum_\alpha (-1)^{|\alpha|}\rho_\alpha \partial^\alpha \varphi$$

where $\rho_\alpha \in \mathscr{D}(X)$ are defined accordingly (just by the Leibniz rule). Thus we see:

$$\begin{aligned}\langle u, \varphi\rangle &= \left\langle f\ ,\ \sum_\alpha (-1)^{|\alpha|}\rho_\alpha\partial^\alpha\varphi \right\rangle \\ &= \left\langle \sum_\alpha (-1)^{|\alpha|}\partial^\alpha(\rho_\alpha f)\ ,\ \varphi \right\rangle.\end{aligned}$$

So setting $f_\alpha = \rho_\alpha f$, we see $f_\alpha \in C(X)$ and $\operatorname{supp}(f_\alpha) \subset X$, and the above gives

$$u = \sum_\alpha (-1)^{|\alpha|}\partial^\alpha f_\alpha \qquad \text{in } \mathscr{E}'(X)$$

and so we are done.

$\square$

Remark: Note that this proof is constructive: we can actually write down these f_α from u.

Aside: For $u \in \mathscr{D}'(X)$, a similar approach can be taken, and we can show that:

$$u = \sum_\alpha \partial^\alpha f_\alpha \qquad \text{in } \mathscr{D}'(X)$$

but the sum can be infinite, although the sum is locally finite, i.e.

$$\langle u, \varphi\rangle = \sum_\alpha \langle \partial^\alpha f_\alpha, \varphi\rangle \qquad \text{for each } \varphi \in \mathscr{D}(X)$$

with only finitely many terms on the RHS being non-zero (so a finite sum for each φ, although which f_α are included can change). To prove this we just need to invoke a partition of unity of X (i.e. for $(\rho_\beta)_\beta$ a partition of unity on X, $u = \sum_\beta \rho_\beta u$ is a sum of distributions of compact support on which the above can be applied), although we shan't go into the details here.

[Look at Reed and Simon's book, in particular the chapter on tempered distributions and the quantum harmonic oscillator.]

4.4. **The Paley-Wiener-Schwartz Theorem.**

For $u \in \mathscr{E}'(\mathbb{R}^n)$ we know we can identify the distribution $\hat{u}$ with the smooth function

$$\lambda \longmapsto \hat{u}(\lambda) := \langle u(x), e^{-i\lambda\cdot x}\rangle \quad \text{for } \lambda \in \mathbb{R}^n.$$

So what happens if we let $\lambda \in \mathbb{C}^n$? Then we can take the complex extension ($\lambda \to z \in \mathbb{C}$), i.e. define

$$\hat{u}(z) := \langle u(x), e^{-iz\cdot x}\rangle$$

which is called the **Fourier-Laplace transform** of u (i.e. the complex extension of the function identifying the Fourier transform of u).

We can see that in fact the Fourier-Laplace transform is complex analytic, since it satisfies the Cauchy-Riemann conditions $\frac{\partial}{\partial \bar{z}}\hat{u} = 0$, since:

$$\frac{\partial}{\partial \bar{z}_i}\langle u(x), e^{-iz\cdot x}\rangle = \left\langle u(x), \frac{\partial}{\partial \bar{z}_i} e^{-iz\cdot x}\right\rangle = 0$$

since $z \mapsto e^{-iz\cdot x}$ is analytic and so has $\frac{\partial}{\partial \bar{z}_i} e^{-iz\cdot x} = 0$, for each $i = 1, \dots, n$.

We also have a nice growth bound on $\hat{u}(z)$:

$$|\hat{u}(z)| \leq C \sum_{|\alpha|\leq N} \sup_K \left|\partial_x^\alpha\left(e^{-iz\cdot x}\right)\right|$$

for some compact $K \subset \mathbb{R}^n$ and $C, N \geq 0$. This comes from the semi-norm definition of $\mathscr{E}'(\mathbb{R}^n)$.

Lemma 4.4. *Suppose $u \in \mathscr{E}'(\mathbb{R}^n)$ has $\operatorname{supp}(u) \subset \overline{B}_\delta(0)$. Then $\exists$ constants $C, N \geq 0$ such that*

$$|\hat{u}(z)| \leq C(1+|z|)^N e^{\delta|\mathrm{Im}(z)|} \qquad \textit{for all } z \in \mathbb{C}^n.$$

Proof. Fix $\varphi \in C^\infty(\mathbb{R})$ such that $\varphi \equiv 1$ for $\tau \geq -1/2$ and $\varphi \equiv 0$ on $\tau \leq -1$. Then define φ_ε for $\varepsilon > 0$ by:

$$\varphi_\varepsilon(x) := \varphi(\varepsilon(\delta - |x|)) \qquad \text{for } x \in \mathbb{R}^n.$$

Then we have $\varphi_\varepsilon \in \mathscr{D}(\mathbb{R}^n)$ and:

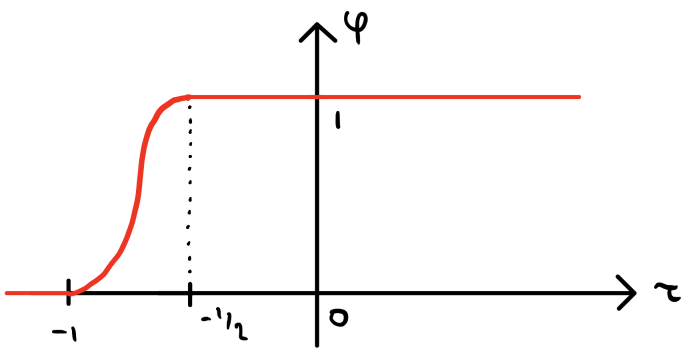

FIGURE 6. A sketch of φ.

$$\varphi_\varepsilon \equiv \begin{cases} 0 & \text{on } |x| \geq \delta + \frac{1}{\varepsilon} \\ 1 & \text{on } |x| \leq \delta + \frac{1}{2\varepsilon}. \end{cases}$$

Thus in particular $\varphi_\varepsilon \equiv 1$ on $\operatorname{supp}(u)$. So we have:

$$\hat{u}(z) := \langle u(x), e^{-iz\cdot x}\rangle = \langle u(x), \varphi_\varepsilon(x) e^{-iz\cdot x}\rangle.$$

So by the semi-norm estimate,

$$|\hat{u}(z)| \leq C \sum_{|\alpha|\leq N} \sup_{K_\varepsilon} \left|\partial_x^\alpha \left(\varphi_\varepsilon(x)e^{-iz\cdot x}\right)\right|$$

where $K_\varepsilon = \mathrm{supp}(\varphi_\varepsilon)$. Then on K_ε,

$$|\partial_x^\beta \varphi_\varepsilon| \lesssim \varepsilon^{|\beta|}$$

and

$$\left|\partial_x^\alpha (e^{-iz\cdot x})\right| \leq |z|^{|\alpha|}e^{\mathrm{Im}(z)\cdot x} \leq |z|^{|\alpha|}e^{\left(\delta+\frac{1}{\varepsilon}\right)|\mathrm{Im}(z)|}$$

and so by Liebniz we have

$$|\hat{u}(z)| \leq C \sum_{|\alpha|+|\beta|\leq N} \varepsilon^{|\beta|}|z|^{|\alpha|}e^{\left(\delta+\frac{1}{\varepsilon}\right)|\mathrm{Im}(z)|}$$

Now choose $\varepsilon = |z|$, and then combining we get:

$$|\hat{u}(z)| \lesssim (1+|z|)^N e^{\delta|\mathrm{Im}(z)|}$$

for any $z \in \mathbb{C}^n$. □

So in particular we see from Lemma 4.4 and the proceeding discussion that if $u \in \mathscr{E}'(\mathbb{R}^n)$ and $\mathrm{supp}(u) \subset \overline{B}_\delta$, then $\hat{u}(z)$ is entire (i.e. analytic for all $z \in \mathbb{C}^n$) and

$$|\hat{u}(z)| \lesssim (1+|z|)^N e^{\delta|\mathrm{Im}(z)|}.$$

The Paley-Wiener-Schwartz theorem is about the converse to this.

Theorem 4.5 (Paley-Wiener-Schwartz). *We have the following:*

(A) If $u \in \mathscr{E}'(\mathbb{R}^n)$ and $\mathrm{supp}(u) \subset \overline{B}_\delta(0)$, then $\hat{u}(z)$ is entire and obeys the estimate:

$$(\dagger) \qquad |\hat{u}(z)| \lesssim (1+|z|)^N e^{\delta|\mathrm{Im}(z)|}$$

for all $z \in \mathbb{C}^n$, for some $N \geq 0$.

Conversely, if $U = U(z)$ is an entire function and obeys (†), then $U = \hat{u}$ for some $u \in \mathscr{E}'(\mathbb{R}^n)$ with $\mathrm{supp}(u) \subset \overline{B}_\delta(0)$.

(B) If $u \in \mathscr{D}(\mathbb{R}^n)$ and $\mathrm{supp}(u) \subset \overline{B}_\delta(0)$, then $\hat{u}(z)$ is entire and obeys the estimate:

$$(\ddagger) \qquad |\hat{u}(z)| \lesssim_m (1+|z|)^{-m} e^{\delta|\mathrm{Im}(z)|} \quad \text{for all } m = 0,1,2,\ldots,\ z \in \mathbb{C}^n.$$

Conversely, if $U = U(z)$ is entire and satisfies (‡), then $U = \hat{u}$, where $u \in \mathscr{D}(\mathbb{R}^n)$ and $\mathrm{supp}(u) \subset \overline{B}_\delta(0)$.

Proof. (B): For $u \in \mathscr{D}(\mathbb{R}^n)$, then we know (as before) that

$$\hat{u}(z) = \int_{\mathbb{R}^n} e^{-ix\cdot z}u(x)\,\mathrm{d}x$$

is an entire function of $z \in \mathbb{C}^n$, and

$$\begin{aligned}
|z^\alpha \hat{u}(z)| &= \left| \int_{\mathbb{R}^n} u(x) D_x^\alpha \left(e^{-iz\cdot x}\right)\, \mathrm{d}x \right| \\
&= \left| \int_{\mathbb{R}^n} e^{-iz\cdot x} D^\alpha u \, \mathrm{d}x \right| \qquad \text{integrating by parts} \\
&\leq e^{\delta|\mathrm{Im}(z)|} \int_{\mathbb{R}^n} |D^\alpha u| \, \mathrm{d}x \\
&\lesssim_m e^{\delta|\mathrm{Im}(z)|}
\end{aligned}$$

where $m = |\alpha|$. This then implies that $|\hat{u}(z)| \lesssim_m (1+|z|)^{-m} e^{\delta|\mathrm{Im}(z)|}$ for all $z \in \mathbb{C}$.

For the converse, if U is as stated then $U = U(\lambda)$ ($\lambda \in \mathbb{R}^n$) decays rapidly as $|\lambda| \to \infty$, $\lambda \in \mathbb{R}^n$, and so by Fourier inversion we have $U = \hat{u}$, where

$$u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i\lambda\cdot x} U(\lambda) \, \mathrm{d}\lambda.$$

By differentiating under the integral and using dominated convergence, we have $u \in C^\infty(\mathbb{R}^n)$.

Then by Cauchy's theorem and (‡) we have (using Cauchy's theorem to change the line we are integrating over):

$$\int_{\mathbb{R}^n} e^{i\lambda\cdot x} U(\lambda) \, \mathrm{d}\lambda = \int_{\mathbb{R}^n} e^{i(\lambda+i\eta)\cdot x} U(\lambda+i\eta) \, \mathrm{d}\lambda.$$

So:

$$\begin{aligned}
|u(x)| &\lesssim \left| \int_{\mathbb{R}^n} U(\lambda+i\eta) e^{ix\cdot(\lambda+i\eta)} \, \mathrm{d}\lambda \right| \\
&\lesssim_m \int_{\mathbb{R}^n} (1+|\lambda+i\eta|)^{-m} e^{\delta|\eta|} e^{-x\cdot\eta} \, \mathrm{d}\lambda \\
&\lesssim_m e^{\delta|\eta| - x\cdot\eta}.
\end{aligned}$$

Now take $\eta = tx/|x|$ to get $|\hat{u}(z)| \lesssim_m e^{\delta t - t|x|} = e^{t(\delta-|x|)}$. Now if $|x| > \delta$, we can take $t \to \infty$ to deduce $u(x) = 0$. Hence $\mathrm{supp}(u) \subset \overline{B}_\delta(0)$.

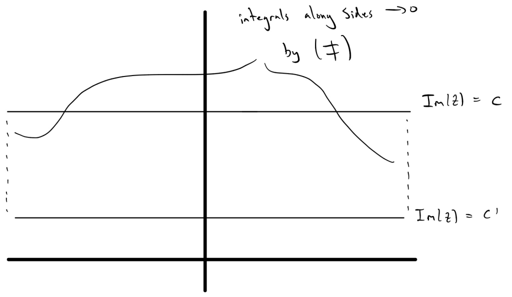

FIGURE 7. An illustration of how we use Cauchy's theorem here. As the integrals along the vertical sides $\to 0$ as we take the width to ∞, we get for f such that $|f| \to 0$ sufficiently rapidly in the horizontal directions, $\int_{\mathrm{Im}(\lambda_i)=c} f(\lambda_i)\, \mathrm{d}\lambda_i = \int_{\mathrm{Im}(\lambda_i)=c'} f(\lambda_i)\, \mathrm{d}\lambda_i$.

(A): We have already seen in Lemma 4.4 that $\hat{u}(z)$ is entire and (†) is satisfied if u is as stated.

For the converse, suppose that U is entire and satisfies (†). Then in particular:

$$|U(\lambda)| \leq C(1+|\lambda|)^N \quad \text{for all } \lambda \in \mathbb{R}^n$$

and so $U \in S'(\mathbb{R}^n)$. So by Theorem 3.2 (Fourier transform is an isomorphism on $S'(\mathbb{R}^n)$) we have $\exists u \in S'(\mathbb{R}^n)$ with $U = \hat{u}$.

Now fix $\varphi \in \mathscr{D}(\mathbb{R}^n)$ such that $\int_{\mathbb{R}^n} \varphi \, dx = 1$ and $\text{supp}(\varphi) \subset \overline{B}_1(0)$. Then for $\varepsilon > 0$, set $\varphi_\varepsilon(x) := \varepsilon^{-n}\varphi(x/\varepsilon)$, so that $\text{supp}(\varphi_\varepsilon) \subset \overline{B}_\varepsilon(0)$ and $\varphi_\varepsilon \to \delta_0$ in $S'(\mathbb{R}^n)$. Then set $u_\varepsilon := u * \varphi_\varepsilon$. [The idea here is to approximate and use (B).]

Then (by a question on Example Sheet 2), $\hat{u}_\varepsilon(z) = \hat{u}(z)\hat{\varphi}_\varepsilon(z) = U(z)\hat{\varphi}_\varepsilon(z)$, for $z \in \mathbb{C}^n$. Thus we see that $\hat{u}_\varepsilon$ is entire, and:

$$|\hat{u}_\varepsilon(z)| \lesssim_m C(1+|z|)^N e^{\delta|\text{Im}(z)|} \cdot (1+|z|)^{-m} e^{\varepsilon|\text{Im}(z)|}$$

for some $N \geq 0$ and each $m = 0, 1, 2, \ldots$ (from (B)). So by taking $m = N + m'$, $m' = 0, 1, 2, \ldots$, we see that

$$|\hat{u}_\varepsilon(z)| \lesssim_{m'} (1+|z|)^{-m'} e^{(\delta+\varepsilon)|\text{Im}(z)|} \qquad \text{for all } m' = 0, 1, 2, \ldots.$$

So hence $u_\varepsilon \in \mathscr{D}(\mathbb{R}^n)$ and $\text{supp}(u_\varepsilon) \subset \overline{B}_{\delta+\varepsilon}(0)$, by (B). Then since $u_\varepsilon \to u$ in $S'(\mathbb{R}^n)$ as $\varepsilon \downarrow 0$, we deduce from this by taking $\varepsilon \downarrow 0$ that $\text{supp}(u) \subset \overline{B}_\delta(0)$. So done.

□

We finish looking at the Paley-Wiener-Schwartz theorem with an example of its use.

Example 4.3. *Suppose we have $(f_m)_{m=1}^N$ which are entire functions $f_m : \mathbb{C}^n \to \mathbb{C}$. Suppose that we have*

$$\left|e^{iy_m \cdot z}\hat{f}_m(z)\right| \lesssim_m (1+|z|)^m e^{\frac{|\text{Im}(z)|}{m+1}}$$

for $m = 1, \ldots, N$ and $y_m \in \mathbb{Z}^n$.

Question: *Are the $(f_m)_m$ linearly independent?*

Answer: *The bound we have on the f_m looks very similar to the bounds in the Paley-Wiener-Schwartz theorem, except we have extra factors of $e^{iy_m \cdot z}$ on the LHS. However these just correspond to Fourier transforms of the translated f_m, i.e. $\widehat{(\tau_{-y_m} f_m)}(z) = e^{iy_m \cdot z}\hat{f}_m(z)$.*

So hence if $f_m = \hat{u}_m$, then the Paley-Wiener-Schwartz theorem gives $\text{supp}(u_m) \subset B_{\frac{1}{m+1}}(y_m)$.

Hence if we had $\sum_m \alpha_m f_m(z) = 0$, then by taking the Fourier transform this is true if and only if

$$\sum_m \alpha_m u_m = 0.$$

But then by the above the u_m have disjoint supports, and so this is true if and only if $\alpha_m = 0$ for all m. So hence the $(f_m)_m$ are linearly independent.

[Pictorially we have a lattice with the supports being disjoint balls centred at the lattice points.] □

5. Oscillatory Integrals

At the start of the course, we considered "strange" integrals, like
$$\int_{-\infty}^{\infty} e^{i\lambda x}\ \mathrm{d}x.$$
This is an example of an **oscillatory integral**, and it is <u>not</u> to be interpreted in terms of Riemann/Lebesgue integrability. We will look at objects of the form:
$$\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\ \mathrm{d}\theta$$
where $(x,\theta)\in X\times\mathbb{R}^k$, for $X\subset\mathbb{R}$ open. Here Φ is called the **phase function** and a is called the **symbol**.

We will allow $a(x,\theta)$ to grow as $|\theta|\to\infty$, e.g. we could take $a(x,\theta)=p(\theta)$, for some polynomial p.

Key idea: Just like alternating series don't need as much decay to converge(xvii), we can the oscillation and integration by parts to control the integral and hopefully "cancel out" the growth of a.

For example, for $\varphi\in\mathscr{D}(\mathbb{R})$ we know that
$$\int_{\mathbb{R}} e^{i\lambda\theta}\varphi(\theta)\ \mathrm{d}\theta=\int_{\mathbb{R}}\frac{1}{i\lambda}\frac{\mathrm{d}}{\mathrm{d}\theta}\left(e^{i\lambda\theta}\right)\varphi(\theta)\ \mathrm{d}\theta\equiv\int_{\mathbb{R}} L\left[e^{i\lambda\theta}\right]\varphi(\theta)\ \mathrm{d}\theta$$
where $L:=\frac{1}{i\lambda}\cdot\frac{\partial}{\partial\theta}$ is a differential operator which has $L(e^{i\lambda\theta})=e^{i\lambda\theta}$. So, on integrating by parts:
$$\begin{aligned}\int_{\mathbb{R}} e^{i\lambda\theta}\varphi(\theta)\ \mathrm{d}\theta&=\int_{\mathbb{R}} L^{(N)}\left[e^{i\lambda\theta}\right]\varphi(\theta)\ \mathrm{d}\theta\\ &\int_{\mathbb{R}}(L^*)^{(N)}[\varphi]e^{i\lambda\theta}\ \mathrm{d}\theta\\ &=\mathscr{O}(\lambda^{-N})\end{aligned}$$
as $L^*=-\frac{1}{i\lambda}\frac{\partial}{\partial\theta}$. So we see that we have rapid decay here as $\lambda\to\infty$, since $N\in\mathbb{N}$ was arbitrary.

More generally considering $\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}\varphi(\theta)\ \mathrm{d}\theta$ for $\Phi\in C^\infty(\mathbb{R})$, we can play the same game: we have
$$\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}\varphi(\theta)\ \mathrm{d}\theta=\int_{\mathbb{R}} L\left[e^{i\lambda\Phi(\theta)}\right]\varphi(\theta)\ \mathrm{d}\theta$$
where
$$L=\frac{1}{i\lambda\Phi'(\theta)}\frac{\mathrm{d}}{\mathrm{d}\theta}$$
(let us just assume for the moment that $\Phi'\neq 0$ on $\mathrm{supp}(\varphi)$ so that L is well-defined here). Again, integrating by parts N times, using the fact that $L^*=-\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{1}{i\lambda\Phi'(\theta)}\right]$, we get
$$\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}\varphi(\theta)\ \mathrm{d}\theta=\mathscr{O}(\lambda^{-N}).$$
So what we are seeing is that "oscillation implies decay", and lots of it if $\Phi'\neq 0$.

(xvii)Consider $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\cdots$ vs $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots$

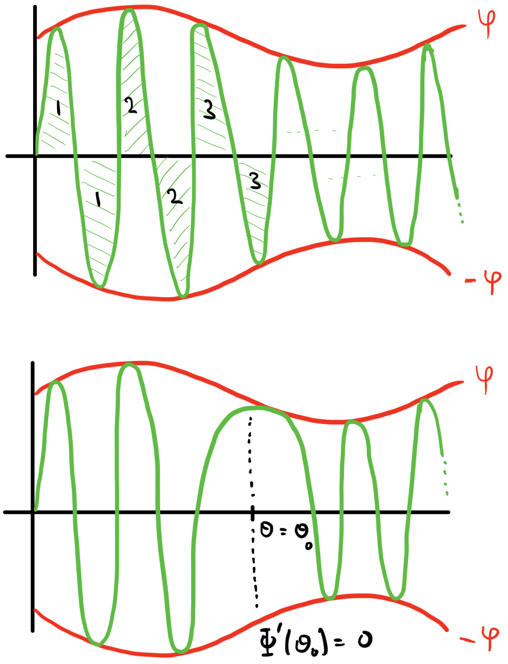

FIGURE 8. These two illustrations demonstrate the type types of behaviour we can see. In the top image, we see that Φ larger and larger (i.e. period smaller and smaller) the bumps labelled 1, those labelled 2, etc, cancel each other out more and more as the peaks are then on smaller and smaller intervals and φ does not change so much on them. This means that both the peaks will be around the same size and so cancel. In the second image, we see what happens when $\Phi' = 0$ somewhere - the shape of $e^{i\lambda\Phi(\theta)}\varphi(\theta)$ (or at least the real/imaginary parts of) flattens out at these points, and so we don't have the cancellation as before. In some sense there is a double zero for Φ here (note that $\frac{\mathrm{d}}{\mathrm{d}\theta}(e^{i\lambda\Phi}\varphi) = \varphi' e^{i\lambda\Phi} + i\lambda\Phi' e^{i\lambda\Phi}\varphi)$ since both Φ and Φ' are zero at such a point. So we need to check the integral about these points still decay as $\lambda \to \infty$. This is what stationary phase says.

Lemma 5.1 (Stationary Phase). *Let $\Phi \in C^\infty(\mathbb{R})$ have $\Phi(0) = \Phi'(0) = 0$ (wlog by translating), with $\Phi''(0) \neq 0$, and $\Phi'(\theta) \neq 0$ for $\theta \in \mathbb{R}\backslash\{0\}$. Then for $\varphi \in \mathscr{D}(\mathbb{R})$ we have:*

$$\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}\varphi(\theta)\,\mathrm{d}\theta = \mathscr{O}\left(\lambda^{-\frac{1}{2}}\right) \qquad \textit{as } \lambda \to \infty.$$

Note: Before when $\Phi' \neq 0$ we saw that we got $\mathscr{O}(\lambda^{-N})$ decay for arbitrary N. So these points of stationary phase are bad indeed, but we can get some control!

Proof. Fix $\rho \in \mathscr{D}(\mathbb{R})$ with $\rho \equiv 1$ on $|\theta| < 1$ and $\rho \equiv 0$ on $|\theta| > 2$. Then for $\delta > 0$ write:

$$\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}\varphi(\theta)\,\mathrm{d}\theta = \underbrace{\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}\rho\left(\frac{\theta}{\delta}\right)\varphi(\theta)\,\mathrm{d}\theta}_{=:I_1(\lambda)} + \underbrace{\int_{\mathbb{R}}\left(1-\rho\left(\frac{\theta}{\delta}\right)\right)e^{i\lambda\Phi(\theta)}\varphi(\theta)\,\mathrm{d}\theta}_{=:I_2(\lambda)}$$

then since $\rho(\theta/\delta) = 0$ on $|\theta| > 2\delta$, we have $I_1(\lambda) = \mathscr{O}(\delta)$.

Now note that since $1-\rho(\theta/\delta)=0$ on $|\theta|<\delta$, we can integrate by parts in $I_2(\lambda)$ (since we can write the integral as one over $|\theta|>\delta$, and we can integrate by parts everywhere where $\Phi'\neq 0$). So setting $L=\frac{1}{i\lambda\Phi'(\theta)}\frac{\mathrm{d}}{\mathrm{d}\theta}$, upon integrating by parts (since $L[e^{i\lambda\Phi}]=e^{i\lambda\Phi}$) we have (there are never any boundary terms due to compact support):

$$I_2(\lambda)=\int_{\mathbb{R}} e^{i\lambda\Phi(\theta)}L^*\left[\left(1-\rho\left(\frac{\theta}{\delta}\right)\right)\varphi(\theta)\right]\,\mathrm{d}\theta.$$

But then

$$\begin{aligned}L^*\left[\left(1-\rho\left(\frac{\theta}{\delta}\right)\right)\varphi(\theta)\right]&:=-\frac{1}{i\lambda}\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{(1-\rho(\theta/\delta))\varphi(\theta)}{\Phi'(\theta)}\right]\\&=\frac{1}{i\lambda}\cdot\frac{\Phi''(\theta)\left(1-\rho\left(\frac{\theta}{\delta}\right)\right)\varphi(\theta)}{\Phi'(\theta)^2}+\frac{1}{i\lambda}\cdot\frac{1}{\delta}\cdot\frac{\rho'\left(\frac{\theta}{\delta}\right)\varphi(\theta)}{\Phi'(\theta)}-\frac{1}{i\lambda}\cdot\frac{\left(1-\rho\left(\frac{\theta}{\delta}\right)\right)\varphi'(\theta)}{\Phi'(\theta)}.\end{aligned}$$

Note:

$$\Phi'(\theta)=\Phi'(\theta)-\underbrace{\Phi'(0)}_{=0\text{ by assumption}}=\int_0^\theta\Phi''(\tau)\,\mathrm{d}\tau=\theta\int_0^1\Phi''(\tau\theta)\,\mathrm{d}\tau$$

and so

$$\frac{\Phi'(\theta)}{\theta}=\int_0^1\Phi''(\tau\theta)\,\mathrm{d}\tau.$$

So since $\Theta''(\theta)\neq 0$ on a neighbourhood of 0 (as $\Phi''(0)\neq 0$ and by continuity of Φ'') we see that:

$$\frac{|\theta|}{|\Phi'(\theta)|}\lesssim 1$$

near 0. Now since $\mathrm{supp}(\varphi)$ is compact, this gives that we have $\frac{1}{|\Phi'(\theta)|}\lesssim\frac{1}{|\theta|}$ on the entire range of integration in I_2. Hence:

$$L^*\left[\left(1-\rho\left(\frac{\theta}{\delta}\right)\right)\varphi(\theta)\right]=\mathcal{O}\left(\frac{1}{\lambda|\theta|^2}\right)+\mathcal{O}\left(\frac{1}{\lambda}\cdot\frac{1}{\delta|\theta|}\right)=\mathcal{O}\left(\frac{1}{\lambda|\theta|}\left[\frac{1}{|\theta|}+\frac{1}{\delta}\right]\right)$$

since the third term does not matter.

Then iterating this N times we see:

$$(L^*)^{(N)}[\cdots]\lesssim\frac{1}{\lambda^N\theta^N}\left(\frac{1}{|\theta|}+\frac{1}{\delta}\right)^N\lesssim\frac{1}{\lambda^N|\theta|^N\delta^N}$$

since on I_2 we have $|\theta|>\delta$ and so $\frac{1}{|\theta|}<\frac{1}{\delta}$. So:

$$I_2(\lambda)\lesssim\lambda^{-N}\delta^{-N}\int_{|\theta|>\delta}|\theta|^{-N}\,\mathrm{d}\theta\lesssim\lambda^{-N}\delta^{-2N+1}$$

since when we integrate x^{-N} we get $C\cdot x^{-N+1}$. So combining we see:

$$\int_{\mathbb{R}}e^{i\lambda\Phi(\theta)}\varphi(\theta)\,\mathrm{d}\theta\lesssim\max\{\delta,\lambda^{-N}\delta^{-2N+1}\}.$$

To minimise the RHS we should choose δ such that $\delta=\lambda^{-N}\delta^{-2N+1}$, i.e. $\delta=\lambda^{-1/2}$. Hence:

$$\int_{\mathbb{R}}e^{i\lambda\Phi(\theta)}\varphi(\theta)\,\mathrm{d}\theta=\mathcal{O}(\lambda^{-1/2})\quad\text{as }\lambda\to\infty.$$

□

This lemma of stationary phase suggests that control of $\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\,\mathrm{d}\theta$ will depend heavily on $|\nabla_\theta a(x,\theta)|$, where

$$\nabla_\theta a = \left(\frac{\partial a}{\partial \theta_1},\dots,\frac{\partial a}{\partial \theta_k}\right).$$

Now let us look at defining oscillatory integrals properly via distributions.

Definition 5.1. *Let $X \subset \mathbb{R}^n$ be open. Then a smooth function $a : X \times \mathbb{R}^k \to \mathbb{C}$ is called a* ***symbol of order N*** *for for all $K \subset X$ compact and for all pairs of multi-indices (α,β), $\exists$ a constant C such that:*

$$\left|D_x^\alpha D_\theta^\beta a(x,\theta)\right| \le C\langle\theta\rangle^{N-|\beta|}$$

for $(x,\theta)\in K\times\mathbb{R}^k$, for $K \subset X$ compact. We call the space of all such symbols $\mathrm{Sym}(X,\mathbb{R}^k;N)$.

Remark: If $a(x,\theta) = P(\theta)$ for a polynomial P of degree N, then $a \in \mathrm{Sym}(X,\mathbb{R}^k;N)$. Intuitively, from the definition of a symbol we expect symbols to look like a polynomial in θ with coefficients that are smooth functions in x (although they are more general than this, since a symbol can have negative order!).

We only care about the large $|\theta|$ behaviour of such symbols because we can always write:

$$\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\,\mathrm{d}\theta = \underbrace{\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\rho(\theta)\,\mathrm{d}\theta}_{\text{classically well-defined}} + \underbrace{\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)(1-\rho(\theta))\,\mathrm{d}\theta}_{\text{oscillatory integral with symbol } = 0 \text{ on } |\theta|<R}$$

where $\rho \in \mathscr{D}(\mathbb{R}^k)$ has $\rho \equiv 1$ on $|\theta| < R$. So we only care about the symbol for $|\theta| > R$, but $R > 0$ was arbitrary!

Definition 5.2. *A function $\Phi : X \times \mathbb{R}^k \to \mathbb{R}$ is called a* ***phase function*** *if:*

(i) Φ *is continuous and (positively) homogeneous of degree 1 in θ (i.e. $\Phi(x,\tau\theta) = \tau\Phi(x,\theta)$ for all $\tau > 0$).*

(ii) Φ *is smooth on $X \times \left(\mathbb{R}^k\backslash\{0\}\right)$.*

(iii) $\mathrm{d}\Phi = \nabla_x\Phi\cdot\mathrm{d}x + \nabla_\theta\Phi\cdot\mathrm{d}\theta$ *is non-vanishing on $X \times \left(\mathbb{R}^k\backslash\{0\}\right)$.*

Remark: The reason behind these conditions for a phase function are as follows. To start with we want to try to make sense of $\int_{\mathbb{R}^n} e^{i\lambda\cdot x}$, i.e. here we have $\Phi(x,\theta) = x\cdot\theta$. So we try to prove results for <u>this</u> Φ. Then afterwards we look back at the proofs and ask "what properties of the Φ did we actually use/need?". Then we would see that the above conditions of a phase function are exactly those. [We will work with the general definition above and not a specific Φ, but this is a reason for why the phase function definition is how it is.]

The standard example of a phase function is $\Phi(x,\theta) = x\cdot\theta$, for $(x,\theta)\in\mathbb{R}^n\times\mathbb{R}^n$. To see that this is a phase function, note that (i), (ii) are immediate, and

$$\mathrm{d}\Phi = \theta\cdot\mathrm{d}x + x\cdot\theta\theta.$$

But $\{d\theta_1, \ldots, d\theta_n, dx_1, \ldots, dx_n\}$ are linearly independent, and so this is non-zero on $\mathbb{R}^n \times (\mathbb{R}^n \backslash \{0\})$.

Then the classical example of an oscillatory integral is:

$$\int_{\mathbb{R}^n} \frac{\theta^\alpha}{(2\pi)^n} e^{ix\cdot\theta}\, d\theta \qquad \text{for } (x,\theta) \in \mathbb{R}^n \times \mathbb{R}^n.$$

A natural question is whether this equals $D^\alpha \delta_0(x)$ in the sense of distributions.

Lemma 5.2. *Let $a \in \mathrm{Sym}(X, \mathbb{R}^k; N)$ and $a_i \in \mathrm{Sym}(X, \mathbb{R}^k; N_i)$ for $i = 1, 2$, Then:*

(i) $D_x^\alpha D_\theta^\beta a \in \mathrm{Sym}(X, \mathbb{R}^k; N - |\beta|)$, *i.e. only derivatives in θ change order of the symbol,*
(ii) $a_1 a_2 \in \mathrm{Sym}(X, \mathbb{R}^k; N_1 + N_2)$.

Proof. (i): If $a \in \mathrm{Sym}(X, \mathbb{R}^k; N)$, then $D^\alpha x D_\theta^\beta a$ is smooth and:

$$\left| D_x^{\alpha'} D_\theta^{\beta'} \left(D_x^\alpha D_\theta^\beta a \right) \right| = \left| D_x^{\alpha+\alpha'} D_\theta^{\beta+\beta'} a \right| \lesssim_{\alpha',\beta',K} \langle\theta\rangle^{N-|\beta|-|\beta'|}$$

for $K \subset X$ compact. So $D_x^\alpha D_\theta^\beta a \in \mathrm{Sym}(X, \mathbb{R}^k; N - |\beta|)$.

(ii): Now if $a_i \in \mathrm{Sym}(X, \mathbb{R}^k; N_i)$ for $i = 1, 2$, then $a_1 a_2$ is smooth and:

$$\begin{aligned}
\left| D_x^\alpha D_\theta^\beta (a_1 a_2) \right| &= \left| \sum_{\alpha' \le \alpha} \sum_{\beta' \le \beta} \binom{\alpha}{\alpha'} \binom{\beta}{\beta'} D_x^{\alpha'} D_\theta^{\beta'} a_1 \cdot D_x^{\alpha-\alpha'} D_\theta^{\beta-\beta'} a_2 \right| \\
&\lesssim_{\alpha,\beta} \sum_{\alpha' \le \alpha} \sum_{\beta' \le \beta} \left| D_x^{\alpha'} D_\theta^{\beta'} a_1 \right| \cdot \left| D_x^{\alpha-\alpha'} D_\theta^{\beta-\beta'} a_2 \right| \\
&\lesssim_{\alpha,\beta,K} \sum_{\alpha' \le \alpha} \sum_{\beta' \le \beta} \langle\theta\rangle^{N_1 - |\beta'|} \langle\theta\rangle^{N_2 - (|\beta| - |\beta'|)} \\
&\lesssim_{\alpha,\beta,K} \langle\theta\rangle^{N_1+N_2-|\beta|}
\end{aligned}$$

and thus we see $a_1 a_2 \in \mathrm{Sym}(X, \mathbb{R}^k; N_1 + N_2)$.

□

Now to define oscillatory integrals, we could define $\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)} a(x,\theta)\, d\theta$ as a linear form on $\mathscr{D}(X)$, via:

$$\varphi \longmapsto \langle I_\Phi(a), \varphi \rangle := \int_{\mathbb{R}^k} \left(\int_X e^{i\Phi(x,\theta)} a(x,\theta) \varphi(x)\, dx \right) d\theta$$

for $\varphi \in \mathscr{D}(X)$.

However this definition is cumbersome due to the lack of absolute convergence in the double integral (as there is no decay guaranteed in the θ-direction. The x-integral is fine). So instead we define:

$$\langle I_\Phi(a), \varphi \rangle := \lim_{\varepsilon \downarrow 0} \langle I_{\Phi,\varepsilon}(a), \varphi \rangle$$

where

$$I_{\Phi,\varepsilon}(a) := \int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\chi(\varepsilon\theta)\,\mathrm{d}\theta$$

where $\chi \in \mathscr{D}(\mathbb{R}^k)$ has $\chi \equiv 1$ on $|\theta| < 1$.

Using this, we can interchange the orders of integration, etc, before taking $\varepsilon \downarrow 0$.

Theorem 5.1 (Oscillatory integrals are distributions of finite order)**.** *For Φ a phase function and $a \in \mathrm{Sym}(X, \mathbb{R}^k; N)$, the oscillatory integral $I_\Phi(a) := \lim_{\varepsilon\downarrow 0} I_{\Phi,\varepsilon}(a)$ defines an element of $\mathscr{D}'(X)$ of order $\leq N + k + 1$.*

Proof. Later. □

Remark: We can actually improve this result to get that the order is $\leq N + k$, but we won't discuss this.

To establish Theorem 5.1, we need an integration by parts trick (similar to that in Lemma 5.1/Stationary Phase) to get more control of the integral as $|\theta| \to \infty$.

Lemma 5.3. *If a differential operator L has the form:*

$$(\star) \qquad L = \sum_{j=1}^{k} a_j(x,\theta)\frac{\partial}{\partial\theta_j} + \sum_{j=1}^{n} b_j(x,\theta)\frac{\partial}{\partial x_j} + c(x,\theta)$$

where $a_j \in \mathrm{Sym}(X, \mathbb{R}^k; 0)$ and $b_j, c \in \mathrm{Sym}(X, \mathbb{R}^k; -1)$, then the formal adjoint L^* has/takes the same form.

Note: This is just like how in Lemma 5.1/stationary phase we had $L = \frac{1}{i\lambda\Phi'}\frac{\mathrm{d}}{\mathrm{d}\theta}$.

Proof. The formal adjoint of such an L is:

$$\begin{aligned} L^* &= -\sum_{j=1}^{k}\frac{\partial}{\partial\theta_j}\left[a_j(x,\cdot)\,\bullet\right] - \sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left[b_j(x,\theta)\,\bullet\right] + c(x,\theta) \\ &= \sum_{j=1}^{k}\tilde{a}_j\frac{\partial}{\partial\theta_j} + \sum_{j=1}^{n}\tilde{b}_j\frac{\partial}{\partial x_j} + \tilde{c} \end{aligned}$$

where we can see that $\tilde{a}_j = -a_j \in \mathrm{Sym}(X, \mathbb{R}^k, 0)$, $\tilde{b}_j = -b_j \in \mathrm{Sym}(X, \mathbb{R}^k; -1)$, and

$$\tilde{c} = c - \sum_{j=1}^{k}\frac{\partial a_j}{\partial\theta_j} - \sum_{j=1}^{n}\frac{\partial b_j}{\partial x_j} \qquad \in \mathrm{Sym}(X, \mathbb{R}^k; -1)$$

since $c \in \mathrm{Sym}(X, \mathbb{R}^k; -1)$, $\frac{\partial a_j}{\partial\theta_j} \in \mathrm{Sym}(X, \mathbb{R}^k; -1)$ by Lemma 5.2(i), and $\frac{\partial b_j}{\partial x_j} \in \mathrm{Sym}(X, \mathbb{R}^k; -1)$ by Lemma 5.2(i). Hence this takes on the same form and so we are done.

□

Note: By "formal adjoint", we just mean the differential operator which obeys the integration by parts rule:

$$\int (Lf)\cdot g = \int f\cdot (L^* g).$$

It is not the adjoint with respect to the L^2-inner product (where complex conjugates get involved).

Lemma 5.4. *For a given phase function Φ, $\exists L$ of the form of $(\star)$ in Lemma 5.3 such that*

$$L^*\left[e^{i\Phi}\right] = e^{i\Phi}.$$

Aside: What is the point of having L as the form of $(\star)$? The point is to get absolute integrability in the double integral:

$$\begin{aligned}\int_{\mathbb{R}^k}\int_X e^{i\Phi(x,\theta)}a(x,\theta)\chi(\varepsilon\theta)\varphi(x)\,\mathrm{d}x\mathrm{d}\theta &= \int_{\mathbb{R}^k}\int_X \left(L^* e^{i\Phi}\right)a(x,\theta)\chi(\varepsilon\theta)\varphi(x)\,\mathrm{d}x\mathrm{d}\theta\\ &= \int_{\mathbb{R}^k}\int_X e^{i\Phi}L\left(a(x,\theta)\chi(\varepsilon\theta)\varphi(x)\right)\,\mathrm{d}x\mathrm{d}\theta.\end{aligned}$$

Now, $a(x,\theta)\chi(\varepsilon\theta)\varphi(x)$ is like a symbol of order N. So what does L do to a symbol of order N? Well, acting by $\frac{\partial}{\partial\theta_j}$ lowers the order by 1, whilst multiplying by a_j does not change the order. Hence $a_j\frac{\partial}{\partial\theta_j}$ lowers the order by 1. Similarly, acting by $\frac{\partial}{\partial x_j}$ does not change the order, whilst multiplying by b_j lowers the order by 1. Hence acting by $b_j\frac{\partial}{\partial x_j}$ lowers the order by 1. Hence L lowers the order by 1, and so we expect $L(a(x,\theta)\chi(\varepsilon\theta)\varphi(x))$ to act like a symbol of order $N-1$!

But then we could repeat this to get

$$\int_{\mathbb{R}^k}\int_X e^{i\Phi}\underbrace{L^{(M)}\left[a(x,\theta)\chi(\varepsilon\theta)\varphi(x)\right]}_{\text{symbol of order } N-M}\,\mathrm{d}x\mathrm{d}\theta$$

and so if we keep doing this until $N-M<0$, we get absolute integrability and so can interchange the order of integration, etc.

Proof of Lemma 5.4. Note that

$$\frac{\partial}{\partial\theta_j}e^{i\Phi} = i\frac{\partial\Phi}{\partial\theta_j}e^{i\Phi} \qquad\text{and}\qquad \frac{\partial}{\partial x_j}e^{i\Phi} = i\frac{\partial\Phi}{\partial x_j}e^{i\Phi}.$$

Then note

$$\begin{aligned}\left(-i|\theta|^2\sum_{j=1}^k\frac{\partial\Phi}{\partial\theta_j}\cdot\frac{\partial}{\partial\theta_j} - i\sum_{j=1}^n\frac{\partial\Phi}{\partial x_j}\cdot\frac{\partial}{\partial x_j}\right)e^{i\Phi} &= \left(|\theta|^2\sum_{j=1}^k\left|\frac{\partial\Phi}{\partial\theta_j}\right|^2 + \sum_{j=1}^n\left|\frac{\partial\Phi}{\partial x_j}\right|^2\right)e^{i\Phi}\\ &= \left(|\theta|^2|\nabla_\theta\Phi|^2 + |\nabla_x\Phi|^2\right)e^{i\Phi}\\ &=: \frac{1}{\pi(x,\theta)}e^{i\Phi(x,\theta)}.\end{aligned}$$

Now note that, since $\Phi(x,\theta)$ is positively homogeneous of degree 1 in θ, we have for $\tau>0$ (a constant):

$$\tau\frac{\partial}{\partial x_j}\Phi(x,\theta) = \frac{\partial}{\partial x_j}\Phi(x,\tau\theta) = \frac{\partial\Phi}{\partial x_j}(x,\tau\theta)$$

and so $\frac{\partial \Phi}{\partial x_j}$ is positively homogeneous of degree 1 in θ. Also,

$$\tau \frac{\partial}{\partial \theta_j}\Phi(x,\theta) = \frac{\partial}{\partial \theta_j}\Phi(x,\tau\theta) = \tau\frac{\partial \Phi}{\partial \theta_j}(x,\tau\theta)$$

and so $\frac{\partial \Phi}{\partial \theta_j}$ is positively homogeneous of degree 0 in θ. So,

$$(x,\theta) \longmapsto |\theta|^2 \cdot |\nabla_\theta \Phi|^2 + |\nabla_x \Phi|^2$$

is positively homogeneous of degree 2 in θ, since the first term on the RHS is a product of a term of homogeneous degree 2 and one of degree 0, whilst the one of the right is the square of a term of homogeneous degree 1, and so also has degree 2. Hence $\pi(x,\theta)$ is homogeneous of degree -2. Then define:

$$\tilde{L} := \pi(x,\theta)\left[-i|\theta|^2\sum_{j=1}^k \frac{\partial \Phi}{\partial \theta_j}\cdot\frac{\partial}{\partial \theta_j} - i\sum_{j=1}^n \frac{\partial \Phi}{\partial x_j}\cdot\frac{\partial}{\partial x_j}\right].$$

By the above we know that $\tilde{L}e^{i\Phi} = e^{i\Phi}$, but $\pi(x,\theta)$ could be singular at $\theta = 0$ (i.e. if $|\nabla_x \Phi| = 0$ at $\theta = 0$, which isn't ruled out by the assumption on Φ that $\mathrm{d}\Phi \neq 0$ on $X \times \left(\mathbb{R}^k \backslash \{0\}\right)$).

So to deal with this, fix $\rho \in \mathscr{D}(\mathbb{R}^k)$ with $\rho = 1$ on $|\theta| < 1$. Set:

$$L^* := (1-\rho(\theta))\tilde{L} + \rho(\theta).$$

Then we have

$$L^* e^{i\Phi} = (1-\rho(\theta))e^{i\Phi} + \rho(\theta)e^{i\Phi} = e^{i\Phi}$$

and

$$L^* = \sum_j \tilde{a}_j \frac{\partial}{\partial \theta_j} + \sum_j \tilde{b}_j \frac{\partial}{\partial x_j} + \tilde{c}$$

where

$$\tilde{a}_j = -\frac{i|\theta|^2(1-\rho(\theta))\frac{\partial \Phi}{\partial \theta_j}}{|\theta|^2\cdot|\nabla_\theta \Phi|^2 + |\nabla_x \Phi|^2}$$

is a smooth function (as we have removed the problem at $\theta = 0$ with the $1-\rho(\theta)$ factor), and is homogeneous of degree 0 by construction. Moreover we have

$$\tilde{b}_j = \frac{-i(1-\rho(\theta))\frac{\partial \Phi}{\partial x_j}}{|\theta|^2\cdot|\nabla_\theta \Phi|^2 + |\nabla_x \Phi|^2}$$

which is smooth and homogeneous of degree -1, since the numerator has degree 1 and the denominator has degree 2. Finally we have

$$\tilde{c} = \rho(\theta).$$

So since $\tilde{a}_j$ is homogeneous of degree 0 on $|\theta| > 1$ and smooth, we have $\tilde{a}_j \in \mathrm{Sym}(X,\mathbb{R}^k;0)$. Similarly $\tilde{b}_j, \tilde{c} \in \mathrm{Sym}(X,\mathbb{R}^k;-1)$. So by Lemma 5.3, we have $L = (L^*)^*$ has the same form and so we are done.

$\square$

Note: By definition of L, if $a \in \mathrm{Sym}(X,\mathbb{R}^k;N)$, then we have already argued that

$$L(a) \in \mathrm{Sym}(X,\mathbb{R}^k;N-1)$$

i.e. L lowers the order by 1 for a symbol. More generally if φ is smooth, we have

$$L^{(M)}[a(x,\theta)\varphi(x)] = \sum_{|\alpha|\leq M} a_\alpha(x,\theta)\partial^\alpha \varphi(x)$$

where $a_\alpha \in \mathrm{Sym}(X,\mathbb{R}^k;N-M)$. This can be proved via induction on M.

Proof of Theorem 5.1. By definition we have $I_\Phi(a) := \lim_{\varepsilon\downarrow 0} I_{\Phi,\varepsilon}(a)$, where

$$I_{\Phi,\varepsilon}(a) := \int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\chi(\varepsilon\theta)$$

and for $\varphi \in \mathscr{D}(X)$,

$$\langle I_{\Phi,\varepsilon}(a),\varphi\rangle = \int_{\mathbb{R}^k}\int_X e^{i\Phi(x,\theta)}a(x,\theta)\chi(\varepsilon\theta)\varphi(x)\,\mathrm{d}x\mathrm{d}\theta.$$

By Lemma 5.4, we know $\exists L$ of the form of $(\star)$ in Lemma 5.3 such that $L^*e^{i\Phi} = e^{i\Phi}$. Using this (M times) and integrating by parts we see that

$$\langle I_{\Phi,\varepsilon}(a),\varphi\rangle = \int_{\mathbb{R}^k}\int_X e^{i\Phi(x,\theta)}L^{(M)}[a(x,\theta)\chi(\varepsilon\theta)\varphi(x)]\,\mathrm{d}x\mathrm{d}\theta.$$

Now observe that

$$\begin{aligned}\left|\left(\frac{\partial}{\partial\theta}\right)^\alpha \chi(\varepsilon\theta)\right| &= |\varepsilon|^{|\alpha|}(\partial^\alpha\chi)(\varepsilon\theta)\\ &\lesssim \varepsilon^{|\alpha|}\langle\varepsilon\theta\rangle^{-|\alpha|}\\ &= C_\alpha\cdot\frac{\varepsilon^{|\alpha|}}{(1+|\varepsilon\theta|^2)^{\alpha/2}}\\ &= C_\alpha\left(\frac{1}{\varepsilon^2}+|\theta|^2\right)^{-|\alpha|/2}\\ &\lesssim_\alpha \langle\theta\rangle^{-|\alpha|}\end{aligned}$$

for $\varepsilon\in(0,1]$, where the second line is true since χ is smooth of compact support and so $|\partial^\alpha\chi(\theta)| \lesssim_\alpha \langle\theta\rangle^{-|\alpha|}$. Moreover the last line holds uniformly for $\varepsilon\in(0,1]$. So hence this tells us that, uniformly in ε, we can treat $\theta\mapsto\chi(\varepsilon\theta)$ as a symbol of order 0. Hence

$$a(x,\theta)\chi(\varepsilon\theta)\in\mathrm{Sym}(X,\mathbb{R}^k;N)$$

uniformly in $\varepsilon\in(0,1]$. So:

$$L^{(M)}[a(x,\theta)\chi(\varepsilon\theta)\varphi(x)] = \sum_{|\alpha|\leq M} a_\alpha(x,\theta,\varepsilon)\partial^\alpha\varphi(x)$$

uniformly in ε, where $a_\alpha\in\mathrm{Sym}(X,\mathbb{R}^k;N-M)$. Now if we choose M such that $N-M<-k$, then the a_α are integrable. Thus by dominated convergence,

$$\begin{aligned}\langle I_\Phi(a),\varphi\rangle &= \lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^k}\int_X e^{i\Phi(x,\theta)}\sum_{|\alpha|\leq M} a_\alpha(x,\theta;\varepsilon)\partial^\alpha\varphi(x)\,\mathrm{d}x\mathrm{d}\theta\\ &= \sum_{|\alpha|\leq M}\int_{\mathbb{R}^k}\int_X e^{i\Phi(x,\theta)}a_\alpha(x,\theta)\partial^\alpha\varphi(x)\,\mathrm{d}x\mathrm{d}\theta\end{aligned}$$

where $a_\alpha(x,\theta)\equiv a_\alpha(x,\theta,0)$. Then if $\mathrm{supp}(\varphi)\subset K$ in X we have:

$$\begin{aligned}|\langle I_\Phi(a),\varphi\rangle| &\leq \sum_{|\alpha|\leq M}\int_{\mathbb{R}^k}\int_X |a_\alpha(x,\theta)|\cdot|\partial^\alpha\varphi|\,\mathrm{d}x\mathrm{d}\theta\\ &\lesssim_K \sum_{|\alpha|\leq M}\sup_K|\partial^\alpha\varphi|\end{aligned}$$

as the $|a_\alpha|$ are integrable by construction. Then since we just need $M > N+k$, Taking $M = N+k+1$ this shows that $\mathrm{ord}(I_\Phi(a)) \leq N+k+1$, since the above is the semi-norm estimate for a distribution. Hence we are done.

□

So now we know what $\int_{\mathbb{R}^k} e^{i\Phi}a(x,\theta)\,\mathrm{d}\theta$ is. But what are the (distributional) derivatives of this? Do they agree with what we might naïvely expect? Well we have:

$$\begin{aligned}
\left\langle \frac{\partial I_\Phi(a)}{\partial x_i}, \varphi \right\rangle &:= -\left\langle I_\Phi(a), \frac{\partial \varphi}{\partial x_i}\right\rangle \\
&:= -\lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^k}\int_X e^{i\Phi}a(x,\theta)\chi(\varepsilon\theta)\frac{\partial\varphi}{\partial x_i}\,\mathrm{d}x\mathrm{d}\theta \\
&= \lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^k}\int_X \frac{\partial}{\partial x_i}\left[e^{i\Phi}a(x,\theta)\chi(\varepsilon\theta)\right]\cdot\varphi\,\mathrm{d}x\mathrm{d}\theta \qquad \text{via integrating by parts} \\
&= \lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^k}\int_X e^{i\Phi}\left[\frac{\partial a}{\partial x_i} + ia\frac{\partial\Phi}{\partial x_i}\right]\chi(\varepsilon\theta)\varphi\,\mathrm{d}x\mathrm{d}\theta \qquad (\dagger) \\
&= \lim_{\varepsilon\downarrow 0}\int_{\mathbb{R}^k}\int_X e^{i\Phi}\tilde{a}(x,\theta)\chi(\varepsilon\theta)\varphi\,\mathrm{d}x\mathrm{d}\theta \\
&=: \langle I_\Phi(\tilde{a}), \varphi\rangle
\end{aligned}$$

where $\tilde{a} := \frac{\partial a}{\partial x_i} + ia\frac{\partial \Phi}{\partial x_i} \in \mathrm{Sym}(X,\mathbb{R}^l;N-1)$ and in ($\dagger$) we are above to pull out $\chi(\varepsilon\theta)$ from the derivative since it does not depend on x. So hence we see $\frac{\partial I_\Phi(a)}{\partial x_i} = I_\Phi(\tilde{a})$ as distributions, which says:

$$\begin{aligned}
\frac{\partial}{\partial x_i}\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\,\mathrm{d}\theta &= \int_{\mathbb{R}^k} e^{i\Phi}\left[\frac{\partial a}{\partial x_i} + ia\frac{\partial \Phi}{\partial x_i}\right]\mathrm{d}\theta \\
&= \int_{\mathbb{R}^k}\frac{\partial}{\partial x_i}\left[e^{i\Phi(x,\theta)}a(x,\theta)\right]\mathrm{d}\theta
\end{aligned}$$

i.e. we can just differentiate under the integral sign, as we might have naïvely expected! So Physicists were fine all along.

Example 5.1 (An Oscillatory Integral). *Consider*

$$I_\Phi(a) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\theta}\,\mathrm{d}\theta$$

for $(x,\theta)\in\mathbb{R}^n\times\mathbb{R}^n$, i.e. $a\equiv 1$, $\Phi(x,\theta) = x\cdot\theta$. Then by our definition we have:

$$\begin{aligned}
\langle I_\Phi(a),\varphi\rangle &:= \lim_{\varepsilon\to 0}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{ix\cdot\theta}\varphi(x)\chi(\varepsilon\theta)\,\mathrm{d}x\mathrm{d}\theta \\
&= \lim_{\varepsilon\to 0}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{ix\cdot(\theta/\varepsilon)}\varphi(x)\varepsilon^{-n}\chi(\theta)\,\mathrm{d}\theta\mathrm{d}x \qquad \textit{(setting } \theta' = \varepsilon\theta\textit{)} \\
&= \lim_{\varepsilon\to 0}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{ix\cdot\theta}\varphi(\varepsilon x)\chi(\theta)\,\mathrm{d}\theta\mathrm{d}x \qquad \textit{(setting } x' = x/\varepsilon\textit{)}
\end{aligned}$$

$$= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^n} \varphi(\varepsilon x)\left[(2\pi)^{-n}\hat{\chi}(-x)\right] \, \mathrm{d}x \qquad \textit{by definition of the Fourier transform}$$
$$= \varphi(0)\chi(0) \qquad \textit{by the inverse Fourier transform}$$
$$= \varphi(0)$$

since $1 = \chi(0) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{\chi}(-x)\,\mathrm{d}x$. *Hence:*

$$\langle I_\Phi(a), \varphi\rangle = \varphi(0) = \langle \delta_0, \varphi\rangle$$

i.e.

$$\delta_0 = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\theta}\,\mathrm{d}\theta.$$

5.1. **Singular Supports.**

It is natural to ask when $I_\Phi(a)$ coincides with a smooth function. We use the idea of **singular supports**:

Definition 5.3. *For a distribution* $u \in \mathscr{D}'(X)$*, the* ***singular support*** *of* u*, denoted* $\operatorname{sing\,supp}(u)$ *is the complement of the largest open set on which* u *can be identified with a smooth function.*

Example 5.2. *Consider* $\delta_0 \in \mathscr{D}'(\mathbb{R}^n)$*. Then for* $\varphi \in \mathscr{D}(\mathbb{R}^n\backslash\{0\})$ *we have* $\langle \delta_0, \varphi\rangle = 0$*, and so as the constant function* 0 *is a smooth function we see that* $\delta_0 \equiv 0$ *on* $\mathbb{R}^n\backslash\{0\}$*. Hence* $\operatorname{sing\,supp}(\delta_0) \subset \{0\}$.

By the stationary phase lemma (Lemma 5.1), we expect $I_\Phi(a)$ to be better at $x_0 \in X$ for which: $\nabla_\theta\Phi(x_0,\theta) \neq 0$ for any $\theta \in \mathbb{R}^k\backslash\{0\}$. The reason we are not interested in $\theta = 0$ is because we are only interested in the large θ behaviour (as we saw before, by splitting the integral into two parts with $\rho(\theta)$ and $1-\rho(\theta)$). This is seen by the following lemma.

Lemma 5.5. *For* $\rho \in \mathscr{D}(\mathbb{R}^k)$*, the function*

$$x \longmapsto \int_{\mathbb{R}^k} e^{i\Phi(x,\theta)}a(x,\theta)\rho(\theta)\,\mathrm{d}\theta$$

is smooth.

Proof. This follows from the definitions and the dominated convergence theorem.

□

Equipped with this we can always write:

$$I_\Phi(a) = \underbrace{\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)} a(x,\theta)\rho(\theta)\, \mathrm{d}\theta}_{\text{smooth by Lemma 5.5}} + \int_{\mathbb{R}^k} e^{i\Phi(x,\theta)} \underbrace{a(x,\theta)(1-\rho(\theta))}_{=:\tilde{a}(x,\theta)}\, \mathrm{d}\theta.$$

So if we choose $\rho \in \mathscr{D}(\mathbb{R}^k)$ with $\rho \equiv 1$ on $|\theta| < 1$, then $\tilde{a} \in \mathrm{Sym}(X, \mathbb{R}^k; N)$ if $a \in \mathrm{Sym}(X, \mathbb{R}^k; N)$ and $\tilde{a}(x,\theta) = 0$ on $|\theta| < 1$.

So if there is any part of the oscillatory integral which is not smooth (i.e. truly a distribution), then it comes from the second part of the integral above, which vanishes on a neighbourhood of 0. So we can assume wlog a is 0 near 0.

So just like our intuition above about when oscillatory integrals are nice, our main result is:

Theorem 5.2. *For an oscillatory integral $I_\Phi(a)$ we have:*

$$\text{sing supp}(I_\Phi(a)) \subset \{x : \nabla_\theta \Phi(x,\theta) = 0 \textit{ for some } \theta \in \mathbb{R}^k \backslash \{0\}\}.$$

Remark: Thus if $\nabla_\theta \Phi(x,\theta) \neq 0$ for any $\theta \in \mathbb{R}^k\backslash\{0\}$, then $I_\Phi(a)$ is smooth about x.

Proof. We will prove the converse (i.e. the opposite inclusion for the complements). Wlog as above we can assume that $a = 0$ on $|\theta| < 1$.

Suppose that $x_0 \in X$ has $\nabla_\theta \Phi(x_0, \theta) \neq 0$ for all $\theta \in \mathbb{R}^k\backslash\{0\}$. Now since $\Phi(x,\theta)$ is homogeneous of degree 1 in θ, we see that $|\nabla_\theta \Phi(x,\theta)|$ is homogeneous of degree 0 in θ. So $|\nabla_\theta \Phi(x,\theta)|$ is determined by its values on any sphere $|\theta| = R$, $R > 0$. So by compactness $|\nabla_\theta \Phi(x,\theta)|$ is bounded below on $|\theta| = R$ (for each x fixed), and so

$$|\nabla_\theta \Phi(x_0, \theta)| \gtrsim 1$$

for all $\theta \in \mathbb{R}^k\backslash\{0\}$. Then by continuity the same estimate holds for $x \in B_\varepsilon(x_0)$ for some $\varepsilon > 0$ small.

Now fix $\psi \in \mathscr{D}(X)$ with $\text{supp}(\psi) \subset B_\varepsilon(x_0)$. Then define:

$$L^* := -\frac{i \sum_{j=1}^k \frac{\partial \Phi}{\partial \theta_j} \cdot \frac{\partial}{\partial \theta_j}}{|\nabla_\theta \Phi|^2}$$

which is similar to the L's we used before in Lemma 5.3 (except we have removed an x—derivatives as we don't want them to hit ψ as we shall see).

Then we know L^* is well-defined on $\text{supp}(\tilde{a})$, where $\tilde{a}(x,\theta) = a(x,\theta)\psi(x)$. Now,

$$\psi I_\Phi(a) = \int_{\mathbb{R}^k} e^{i\Phi(x,\theta)} \underbrace{a(x,\theta)\psi(x)}_{=\tilde{a}}\, \mathrm{d}\theta.$$

So since $\nabla_\theta \Phi$ is homogeneous of degree 0 in θ, since $a \in \text{Sym}(X, \mathbb{R}^k; N)$ we have $L(a) \in \text{Sym}(X, \mathbb{R}^k; N-1)$, and so by definition,

$$\begin{aligned}
\langle \psi I_\Phi(a), \varphi \rangle &:= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^k} \int_X e^{i\Phi(x,\theta)} \tilde{a}(x,\theta)\chi(\varepsilon\theta)\varphi(x)\ \mathrm{d}x\mathrm{d}\theta \\
&= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^k} \int_X e^{i\Phi(x,\theta)} L^{(M)}[\tilde{a}(x,\theta)\chi(\varepsilon\theta)\varphi(x)]\ \mathrm{d}x\mathrm{d}\theta \qquad \text{integrating by parts and } L^*[e^{i\Phi}] = e^{i\Phi} \\
&= \lim_{\varepsilon \to 0} \int_{\mathbb{R}^k} \int_X e^{i\Phi(x,\theta)} L^{(M)}[\tilde{a}(x,\theta)\chi(\varepsilon\theta)]\,\varphi(x)\ \mathrm{d}x\mathrm{d}\theta
\end{aligned}$$

where in the last line we have used that L has no x—derivatives and so we can pull φ out of the L-term. Now choose $M > 0$ sufficiently large so that $L^{(M)}[\tilde{a}(x,\theta)]$ is a symbol of sufficiently negative degree so that this integral is absolutely integrable. So then we can exchange the order of integration and so we get:

$$\begin{aligned}
\langle \psi I_\Phi(a), \varphi \rangle &= \int_X \left(\int_{\mathbb{R}^k} e^{i\Phi(x,\theta)} L^{(M)}[\tilde{a}(x,\theta)]\ \mathrm{d}\theta \right) \varphi(x)\ \mathrm{d}x \\
&= \langle (\cdots), \varphi \rangle
\end{aligned}$$

and so we see by the arbitrariness of φ,

$$\psi I_\Phi(a) = \int_{\mathbb{R}^k} e^{i\Phi(x,\theta)} L^{(M)}[\tilde{a}(x,\theta)]\ \mathrm{d}\theta.$$

Now since this was true for M arbitrary (sufficiently large) we can differentiate under the integral and conclude that $I_\Phi(a)$ is smooth on a neighbourhood of x.

□

Example 5.3. *By previous calculations, we know that if* $\delta_0(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{ix\cdot\theta}\ \mathrm{d}\theta$ *has:*

$$D^\alpha \delta_0(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^k} \theta^\alpha e^{ix\cdot\theta}\ \mathrm{d}\theta.$$

So applying Theorem 5.2 we see:

$$\text{sing supp}(D^\alpha \delta_0) \subset \{x : \underbrace{\nabla_\theta(x\cdot\theta)}_{=x} = 0 \ \textit{ for some } \theta \in \mathbb{R}^k \backslash \{0\}\} = \{x = 0\}.$$

What else can we do with these oscillatory integrals? To finish the course we shall show how they can be used to solve PDEs:

Example 5.4. *Consider the PDE:*

$$\begin{cases} \frac{\partial u}{\partial t} + c \cdot \nabla_y u = 0 & \textit{for } (y,t) \in \mathbb{R}^n \times (0,\infty) \\ u(x,0) = \delta_0(x) \end{cases}$$

for some constant c. Then we guess, via taking the Fourier transform and then solving the equation and taking the inverse Fourier transform, that the solution is:

$$u(y,t) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\theta\cdot(y-ct)}\,\mathrm{d}\theta.$$

Then in the usual notation here we have $X = \mathbb{R}^n \times [0,\infty)$, $x = (y,t)$. So,

$$\Phi(x,\theta) \equiv \Phi(y,t,\theta) = \theta\cdot(y-ct).$$

Then we can check that

$$\frac{\partial u}{\partial t} + c\cdot\nabla_y u = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\theta\cdot(y-ct)}[-c\cdot\theta + c\cdot\theta]\,\mathrm{d}\theta = 0$$

and so this solves the equation, and formally,

$$u(y,0) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{iy\cdot\theta}\,\mathrm{d}\theta = \delta_0(y)$$

i.e. $u(y,t)\to\delta_0(y)$ as $t\to 0$ in $\mathscr{D}'(\mathbb{R}^n)$. Then by Theorem 5.2 we have:

$$\text{sing supp}(u) \subset \{(y,t) : \nabla_\theta(\theta\cdot(y-ct)) = 0 \text{ for some } \theta\in\mathbb{R}^k\backslash\{0\}\} = \{(y,t) : y = ct\}$$

and so we see that the singularity propagates with time! This is perhaps not unexpected as it is a wave, but it is nice to see this mathematically!

End of Lecture Course