Introduction to Microlocal Analysis

A. Uribe[1]

[1]Mathematics Department, University of Michigan, Ann Arbor, Michigan 48109.

Contents

Chapter 1

Fundamentals of Fourier Analysis

1.1 The Fourier transform, part I

We begin by defining the FT on $C_o^\infty(\mathbb{R}^n)$, the space of smooth compactly-supported functions on $\mathbb{R}^n$:

Definition 1.1.1 *If $f \in C_o^\infty(\mathbb{R}^n)$, we define:*

$$\hat{f}(\xi) = \int e^{-ix\cdot\xi} f(x)\, dx.$$

The definition makes sense for $f \in L^1$, and:

Lemma 1.1.2 *If $f \in L^1$ then $\hat{f}$ is a continuous function such that* $\forall \xi$

$$|\hat{f}(\xi)| \leq \|f\|_{L^1}. \tag{1.1}$$

PROOF The bound above is obvious, and continuity is justified by the dominated convergence theorem.

□

Notation:

$$D_j = \frac{1}{i}\frac{\partial}{\partial x_j}, \quad D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}$$

where $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index. We also let $\xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$.

Some basic properties of the FT:

Lemma 1.1.3 $\forall f, g \in C_o^\infty(\mathbb{R}^n)$ *and all multi-indices α,*

1. $\widehat{D^\alpha f}(\xi) = \xi^\alpha \hat{f}(\xi)$,

2. $\widehat{x_j f}(\xi) = i \dfrac{\partial}{\partial \xi_j} \hat{f}$,

3. $\displaystyle\int \hat{f}(y)\, g(y)\, dy = \int f(x)\, \hat{g}(x)\, dx.$

The proof is very easy, just integrate by parts for 1 and 2, and apply Fubini to get 3.

Corollary 1.1.4 *$\forall f \in C_o^\infty(\mathbb{R}^n)$ one has:*

$$\forall\, \alpha,\, \beta \quad \sup |\xi^\alpha\, D_\xi^\beta \hat{f}(\xi)| < \infty. \tag{1.2}$$

To prove this apply repeatedly the first two items of Lemma 1.1.3 and (1.1).

The previous result motivates the following definition:

Definition 1.1.5 *The space of Schwartz functions on $\mathbb{R}^n$, denoted $\mathcal{S}(\mathbb{R}^n)$, consists of all smooth functions on $\mathbb{R}^n$ such that*

$$\forall\, \alpha,\, \beta \quad \|f\|_{\alpha,\beta} := \sup |x^\alpha\, D_x^\beta f(\xi)| < \infty. \tag{1.3}$$

Obviously $C_o^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, but for example a function of the form

$$f(x) \;=\; p(x)\, e^{-|x|^2/2}$$

where p is a polynomial is in $\mathcal{S}(\mathbb{R}^n)$ but not in $C_o^\infty(\mathbb{R}^n)$.

With this terminology Corollary 1.1.4 states that the Fourier transform maps $C_o^\infty(\mathbb{R}^n)$ into $\mathcal{S}(\mathbb{R}^n)$. The same argument also shows that the FT maps $\mathcal{S}(\mathbb{R}^n)$ into itself. We will soon prove the following Theorem:

Theorem 1.1.6 *The Fourier transform is a bijection of $\mathcal{S}(\mathbb{R}^n)$ onto itself; its inverse is*

$$\check{g}(x) \;=\; \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi}\, g(\xi)\, d\xi.$$

There are several different proofs of this Theorem; the one we will give exploits part (3) of Lemma 1.1.3 with g a Gaussian. We first need to compute the Fourier transform of a Gaussian. We will do so in slightly greater generality than needed, for future reference.

Proposition 1.1.7 *Let A be a symmetric positive-definite $n \times n$ matrix, and let*

$$g_A(x) \;=\; e^{-x\cdot Ax/2}.$$

(Think of x as a column vector, so that $x \mapsto x \cdot Ax$ is the quadratic form defined by A.) Then the FT of g_A is:

$$\widehat{g_A}(\xi) \;=\; \frac{(2\pi)^{n/2}}{\sqrt{\det A}}\, g_{A^{-1}}(\xi).$$

PROOF Without loss of generality we can assume A to be diagonal, and then it suffices to prove the result for $n = 1$, that is for

$$g_a(x) = e^{-ax^2/2}, \quad a > 0.$$

(EXERCISE: Verify this claim.) Notice that g_a satisfies the equation: $g_a' = -ax\, g_a$. Dividing by i and applying the FT to both sides we get (using Lemma 1.1.3)

$$\xi\, \widehat{g_a}(\xi) \;=\; -a\,(\hat{g}_a)'(\xi).$$

Any solution of this equation is of the form $\widehat{g_a}(\xi) = C^{-\xi^2/(2a)}$, where $C = \widehat{g_a}(0) = \int g_a(x)dx$. A well-known calculation gives $\int e^{-ax^2/2}\, dx \;=\; \sqrt{2\pi/a}$, and the result follows.

□

PROOF OF THEOREM 1.1.6. Apply part 3 of Lemma 1.1.3 with $g \;=\; g_{aI} \;=\; \exp(-a|x|^2/2)$, to get:

$$\frac{(2\pi)^{n/2}}{a^{n/2}} \int f(x)\, e^{-|x|^2/2a}\, dx \;=\; \int \hat{f}(y)\, e^{-a|y|^2/2}\, dy\,. \tag{1.4}$$

We now let $a \to 0$. On the RHS the integrand tends to $\hat{f}$, and by Lebesgue dominated the integral converges to $\int \hat{f}(y)\, dy$. To see what happens to the LHS in the limit, make the change of variables $u = \frac{1}{\sqrt{a}}x$ to get that the LHS equals

$$(2\pi)^{n/2} \int f(\sqrt{a}u)\, e^{-|u|^2/2}\, du.$$

As $a \to 0$ once again Lebesgue dominated yields that the limit is $(2\pi)^n\, f(0)$. Thus we have shown that

$$f(0) \;=\; \frac{1}{(2\pi)^n} \int \hat{f}(\xi)\, d\xi. \tag{1.5}$$

This is the inversion formula at $x = 0$. To obtain it in general proceed as follows. Given $w \in \mathbb{R}^n$ let

$$f_w(x) \;=\; f(x + w). \tag{1.6}$$

Obviously $f_w \in \mathcal{S}(\mathbb{R}^n)$, and therefore we can apply (1.5) to it. The FT of f_w is easily computed (proof left as an exercise) and Theorem 1.1.6 follows:

Lemma 1.1.8 *Let $f \in \mathcal{S}(\mathbb{R}^n)$, $w \in \mathbb{R}^n$ and define f_w by (1.6). Then*

$$\widehat{f_w}(\xi) = e^{iw\cdot\xi}\,\hat{f}(\xi).$$

Corollary 1.1.9 . *Let $f, g, \in \mathcal{S}(\mathbb{R}^n)$, and let $< f, g >= \int f(x)\overline{g(x)}dx$ denote the L^2 inner product. Then (Parseval's identity)*

$$< f, g >= \frac{1}{(2\pi)^n} < \hat{f}, \hat{g} > .$$

In particular (Plancherel's identity)

$$\|f\|_2 = \frac{1}{(2\pi)^{n/2}} \|\hat{f}\|_2.$$

Proof Set

$$h := \frac{1}{(2\pi)^n}\,\overline{\hat{g}}.$$

so that the RHS of Parseval's identity is $\int \hat{f}(x)h(x)dx$. Now use part 3 of Lemma 1.1.3, applied to f and h. The Fourier inversion formula implies that $\overline{\hat{h}} = g$.

□

1.2 Convolution

Definition 1.2.1 *The convolution of two functions f, g on $\mathbb{R}^n$ is the function*

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

defined whenever the RHS converges absolutely.

Clearly the convolution is well-defined e.g. in the following cases:

- If one function is in $\mathcal{S}(\mathbb{R}^n)$ and the other is bounded by a polynomial.
- If $f \in L^p$ and $g \in L^q$ where p, q are conjugate exponents. In this case, by Hölder's inequality, $f * g$ is in L^∞, and in fact

$$f \in L^p \text{ and } g \in L^q \Longrightarrow \|f * g\|_\infty \le \|f\|_p\,\|g\|_q. \tag{1.7}$$

It is easy to show that the operation of convolution is commutative and associative. Notice that if both f and g have compact support then $f * g$ also has compact support. (Exercise: verify these facts.)

1.2.1 Basic L^p estimates

We now turn to some basic L^p estimates of convolutions.

Proposition 1.2.2 *If $f,\ g \in L^1$ then $f * g \in L^1$ and*

$$\|f * g\|_1 \leq \|f\|_1 \, \|g\|_1 .$$

PROOF The double integral

$$\int\int |f(x-y)\, g(y)|\, d(x,y)$$

exists, (a priori it may be infinite), and by Fubini's theorem is equal to either of the iterated integrals. Integrating first with respect to x we find that in fact it's finite:

$$\int\int |f(x-y)\, g(y)|\, d(x,y) \ = \ \|f\|_1 \, \|g\|_1 . \tag{1.8}$$

In other words, the function $f(x-y)\, g(y)$ is in $L^1(\mathbb{R}^n \times \mathbb{R}^n)$. Therefore, by Fubini's theorem again the integral

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

exists for almost all x and defines an integrable function; this function is of course $f * g$. The bound of its L^1 norm is given by (1.8).

□

Proposition 1.2.3 *If $f \in L^p$ and $g \in L^1$, then $f * g \in L^p$ and*

$$\|f * g\|_p \leq \|f\|_p \, \|g\|_1 .$$

PROOF We first must show that the convolution exists, i.e. that $|g(y)|\, |f(x-y)|$ is integrable in y for almost all x. Since $|g|$ and $|f|^p$ are in L^1, by the previous proposition $|g(y)|\, |f(x-y)|^p$ is integrable in y for almost all x. Therefore:

$$|g(y)|^{1/p}\, |f(x-y)| \text{ is in } L^p(\mathbb{R}^n_y) \text{ for almost all } x.$$

Let q be the exponent conjugate to p and notice that

$$|g(y)|\, |f(x-y)| \ = \ |g(y)|^{1/p}\, |f(x-y)|\, |g(y)|^{1/q} .$$

Since $|g(y)|^{1/q} \in L^q(\mathbb{R}^n_y)$, by Hölder's inequality we get that for almost all x

$$\int |g(y)|\, |f(x-y)|\, dy \leq \Big(\int |g(y)|\, |f(x-y)|^p\, dy\Big)^{1/p} \|g\|_1^{1/q} . \tag{1.9}$$

This shows that the convolution $f * g$ is defined for almost all x and for those x $|f * g(x)|$ is bounded by the right-hand side of (1.9). Therefore

$$|f * g(x)|^p \leq \Big(\int |g(y)|\, |f(x-y)|^p \, dy\Big)\, \|g\|_1^{p/q}.$$

Now integrate the above with respect to x; by Fubini's theorem we can interchange the order of integration and get:

$$\|f * g\|_p^p \leq \|f\|_p^p \, \|g\|_1^{1+\frac{p}{q}}.$$

Now just take pth roots.

□

1.2.2 Some approximation theorems

One of the most common applications of convolutions is to the construction of smooth approximations to functions. We'll develop this idea next. We first observe that convolutions smooth things out.

Lemma 1.2.4 *If $u \in C_o^\infty(\mathbb{R}^n)$ and f is locally integrable (denoted: $f \in L^1_{\rm loc}$) then $f * u \in C^\infty(\mathbb{R}^n)$, and in fact*

$$D^\alpha(f * u) \;=\; f * (D^\alpha u).$$

Proof Both continuity of the convolution and differentiation under the integral sign defining $f * u$ can be justified by the dominated convergence theorem.

□

Theorem 1.2.5 *Let $u \in C^k$ for some $0 \leq k < \infty$, $K \subset \mathbb{R}^n$ a compact set and let $\epsilon > 0$. Then there exists $\delta > 0$ such that for any $\phi \in C_o^\infty(\mathbb{R}^n)$ with $\phi \geq 0$, $\int \phi(x)\, dx = 1$, and such that the support of ϕ is contained in the ball $\{\, x\,;\, |x| \leq \delta \,\}$, one has: $\forall \alpha$ such that $|\alpha| \leq k$*

$$\sup_{x \in K} |D^\alpha(u * \phi)(x) - D^\alpha u(x)| < \epsilon.$$

Proof It suffices to prove the case $\alpha = 0$. Using that $\phi \geq 0$ and that $\int \phi(x)\, dx = 1$, get:

$$|u_\phi(x) - u(x)| = \Big|\int \Big(u(x-y) - u(x)\Big)\, \phi(y)\, dy\Big| \leq \sup_{|y| \leq \delta} |u(x-y) - u(x)|.$$

Now just use uniform continuity of $u|_K$.

□

Notice that this gives an explicit approximation of a continuous, compactly supported function by smooth functions of compact support.

We'll next review the following fact:

Proposition 1.2.6 *$C_o^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, $1 \le p < \infty$.*

In view of the fact that C_0^0 is dense in $C_o^\infty(\mathbb{R}^n)$ in the sup norm and hence in all L^p norms, this proposition follows from the following:

Lemma 1.2.7 *The space C_0^0 of continuous functions on $\mathbb{R}^n$ of compact support is dense in L^p, $1 \le p < \infty$.*

PROOF [Sketch; see e.g. TORCHINSKY, A. *Real Variables*, pg 133 for the $p = 1$ case; the case $p > 1$ is similar.] We begin with the basic fact from integration theory that the space of simple functions is dense in L^p. (Indeed it suffices to approximate non-negative functions $f \in L^p$, and then one knows that there exists an increasing sequence of simple functions approximating f from below.) Thus it suffices to show that the characteristic function of a bounded measurable set, E, can be approximated in the L^p norm by an element of C_0^0. By regularity of Lebesgue measure, there exists an open set U containing E whose measure is very close to that of E, so WLOG can assume E is open. At this point E can be unbounded, but it has finite measure.

Next, write $E = \bigcup_{k=1}^\infty I_k$, where the I_k are compact n-intervals (Cartesian products of n closed intervals). By the finiteness of the measure of E, we can throw away all but finitely-many of the I_k and make a small error in measure, that is WLOG $E = \bigcup_{k=1}^N I_k$ for some N. All this reduces the problem to proving the following: *The characteristic function of an n-interval can be approximated in L^p norm by a continuous function.* But this is obvious.

□

Proposition 1.2.6 can now be used to prove some further results on convolutions:

Proposition 1.2.8 *Let $u \in L^p$ and let $\epsilon > 0$. Then there exists $\delta > 0$ such that for any $\phi \in C_o^\infty(\mathbb{R}^n)$ with $\phi \ge 0$, $\int \phi(x)\,dx = 1$, and such that the support of ϕ is contained in the ball $\{\, x\,;\, |x| \le \delta \,\}$, one has:*

$$\|u * \phi - u\|_p < \epsilon.$$

PROOF Let $f \in C_o^\infty(\mathbb{R}^n)$ such that $\|f - u\|_p < \epsilon/3$. To find δ, apply Theorem 1.2.5 to the continuous function f and to a suitable fraction of ϵ so that $\|f_\phi - f\|_p < \epsilon/3$. Then:

$$\|u_\phi - u\|_p \le \|u_\phi - f_\phi\|_p + \|f_\phi - f\|_p + \|f - u\|_p$$

Notice that $u_\phi - f_\phi = (u - f)_\phi$, and by Proposition 1.2.3 we get:

$$\|u_\phi - f_\phi\|_p \le \|u - f\|_p = \epsilon/3,$$

since $\|\phi\|_1 = 1$. Therefore $\|u_\phi - u\|_p \le \epsilon$.

□

Exercise 1.2.1 Prove the following: If $f \in L^p$ and $g \in L^q$ where p and q are conjugate exponents, then $f * g$ is bounded and uniformly continuous.

Exercise 1.2.2 Let $n = 1$ and

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

(a) Prove that $\forall f \in C_0^\infty(\mathbb{R})$ the function $f * H$ is a primitive (anti-derivative) of f.
(b) Generalize: For each positive integer k find a function H_k such that the k-th derivative of $f * H_k$ is f, for each $f \in C_o^\infty(\mathbb{R}^n)$.

1.3 The Fourier transform in L^2 and L^1.

1.3.1 L^2

Parseval's identity allows to extend the FT from $\mathcal{S}(\mathbb{R}^n)$ to L^2 by a density argument. This is based on the following:

Proposition 1.3.1 *Let $f \in L^2$ and (g_k) a sequence in $\mathcal{S}(\mathbb{R}^n)$ converging to f in L^2. Then:*

1. *The sequence $(\widehat{g_k})$ converges in L^2.*
2. *The limit of $(\widehat{g_k})$ in L^2 only depends on f.*
3. *If $f \in L^2 \cap L^1$, $\lim\limits_{k\to\infty} \widehat{g_k} = \hat{f}$.*

Proof Since (g_k) converges, it is a Cauchy sequence in L^2. By Parseval's identity and the linearity of the FT, $(\widehat{g_k})$ is likewise Cauchy in L^2. This proves (1). To prove (2), suppose that (g_k) a sequence in $\mathcal{S}(\mathbb{R}^n)$ converging to 0 in L^2. Again by Parseval we get that $(\|\widehat{g_k}\|_2)$ converges to zero, which implies (2).

Let us now prove (3). Assume $f \in L^1 \cap L^2$, and let $h \in \mathcal{S}$. Observe that the relation

$$\int \hat{f}(y)\, h(y)\, dy = \int f(x)\, \hat{h}(x)\, dx \tag{1.10}$$

remains valid in this context, since the function $e^{-ix\cdot y} f(x) h(y)$ is integrable (this is true for all $f, h \in L^1$; just apply Fubini). Therefore:

$$\left|\int \widehat{g_k}\, h\, dx - \int \hat{f}\, h\, dx\right| = \left|\int (g_k - f)\, \hat{h}\, dx\right|.$$

Since $g_k - f \in L^2$ (by the assumption that $f \in L^2$) then we can apply the Cauchy-Schwartz inequality:

$$\Big| \int \widehat{g_k}\, h\, dx - \int \hat{f}\, h\, dx \Big| \le \|g_k - f\|_2\, \|\hat{h}\|_2$$

Therefore $\forall h \in \mathcal{S}$

$$\lim_{k\to\infty} \int \widehat{g_k}\, h\, dx \;=\; \int \hat{f}\, h\, dx.$$

Now if we denote temporarily by $\tilde{f}$ the L^2 limit of $(\widehat{g_k})$, then we obviously we also have that $\forall h \in \mathcal{S}$ the limit above is $\int \tilde{f}\, h\, dx$. Therefore:

$$\forall h \in \mathcal{S}(\mathbb{R}^n) \qquad \int (\hat{f} - \tilde{f})\, h\, dx \;=\; 0.$$

From this it easily follows that $\tilde{f} = \hat{f}$.

□

The previous results justify the following:

Definition 1.3.2 *Let $f \in L^2$ and (g_k) a sequence in $\mathcal{S}(\mathbb{R}^n)$ converging to f in L^2. (We have seen that such sequences do exist.) We define:*

$$\hat{f} \;:=\; \lim_{k\to\infty} \widehat{g_k}.$$

Exercise 1.3.1 Let $f \in L^2(\mathbb{R})$. Prove that

$$\hat{f}(\xi) \;=\; \lim_{(a,b)\to\infty} \int_{-a}^{b} e^{-ix\cdot\xi}\, f(x)\, dx.$$

Exercise 1.3.2 Extend Parseval's identity to L^2.

Exercise 1.3.3 Extend the inversion formula to L^2.

1.3.2 L^1

Proposition 1.3.3 *For $f, g \in L^1$:*

1. $\lim_{|\xi|\to\infty} \hat{f}(\xi) \;=\; 0,$

2. $(f * g)\hat{} = \hat{f}\,\hat{g}.$

Proof To prove (1), let (f_k) be a sequence in $C_o^\infty(\mathbb{R}^n)$ converging to f in L^1, say $\|f_k - f\|_1 \le 1/k$. Then $\forall \xi$

$$|\widehat{f_k}(\xi) - \hat{f}(\xi)| \le \|f_k - f\|_1 \le \frac{1}{k}.$$

Therefore $|\hat{f}(\xi)| \le |\widehat{f_k}(\xi)| + 1/k$. Now take $\limsup$ as $|\xi| \to \infty$ and use the fact that $\widehat{f_k}$ is rapidly decreasing to get: $\forall k$

$$\limsup_{|\xi|\to\infty} |\hat{f}(\xi)| \le \frac{1}{k}.$$

To prove (2) just use Fubini.

□

We next show how to recover $f \in L^1$ from its Fourier transform.

Proposition 1.3.4 *Let $f \in L^1$ and $\phi_\delta \in C_o^\infty(\mathbb{R}^n)$ be an approximate identity, that is: $0 \le \phi_\delta$, $\displaystyle\int \phi_\delta \, dx = 1$ and support of ϕ_δ in $\{x\,;\,|x| \le \delta\}$. Then*

$$f \;=\; \lim_{\delta\to 0}\Big(\hat{\phi}_\delta\, \hat{f}\Big)^{\check{}}$$

in the L^1 norm.

Proof First notice that $\hat{\phi}_\delta\, \hat{f}$ is continuous and decreases rapidly at infinity, so it is integrable. Therefore $(\hat{\phi}_\delta\, \hat{f})^{\check{}}$ exists in the ordinary sense. Next we will check that for each $\delta > 0$

$$\Big(\hat{\phi}_\delta\, \hat{f}\Big)^{\check{}} \;=\; \phi_\delta * f. \tag{1.11}$$

This is again an application of Fubini:

$$(2\pi)^n \Big(\hat{\phi}_\delta\, \hat{f}\Big)^{\check{}}(x) \;=\; \int e^{ix\cdot\xi}\, \hat{\phi}_\delta(\xi) \Big[\int e^{-iy\cdot\xi}\, f(y)\, dy\Big]\, d\xi.$$

Now just use the fact that $\hat{\phi}_\delta(\xi)\, f(x)$ is integrable in (x,ξ) space to interchange the order of integration and apply the Fourier inversion formula to ϕ_δ. This proves (1.11). To finish the proof let $\delta \to 0$ and quote Proposition 1.2.8.

□

Proposition 1.3.5 *Let $f \in L^1$ be such that $\hat{f} \in L^1$ (therefore $(\hat{f})^{\check{}}$ is defined). Then f is equal to a continuous function almost everywhere and*

$$f \;=\; (\hat{f})^{\check{}} \quad \textit{almost everywhere.}$$

PROOF Proceeding as in the proof of the previous proposition but now using that $\hat{f}(\xi)\,\phi_\delta(y)$ is integrable, get that $\left(\hat{\phi}_\delta\,\hat{f}\right)\check{} = (\hat{f})\check{} * \phi_\delta$. But this is also equal to $f * \phi_\delta$, by the previous proposition. Therefore

$$\forall \delta > 0 \quad (\hat{f})\check{} * \phi_\delta \;=\; f * \phi_\delta.$$

Notice that $(\hat{f})\check{}$ is a continuous function, so the LHS converges to $(\hat{f})\check{}$ locally uniformly. Finally recall that every sequence that converges in L^1 has a subsequence that converges pointwise a.e. (see Rudin I, Theorem 3.12.)

□

1.4 Tempered distributions

1.4.1 Basic definitions

Consider, as motivation, the formula $(f * H)' \;=\; f$ where H is the Heaviside function (see Exercise 1.2.2). Recalling that for Schwartz functions f, g we have $(f * g)' = f * (g')$, proceeding *purely formally* we may write

$$\forall f \in \mathcal{S}(\mathbb{R}^n) \quad f \;=\; f * (\delta), \quad \delta \;=\; H'. \tag{1.12}$$

It is easy to see that one cannot make sense of δ as an ordinary function: It would have to be zero for all $x \neq 0$, therefore zero almost everywhere, and yet $\int_\infty^1 = H(1) = 1$! However, let us write (formally of course) (1.12) in more detail:

$$\forall f \in \mathcal{S}(\mathbb{R}^n),\; a \in \mathbb{R}^n \quad f(a) \;=\; \int f(a-x)\,\delta(x)\,dx. \tag{1.13}$$

This property would follow easily from the simpler property that

$$\forall f \in \mathcal{S}(\mathbb{R}^n) \quad f(0) \;=\; \int f(x)\,\delta(x)\,dx. \tag{1.14}$$

This is the one property that we require of δ. We therefore turn it into a definition:

Definition 1.4.1 *Dirac's delta function on* $\mathbb{R}^n$ *is the functional*

$$\begin{array}{rcl} \delta : \mathcal{S}(\mathbb{R}^n) & \to & \mathbb{C} \\ f & \mapsto & f(0) \end{array}$$

Of course this definition is useful only because there's a whole theory of functionals $\mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ (that satisfy a mild continuity condition): Laurent Schwartz' theory of (tempered) distributions.

Definition 1.4.2 *For each non-negative integer N, define the following norm on $\mathcal{S}(\mathbb{R}^n)$:*

$$|||f|||_N = \sum_{|\alpha \leq N, \, |\beta| \leq N} \sup_{x \in \mathbb{R}^n} |\, x^\alpha D^\beta f(x) \,| \, .$$

Then, a tempered distribution is a linear functional

$$\begin{aligned} T : \mathcal{S}(\mathbb{R}^n) &\to \quad \mathbb{C} \\ f &\to < T, f > \end{aligned}$$

such that $\exists N, \; C > 0$ *such that*

$$\forall f \in \mathcal{S}(\mathbb{R}^n) \quad |< T, f >| \leq C \, |||f|||_N . \tag{1.15}$$

The set of all tempered distributions will be denoted: $\mathcal{S}'$.

Informally speaking a linear functional on $\mathcal{S}(\mathbb{R}^n)$ is in $\mathcal{S}'$ iff it can be controlled by the sup norm of finitely-many of the derivatives of its argument, weighted by polynomials of bounded degree.

We will next list several examples:

1. Clearly $\delta \in \mathcal{S}'$, as $|f(0)| \leq |||f|||_0$.

2. Let $u \in L^p$, $1 \leq p \leq \infty$. We can then define a functional T_u by the rule

   $$< T_u, f > = \int u(x) \, f(x) \, dx .$$

   Let us prove that this is a distribution. Let q be the conjugate exponent, let N be an integer such that $(1 + |x|^2)^{-N} \in L^q$ and write

   $$\int u(x) \, f(x) \, dx \; = \int f(x)(1 + |x|^2)^N \, \frac{u(x)}{(1 + |x|^2)^N} \, dx .$$

   Since, by Hölder's inequality $u(x) \, (1 + |x|^2)^{-N} \in L^1$,

   $$|< T_u, f >| \leq \Big(\int \Big| \frac{u(x)}{(1 + |x|^2)^N} \Big| \, dx \Big) \sup_{x \in \mathbb{R}^n} \; |f(x)(1 + |x|^2)^N| \, .$$

   This is an estimate of the desired form, (1.15).

3. More generally let $u(x)$ be a continuous function on $\mathbb{R}^n$ with at most polynomial growth at infinity (e.g. a polynomial), and define T_u as in the previous example. Then the same trick shows that T_u is a tempered distribution.

4. Here's an example which is a combination of derivatives and delta function ideas: Fix a multi-index α and define T by the recipe

$$< T, f > = \int_{-\infty}^{+\infty} x_1^2 \, (D^\alpha f) \, (x_1, 0, \ldots, 0) \, dx_1 . \tag{1.16}$$

Exercise 1.4.1 Prove that (1.16) is indeed a distribution.

Definition 1.4.3 *It should be clear that in examples 2 and 3 the functional T_u uniquely determines the function u. We will henceforth identify u with T_u, so that we have embeddings, e.g. $L^p \subset \mathcal{S}'$, $\mathcal{S}(\mathbb{R}^n) \subset \mathcal{S}'$.*

1.4.2 Differentiation, product by functions.

Our next concern will be to extend to $\mathcal{S}'$ some of the basic operations on functions. This will be done according to the following general principle. Let $A : \mathcal{S} \to \mathcal{S}$ be a linear operator, for example $A = \partial/\partial x_1$, and let $A^\dagger$ be the transpose of A, that is the operator defined by:

$$\forall f, \; g \quad \int A(f) \, g \, dx \; = \; \int f \, A^\dagger(g) \, dx \, .$$

Assume that $A^\dagger$ exists and that it maps $\mathcal{S}$ into itself. Then: if u_f is the distribution defined by a function $f \in \mathcal{S}$,

$$< u_{A(f)}, g > = \int A(f) \, g \, dx \; = \; \int f \, A^\dagger(g) \, dx \; = < u_f, A^\dagger(g) > \, . \tag{1.17}$$

Now we wish to extend A to arbitrary distributions, u. By "extend" we mean precisely that the definition of $A(u)$ should be such that $\forall f \in \mathcal{S} \; A(u_f) = u_{A(f)}$. But then (1.17) tells us exactly how to define the extension:

Definition 1.4.4 *Let $A : \mathcal{S} \to \mathcal{S}$ be a linear operator and assume that its transpose $A^\dagger$ maps $\mathcal{S}$ into itself. Then we define $\forall u \in \mathcal{S}'$*

$$\begin{array}{rcl} A(u) : \mathcal{S} & \to & \mathbb{C} \\ g & \mapsto & < u, A^\dagger(g) > \end{array}$$

If $A^\dagger$ is continuous in a suitable sense then $A(u)$ will be a tempered distribution; we'll check directly that this is the case in each of the examples below.

Definition 1.4.5 *Let $\phi \in C^\infty(\mathbb{R}^n)$ be such that the multiplication operator $A(f) = \phi f$ maps $\mathcal{S}$ into itself. Then for every $u \in \mathcal{S}'$ we define ϕu according to the general principle above:*

$$\forall g \in \mathcal{S} \qquad < \phi u \, , \, g > = < u \, , \, \phi g > \, .$$

It it is trivial to check that ϕu is indeed a distribution.

Definition 1.4.6 *For each distribution u and each multi-index α, define $D^\alpha u$ according to the general principle above:*

$$\forall g \in \mathcal{S} \qquad < D^\alpha u \,,\, g >= (-1)^{|\alpha|} < u \,,\, D^\alpha g > .$$

Exercise 1.4.2 Check that $D^\alpha u$ is indeed a distribution, and that $D^{\alpha+\beta} u = D^\alpha(D^\beta u)$.

Examples

1. Let $u = \delta$, the delta function at the origin. Then

$$< x_j \delta \,,\, g >=< \delta \,,\, x_j g >= x_j \, g(x)|_{x=0} \;=\; 0,$$

   i.e. $x_j \, \delta = 0$.

2. If $H(x)$ is the Heaviside function ($n = 1$), then

$$< H' \,,\, g >= -\int_0^\infty g'(x)\, dx \;=\; g(0),$$

   that is: $H' = \delta$.

3. For a given $j = 1, \ldots, n$,

$$< \frac{\partial \delta}{\partial x_j} \,,\, g >= \;-< \delta \,,\, \frac{\partial g}{\partial x_j} >= \;-\frac{\partial g}{\partial x_j}(0).$$

4. Observe that

$$< x_j \frac{\partial \delta}{\partial x_j} \,,\, g >= \;-< \delta \,,\, \frac{\partial}{\partial x_j}(x_j \, g) >= \;-g(0),$$

   that is: $x_j \dfrac{\partial \delta}{\partial x_j} = -\delta!$

Generalizations/converses to some the examples above:

Proposition 1.4.7 *Let u be a bounded function that is in $C^1(\mathbb{R} \setminus \{x_0\})$, and let $v = u'$ in the sense of functions (so that $v \in C^0(\mathbb{R} \setminus \{x_0\})$). Assume that v is an integrable function. Then the distributional derivative of u is:*

$$u' \;=\; v + \Big(u(x_0 + 0) - u(x_0 - 0)\Big)\, \delta_{x_0}.$$

PROOF Integrate over $|x - x_0| > \epsilon$ and take the limit as $\epsilon \to 0$...

□

Proposition 1.4.8 *If $u \in \mathcal{S}'$ is such that $\forall j = 1, \ldots, n$ $x_j u = 0$, then u is a constant multiple of δ.*

We need the following:

Lemma 1.4.9 *$\forall f \in \mathcal{S}$, $\exists f_1, \ldots f_n \in \mathcal{S}$ such that*

$$f(x) \;=\; f(0) \,+\, \sum_{j=1}^{n} x_j\, f_j(x).$$

PROOF Define:

$$f_j(x) \;=\; \int_0^1 \frac{\partial}{\partial x_j}\, f(tx_1, \ldots, tx_n)\, dt.$$

The sum on the RHS of the desired identity is then simply

$$\int_0^1 \frac{d}{dt}\Big[f(tx)\Big]\, dt \;=\; f(x) - f(0).$$

□

It should be obvious that Proposition (1.4.8) follows from the lemma.

Proposition 1.4.10 *If $u \in \mathcal{S}'$ is such that $\forall j = 1, \ldots, n$ $\frac{\partial u}{\partial x_j} = 0$, then $u = c$, a constant.*

PROOF [SKETCH] This proof uses the Fourier transform. As we will see, the FT can be extended to $\mathcal{S}'$ according to the general principle, in such a way that the identity $\widehat{D_j u}(\xi) = \xi_j\, \hat{u}$ is preserved. Therefore the assumption on u is that $\forall j$ $\xi_j\, \hat{u} = 0$, and by Proposition 1.4.8 that means that $\hat{u}$ is a multiple of the delta function. As we'll see this means that u is a constant.

□

1.4.3 Some important examples.

$\underline{x_+^a}$ For every $a > -1$, the function

$$x_+^a \;:=\; \begin{cases} x^a & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$$

is locally integrable and has polynomial growth. It therefore defines a tempered distribution. What is its distributional derivative? A guess is that, at least if $a > 1$, one should have:

$$\left(x_+^a\right)' \;=\; a\, x_+^{a-1}. \tag{1.18}$$

This is patently false if $a = 0$ (for then x_+^a is the Heaviside function and as we've seen its derivative is the delta function). For $-1 < a < 0$ the right-hand side is not locally integrable and therefore it does not immediately define a distribution. Here's how to regularize it:

Definition 1.4.11 *If $-2 < b < -1$, define the distribution x_+^b by:*

$$< x_+^b \,, f > := \int_0^\infty \Big(f(x) - f(0)\Big)\, x^b \, dx. \tag{1.19}$$

Observe that $\int_1^\infty x^b \, dx < \infty$ and that the integrand in (1.19) is integrable near $x = 0$ (this follows from Lemma (1.4.9): $f(x) - f(0) = xg(x)$ with $g \in \mathcal{S}$). Therefore the integral is absolutely convergent. Also observe that for functions $f \in \mathcal{S}(\mathbb{R}^n)$ supported away from zero the definition above agrees with the definition of the distribution defined by the function x_+^b.

Lemma 1.4.12 *With the definition above, for every $a > -1$, $a \neq 0$, (1.18) holds, in the sense of distributions.*

1.4.4 Localization, supports.

Although tempered distributions are defined globally as functionals on $\mathcal{S}$, it turns out that they are local (but not punctual!) objects. This is because of the existence of smooth partitions of unity.

Definition 1.4.13 *Let $\Omega \subset \mathbb{R}^n$ be an open set, and $u, v \in \mathcal{S}'$. We'll say that u and v agree on Ω, and write $u|_\Omega = v|_\Omega$ iff for every $f \in C_o^\infty(\mathbb{R}^n)$*

$$\mathrm{supp} f \subset \Omega \quad \Rightarrow \quad < u \,, f > = < v \,, f > .$$

Our immediate goal is to show that a distribution is determined locally:

Theorem 1.4.14 *Let $\{\Omega_j\}$ be an open cover of $\mathbb{R}^n$, and $u, v \in \mathcal{S}'$. Suppose that $\forall j \;\; u|_{\Omega_j} = v|_{\Omega_j}$. Then $u = v$.*

To prove this theorem we need to establish some preliminary results.

Proposition 1.4.15 *$C_o^\infty(\mathbb{R}^n)$ is dense in $\mathcal{S}(\mathbb{R}^n)$, in the following sense: $\forall f \in \mathcal{S}(\mathbb{R}^n) \; \exists \{f_k\} \subset C_o^\infty(\mathbb{R}^n)$ such that*

$$\forall \epsilon > 0, \; N > 0 \; \exists K > 0 \; \text{such that} \quad \forall k > K \quad |||\, f - f_k \,|||_N < \epsilon. \tag{1.20}$$

PROOF First notice that if $g \in \mathcal{S}(\mathbb{R}^n)$ $\forall \alpha,\ \beta$

$$\lim_{R\to\infty} \sup_{x;|x|>R} |x^\alpha D^\beta g(x)| = 0. \tag{1.21}$$

Let $\chi \in C_o^\infty(\mathbb{R}^n)$ be such that $\chi(x) = 1$ for all x such that $|x| \leq 1$, and let $f_k(x) = \chi(\frac{1}{k}x)\, f(x)$. Then $f_k \in C_o^\infty(\mathbb{R}^n)$ and

$$|x| \leq k \quad \Rightarrow \quad f(x) - f_k(x) = 0.$$

Therefore

$$|||\, f - f_k\, |||_N = \sum_{|\alpha|,|\beta|\leq N} \sup_{x;|x|>k} |x^\alpha D^\beta (f - f_k)|.$$

Now just apply (1.21).

$\square$

Corollary 1.4.16 *If $u, v \in \mathcal{S}'$ agree on each $f \in C_o^\infty(\mathbb{R}^n)$, then $u = v$.*

By the density of $C_o^\infty(\mathbb{R}^n)$ in $\mathcal{S}(\mathbb{R}^n)$ we'll only need finite partitions of unit:

Proposition 1.4.17 *Let $\Omega_1, \ldots \Omega_N$ be open sets in $\mathbb{R}^n$, and let $\Omega = \Omega_1 \cup \cdots \cup \Omega_N$. Let $f \in C_0^\infty(\Omega)$. Then $\exists f_j \in C_0^\infty(\Omega_j)$ such that* $f = \sum_{j=1}^{N} f_j$.

PROOF For each $x \in \operatorname{supp} f$, pick a compact neighborhood of x contained in some Ω_j. By compactness of $\operatorname{supp} f$, a finite collection of such neighborhoods cover $\operatorname{supp} f$. For each j, let K_j be the union of such neighborhoods contained in Ω_j. Then:

$$\operatorname{supp} f \subset \bigcup_{j=1}^{N} K_j \quad \text{and} \quad \forall j\ K_j \subset \Omega_j.$$

For each j let $\chi_j \in C_0^\infty(\Omega_j)$ such that $\forall x \in K_j\ \chi_j(x) = 1$, and let $f_1 = f\, \chi_1$ and inductively

$$f_j = f\, \chi_j\, (1 - \chi_1) \cdots (1 - \chi_{j-1}).$$

Now notice:

$$f - \sum_{j=1}^{N} f_j = f \Pi_{j=1}^{N} (1 - \chi_j)$$

which at each point is zero, because either $f(x) = 0$ or some $\psi_j(x) = 1$.

$\square$

1.4.5 Convolution and approximation results.

Definition 1.4.18 *If f is a function on $\mathbb{R}^n$, we define:*

1. $f^\rho(x) := f(-x)$,
2. $\forall y \in \mathbb{R} \quad f^y(x) := f(x-y)$.

Notice that with these notations the convolution of two functions is:

$$f * g(x) = \int f^y(x)\, g(y)\, dy = \int (f^\rho)^x(y)\, g(y)\, dy. \tag{1.22}$$

Definition 1.4.19 *If $u \in \mathcal{S}'$ and $f \in \mathcal{S}$, the convolution $u * f$ is defined to be the function*

$$u * f(x) := < u\,,\, (f^\rho)^x > .$$

Notice that this definition makes sense: For each $x \in \mathbb{R}^n$ the translate f^x is still a Schwartz function.

Proposition 1.4.20 *If $u \in \mathcal{E}'$ and $f \in C_o^\infty(\mathbb{R}^n)$ $u * f$ is a smooth function of compact support.*

Exercise 1.4.3 Prove this proposition.

Proposition 1.4.21 *If $u \in \mathcal{E}'$ and $f, g \in C_o^\infty(\mathbb{R}^n)$, then*

$$< u * f\,,\, g > = < u\,,\, f^\rho * g > . \tag{1.23}$$

Proof It is easy to prove that (1.23) holds if u is smooth: By Fubini both sides equal

$$\int\int u(y)\, f(x-y)\, g(x)\, dy\, dx$$

which is absolutely convergent. In general, (1.23) is the statement

$$\int < u\,,\, (f^\rho)^x > g(x)\, dx = < u\,,\, \int (f^\rho)^x\, g(x)\, dx > \tag{1.24}$$

(use the first of the identities (1.22)), that is what we have to prove is that we can move in the integral sign into the pairing. To this end write the LHS of (1.24) as a limit of Riemann sums:

$$\lim_{k\to\infty} \frac{1}{k^n} \sum_{l \in \mathbb{Z}^n} g(\frac{1}{k}l) < u\,,\, (f^\rho)^{l/k} > = \lim_{k\to\infty} < u\,,\, \frac{1}{k^n} \sum_{l \in \mathbb{Z}^n} g(\frac{1}{k}l)\, (f^\rho)^{l/k} > .$$

It suffices therefore to prove that

$$\frac{1}{k^n} \sum_{l \in \mathbb{Z}^n} g(\frac{1}{k} l)\,(f^\rho)^{l/k} \to \int (f^\rho)^x\, g(x)\, dx$$

in the C^∞ topology.

$\square$

1.5 Fourier transforms of tempered distributions

1.5.1 Definition and first properties

Proposition 1.5.1 *The Fourier transform $\mathcal{F}: \mathcal{S} \to \mathcal{S}$ is continuous, that is $\forall N$ $\exists M,\ C > 0$ such that $\forall f \in \mathcal{S}$*

$$|||\, \hat{f}\, |||_N \leq C\, |||\, f\, |||_M$$

PROOF For clarity let us first do the case $N = 0$. Thus we have to bound the sup norm of $\hat{f}$ for $f \in \mathcal{S}$. Here's the trick: write

$$\hat{f}(y) = \int \frac{e^{-iy\cdot x}}{(1+|x|^2)^n}\,(1+|x|^2)^n\, f(x)\, dx$$

and notice that the first factor is integrable. Then, with $C = \displaystyle\int \frac{dx}{(1+|x|^2)^n}$ we get: $\forall f \in \mathcal{S}$, $\forall y \in \mathbb{R}^n$

$$|\hat{f}(y)| \leq C \sup_{x\in\mathbb{R}^n} |(1+|x|^2)^n\, f(x)| \leq C'\, |||\, f|||_{2n}, \tag{1.25}$$

and therefore $\sup|\hat{f}| \leq C'\,|||\, f|||_{2n}$. Now for N arbitrary we must bound the sup norm of $y^\alpha\, \hat{f}^{(\beta)}(y)$ for $|\alpha|, |\beta| \leq N$. But

$$|y^\alpha\, \hat{f}^{(\beta)}(y)| = |[x^\beta\, f^{(\alpha)}]\hat{}\,(y)|$$

and $x^\beta\, f^{(\alpha)}(x)$ is a Schwartz function, so we can apply (1.25).

□

This, together with the formula

$$\forall f,\ g \in \mathcal{S} \qquad \int \hat{f}(y)\, g(y)\, dy = \int f(x)\, \hat{g}(x)\, dx \tag{1.26}$$

justifies the following definition:

Definition 1.5.2 *If $u \in \mathcal{S}'$ is a tempered distribution, its Fourier transform is the functional*

$$\forall f \in \mathcal{S} \qquad <\hat{u}\,,\, f> := <u\,,\, \hat{f}>\,.$$

By proposition 1.5.1 $\hat{u}$ is indeed a tempered distribution, and by (1.26) this definition extends the original definition of Fourier transform from $\mathcal{S}$ to $\mathcal{S}'$.

Lemma 1.5.3 *For every $u \in \mathcal{S}'$ and all multi-indices α one has:*

$$\widehat{D^\alpha u} = y^\alpha\, \hat{u}.$$

Lemma 1.5.4 *For every $u \in \mathcal{S}'$ and all $f \in \mathcal{S}$ define:*

$$< \check{u}\,,\, f > := < u\,,\, \check{f} >$$

where $\check{f}(x) = (2\pi)^{-n} \int e^{ix\cdot y} f(y)\, dy$ is the inverse Fourier transform. Then:

1. *$\check{u}$ is a tempered distribution.*

2. *$\hat{}$ and $\check{}$ are inverses of each other.*

Exercise 1.5.1 Prove these two lemmas.

1.5.2 Examples

A. Let δ be the Dirac delta at the origin in $\mathbb{R}^n$. Then

$$< \hat{\delta}\,,\, f > = \hat{f}(0) \;=\; \int f(x)\, dx,$$

and so we can say: $\hat{\delta}$ *is equal to the constant function one.*

B. Let α be a multiindex. Then by the previous example and Lemma 1.5.3 we can conclude:

$$\widehat{D^\alpha \delta} \;=\; y^\alpha. \tag{1.27}$$

Observe that the greater $|\alpha|$ the faster the F.T. of $\delta^{(\alpha)}$ grows at infinity, and also the more singular $\delta^{(\alpha)}$ is.

C. Consider $\mathbb{R}^2$ with coordinates (x_1, x_2), and let $a \in C_0^\infty(\mathbb{R})$. Define a distribution u by the formula

$$< u\,,\, f > = \int_{-\infty}^{\infty} a(x_1)\, f(x_1, 0)\, dx_1.$$

This is a delta-function along the x_1 axis (or "layer"). Then one easily checks that: *the Fourier transform of u is the function of two variables*

$$\hat{u}\,(y_1, y_2) \;=\; \hat{a}(y_1).$$

Notice that this function is rapidly decreasing along any ray that is *not* vertical. So in a sense the only "singular direction" of u is the vertical direction. The wave-front set (to be studied in the next chapter) is based on this general idea.

D. Let's compute the Fourier transform of the Heaviside function, H. It may be tempting to argue as follows: "$H' = \delta$, therefore $iy\hat{H} = 1$ and so $\hat{H} = -i/y$" The first part is correct, but remember that $y\delta(y) = 0$ so we must be careful (also it is not clear what we mean by "$-i/y$", principal value?). Strictly speaking we must compute, for $f \in \mathcal{S}(\mathbb{R})$

$$< \hat{H}\,,\, f >= \int_0^\infty \hat{f}(x)\,dx \;=\; \int_0^\infty \int_{-\infty}^\infty e^{-ixy}\, f(y)\,dy\,dx.$$

We cannot change the order of integration at this point, since the integrand is not integrable over (x, y) space. Consider instead

$$\int_0^\infty e^{-\tau x}\,\hat{f}(x)\,dx \;=\; \int_0^\infty \int_{-\infty}^\infty e^{-ixy}\,e^{-\tau x}\,f(y)\,dy\,dx, \tag{1.28}$$

where $\tau > 0$. The integrand is now integrable over $(x, y) \in (0, \infty) \times \mathbb{R}$, and

$$< \hat{H}\,,\, f >= \lim_{\tau \to 0+} \int_0^\infty e^{-\tau x}\,\hat{f}(x)\,dx,$$

by the dominated convergence theorem.

Changing the order of integration in (1.28) gives $\int_0^\infty e^{-\tau x}\,\hat{f}(x)\,dx \;=< h_\tau\,,\, f >$, where

$$h_\tau(y) \;=\; \int_0^\infty e^{-x(iy+\tau)}\,dx \;=\; \frac{1}{iy + \tau}$$

(recall that $\tau > 0$). Therefore we conclude:

$$\hat{H} \;=\; -i\,(y - i0)^{-1}. \tag{1.29}$$

Exercise 1.5.2 Prove that $(y-i0)^{-1} = 1/y + i\pi\delta$ ($1/y$ in the principal value sense), so indeed the naive argument above misses a delta function.

E. The previous example can be generalized to the calculation of the F.T. of x_+^a, with $\Re a > -1$. We have:

$$\int_0^\infty e^{-\tau x}\,x^a\,\hat{f}(x)\,dx \;=< h_\tau\,,\, f >, \text{ where } \quad h_\tau(y) \;=\; \int_0^\infty e^{-x(iy+\tau)}\,x^a\,dx. \tag{1.30}$$

Now change variables from x to $\zeta := x(iy + \tau)$. (Notice that $iy + \tau$ is on the right half plane and so we can use the standard branch of the logarithm.) Then

$$h_\tau(y) \;=\; \frac{1}{(iy+\tau)^{a+1}} \int_\Sigma e^{-\zeta}\,\zeta^a\,d\zeta$$

where Σ is the ray through the origin and $(iy + \tau)$. The latter is on the right half-plane ($\Re > 0$), therefore by Cauchy's theorem we can move the contour of integration to the positive real semi-axis. (The integral on the line segment from $b \in \mathbb{R}^+$ to icb with c a constant is exponentially small in b.) Therefore:

$$h_\tau(y) \;=\; \frac{1}{(iy+\tau)^{a+1}}\,\Gamma(a+1),$$

where Γ is the Gamma function. Finally we obtain:

$$\hat{x_+^a} = -i\,\Gamma(a+1)\,\frac{e^{-ia\pi/2}}{(y-i0)^{a+1}}, \tag{1.31}$$

generalizing (1.29).

1.5.3 Paley-Wiener Theorems.

Theorem 1.5.5 *Let $u \in \mathcal{E}'$ be a compactly supported distribution. Then its Fourier transform, $\hat{u} \in \mathcal{S}'$ (in the sense of Definition 1.5.2), is actually equal to the function on $\mathbb{R}^n$*

$$\hat{u} = (\xi \mapsto < u\,,\, e^{-ix\cdot\xi} >).$$

Moreover this function extends to an entire function on $\mathbb{C}^n$, and it satisfies the estimate: $\exists N,\, C,\, r > 0$ such that $\forall \zeta \in \mathbb{C}^n$

$$|< u\,,\, e^{-ix\cdot\zeta} >| \leq C\,(1+|\zeta|)^N\, e^{r|\mathrm{Im}\zeta|} \tag{1.32}$$

Proof Begin by constructing an approximation to the identity in the standard way: Let $\chi \in C_o^\infty(\mathbb{R}^n)$ such that:

$$\int \chi(x)\,dx = 1, \quad \chi(-x) = \chi(x)$$

and let $\chi_k(x) = k^n\,\chi(kx)$. Let

$$u_k := \chi_k * u.$$

Then (by a homework problem) $\forall k\; u_k \in C_o^\infty(\mathbb{R}^n)$ and $u_k \to u$ weakly. Since the Fourier transform is continuous in $\mathcal{S}'$ as well, then $\hat{u_k} \to \hat{u}$ weakly. On the other hand $\hat{u_k}$ is the smooth function

$$\hat{u_k}(\xi) = < u * \chi_k\,,\, e^{-ix\cdot\xi} >$$

which clearly extends to an entire function, and these functions are converging weakly to $\hat{u}$. Next invoke Proposition 1.4.21 to conclude that

$$< u * \chi_k\,,\, e^{-ix\cdot\xi} > = < u\,,\, \chi_k * e^{-ix\cdot\xi} >, \tag{1.33}$$

because $\chi_k(-x) = \chi_k(x)$. Moreover

$$\chi_k * e^{-ix\cdot\zeta} \to e^{-ix\cdot\zeta} \quad \text{in the sense of } \mathcal{E}_x$$

and uniformly for ζ on compact sets of $\mathbb{C}^n$. Therefore

$$\hat{u_k}(\zeta) \to < u\,,\, e^{-ix\cdot\zeta} >$$

uniformly on compact sets of $\mathbb{C}^n$.

$\square$

Proposition 1.5.6 *If $u \in \mathcal{E}'$, then:*

$$u \in C_0^\infty \;\Leftrightarrow\; \forall N \; \exists C \quad \textit{such that} \; \forall \xi \in \mathbb{R}^n \quad |\hat{u}(\xi)| \leq C\,(1+|\xi|)^{-N}.$$

Chapter 2

The Principal Symbol and the Wave-Front set

The main goal of this chapter is to introduce the main invariants in microlocal analysis, namely the principal symbol and the wave-front set.

2.1 The Symbol Calculus of Differential Operators

Let $\mathcal{U} \subset \mathbb{R}^n$ be an open set, and denote by $\mathrm{Diff}^m(\mathcal{U})$ the space of differential operators of order m with smooth coefficients. An element $P \in \mathrm{Diff}^m(\mathcal{U})$ will be written in the form

$$P \;=\; \sum_{\alpha\ ,\ |\alpha|\le m} a_\alpha(x)\ D^\alpha,$$

where we are using the standard multi-index notation.

Definition 2.1.1 *The full symbol, or amplitude, of $P \in \mathit{Diff}^m(\mathcal{U})$ is the following function on $\mathcal{U} \times \mathbb{R}^n$:*

$$P(x,\xi) \;=\; \sum_{\alpha\ ,\ |\alpha|\le m} a_\alpha(x)\ \xi^\alpha\ .$$

The principal symbol of P is the top order term of this function, that is

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

We will: (1) show that the principal symbol of a differential operator on a manifold M is naturally a function on the cotangent bundle of M, and (2) answer the questions: what are the symbols of the composition and of the commutator of two operators? To answer the last question will require a discussion of the symplectic structure of the cotangent bundle.

Lemma 2.1.2 *Let $P \in Diff^m(\mathcal{U})$, $\rho \in C_0^\infty(\mathcal{U})$, $\phi \in C_0^\infty(\mathcal{U}, \mathbb{R})$. Then*

$$\sigma_P(x_0, d_{x_0}\phi) = \lim_{\tau\to\infty} \frac{1}{\rho(x_0)} e^{-i\tau\phi(x_0)} \tau^{-m} P(\rho e^{i\tau\phi})(x_0).$$

Proof Simply apply Leibnitz formula to compute $P(\rho e^{i\tau\phi})$, keeping only the terms of highest degree in τ.

□

Remark. If M is a manifold Diff($\mathcal{M}$) is the algebra of operators generated by multiplication by smooth functions and differentiation by smooth vector fields. If $P \in \mathrm{Diff}(M)$ is of order m, $\sigma_P : T^*M \to \mathbb{C}$ is the function on T^*M defined as follows: if $x_0 \in M$ and $\xi \in T^*_{x_0}M$ (i.e. $\xi : T_{x_0}M \to \mathbb{R}$ is linear), let ρ, ϕ be smooth functions on M such that $\rho(x_0) = 1$, $d\phi_{x_0} = \xi$, then

$$\sigma_P(x_0, \xi) := \lim_{\tau\to\infty} (e^{-i\tau\phi(x_0)} \frac{1}{\tau^n} P(\rho e^{i\tau\phi})(x_0)).$$

This is independent of our choice of ρ and ϕ (provided that $d\phi_{x_0} = \xi$ and $\rho(x_0) = 1$).

Lemma 2.1.3 *(The symbol calculus, part 1.)*

1. *The mapping $P \mapsto \sigma_P$ is linear.*
2. $\sigma_{P\circ Q}(x, \xi) = \sigma_P(x, \xi) \cdot \sigma_Q(x, \xi)$

Notice that in particular $\sigma_{P\circ Q} = \sigma_{Q\circ P}$. This raises the following:

Question: What is $\sigma_{[P,Q]}$? Here $[P, Q] = P \circ Q - Q \circ P$. This is a differential operator of order the sum of the orders of P and Q *minus one.*

Definition 2.1.4 *Let $H, G \in C^\infty(\mathcal{U} \times \mathbb{R}^n)$. Then the Poisson bracket of H and G is the function $\{H, G\} \in C^\infty(\mathcal{U} \times \mathbb{R}^n)$ defined by:*

$$\{H, G\} = \sum_{j=1}^n \frac{\partial H}{\partial \xi_j}\frac{\partial G}{\partial x_j} - \frac{\partial H}{\partial x_j}\frac{\partial G}{\partial \xi_j}.$$

Lemma 2.1.5 *(The symbol calculus, part 2.)*

$$\sigma_{[P,Q]} = \frac{1}{i} \{\sigma_P , \sigma_Q\}.$$

Proof Let $Q = \sum_{|\beta|\le\ell} b_\beta D^\beta$, so that

$$P \circ Q = \sum_{\alpha,\beta} a_\alpha D^\alpha b_\beta D^\beta.$$

$$D_n^{\alpha_n} b_\beta D^\beta = \sum_{j=0}^{\alpha_n} \binom{\alpha_n}{j} D_n^j (b_\beta) D_n^{\alpha_n - j} D^\beta = D^\beta D_n^{\alpha_n} D^\beta + \alpha_n (D_n b_\beta) D^{\alpha_n - 1} D^\beta + \mathrm{lc}$$

$$D_{n-1}^{\alpha_n - 1} D_n^{\alpha_n}(b_\beta D^\beta) = b_\beta D_{n-1}^{\alpha_n - 1} D_n^{\alpha_n} D^\beta + \alpha_{n-1} D_{n-1}(b_\beta) D_{n-1}^{n-1-1} D_n^{\alpha_n} D^\beta$$
$$+\alpha_n D_n(b_\beta) D_{n-1}^{\alpha_n - 1} D_n^{\alpha_n - 1} D^\beta + \mathrm{l.o.t.}$$
$$\Rightarrow D^\alpha b_\beta D^\beta = b_\beta D^\alpha D^\beta + \sum_{j=1}^{n} x_j D_j(b_\beta) D^{\alpha - [j]} D^\beta + \mathrm{l.o.t.}$$

□

Proposition 2.1.6 *(Basic Properties of the Poisson bracket.)*

1. $\{H, G\} = < dG, \Xi_H >$ *and* $dG = \omega \rfloor \Xi_G \Rightarrow \{H, G\} = \omega(\Xi_G, \Xi_H) \Rightarrow \{H, G\} = -\{G, H\}$ *because* ω *is alternating.*

2. $\{H, GK\} = \{H, G\}K + \{H, K\}G$

3. *It satisfies the Jacobi Identity:* $\{H, \{G, K\}\} + \{K, \{H, G\} + \{H + G, K\} = \{H, K\} + \{G, K\}$.

4. $\{H,G,K\} = \{H,K\} + \{G,K\}$.

2.2 The Wave-Front Set

Let $\mathcal{U}$ be an open set in $\mathbb{R}^n$, and u a distribution on $\mathcal{U}$. It follows from Proposition 1.5.6 that

$$u \in C^\infty(\mathcal{U}) \quad \Leftrightarrow \quad \forall \rho \in C_0^\infty(\mathcal{U}) \quad |\hat{\rho u}(\xi)| \ \leq \ C_N \ (1 + |\xi|)^{-N}$$

for all integers $N > 0$. Localizing this condition both in the x and the ξ variables leads to the following definition:

Definition 2.2.1 *Given* $u \in \mathcal{D}'(\mathcal{U})$ *and* $(x, \xi_0) \in \mathcal{U} \times \left(\mathbb{R}^n \setminus 0\right)$, *we will say that* u *is microlocally smooth at* (x, ξ_0) *iff there exists a neighborhood,* U, *of* x *in* $\mathcal{U}$, *and a conic neighborhood,* $\mathcal{C}$, *of* ξ_0 *in* $\mathbb{R}^n \setminus 0$ *with the following property:*

$$\forall \rho \in C_0^\infty(U) \quad \text{and} \quad \forall N > 0 \ \exists C_N$$

such that

$$\forall \xi \in \mathcal{C} \quad |\widehat{\rho u}(\xi)| \ \leq \ C_N \ (1 + |\xi|)^{-N} \ .$$

The wave-front set of u, *denoted* WF(u), *is the complement in* $\mathcal{U} \times \left(\mathbb{R}^n \setminus 0\right)$ *of the set of* (x, ξ)*'s at which* u *is microlocally smooth.*

Hence $WF(u)$ is a closed conic subset of $\mathcal{U} \times \left(\mathbb{R}^n \setminus 0\right)$. We first notice that the previous definition has a number of equivalent reformulations.

Lemma 2.2.2 *The following conditions are equivalent:*
(1) u is microlocally smooth at (x_0, ξ_0).
(2) $\exists \rho \in C_0^\infty(\mathcal{U})$ with $\rho(x_0) \neq 0$ and a conic neighborhood of ξ_0, $\mathcal{C}$, such that

$$\forall N > 0 \ \exists C_N \ \textit{such that} \ \forall \xi \in \mathcal{C} \quad |\widehat{\rho u}(\xi)| \ \leq \ C_N \ (1 + |\xi|)^{-N} \ .$$

(3) $\exists \rho \in C_0^\infty(\mathcal{U})$ with $\rho(x_0) \neq 0$ and a neighborhood, W, of ξ_0 in the sphere in $\mathbb{R}^n$ centered at the origin and of radius $|\xi_0|$ such that

$$\forall N > 0 \ \exists C_N \ \textit{such that} \ \forall \xi \in W \ \textit{and} \ \forall \tau > 1 \quad |\widehat{\rho u}(\tau\xi)| \ \leq \ C_N \ \tau^{-N} \ .$$

Proof The equivalence of 2 and 3 is elementary, and (1) $\Rightarrow$ (2) is trivial. Assume (2), and let U be a neighborhood of x_0 in $\mathcal{U}$ where ρ is bounded away from zero. Let $\rho' \in C_0^\infty(U)$. Then we can write

$$\rho' u \ = \ f \ \rho u \ ,$$

with $f \in C_0^\infty$. Taking Fourier transforms on both sides of this equation we get

$$(\rho' u)\hat{} \ = \ \hat{f} \ \star \ \widehat{\rho u} \ .$$

Let $g = \widehat{\rho u}$; then the previous equation reads

$$(\rho' u)\hat{}(\xi) \ = \ \int \hat{f}(\eta) \ g(\xi - \eta) \ d\eta \ .$$

We will estimate this integral with ξ in a smaller cone contained in $\mathcal{C}$, as follows (this is lemma 8.8.1 in Hörmander).

We begin by noting that $\hat{f}$ is rapidly decreasing, while g has polynomial growth in general and is rapidly decreasing in $\mathcal{C}$. Hence in a region where $|\eta|$ dominates $|\xi - \eta|$, the rapid decrease of $\hat{f}$ will make the integral rapidly decreasing. More precisely, let $0 < c < 1$ and distinguish the two regions

$$\begin{aligned} (I) \ &: \ |\eta| < c|\xi| \\ (II) \ &: \ |\eta| \geq c|\xi| \ . \end{aligned}$$

By the triangle inequality, in region (II) $|\xi - \eta| \leq (1 + c^{-1})|\eta|$. Let M, C be such that $|g(\eta)| \leq C(1 + |\eta|)^M$. Hence

$$|\eta| \geq c|\xi| \ \Rightarrow \ |g(\xi - \eta)| \ \leq \ C \ (1 + |\eta|)^M$$

with a different constant. Let $A(\xi)$ resp. $B(\xi)$ be the integral of $|\hat{f}(\eta)\ g(\xi-\eta)|$ over the region (I), resp. (II). Then

$$B(\xi)\ \leq\ C \int_{|\eta|>c|\xi|} |\hat{f}(\eta)|\ (1+|\eta|)^M\ d\eta\ .$$

But $\forall K>0\ \exists C_K$ such that

$$|\hat{f}(\eta)|\ \leq\ C_K\ (1+|\eta|)^{-K}\ ,$$

and so with different constants we get that $\forall K$ and $\forall \xi$

$$B(\xi)\ \leq\ C_K \int_{|\eta|>c|\xi|} (1+|\eta|)^{M-K}\ d\eta\ ,$$

from which rapid decrease of B follows easily.

We still have to estimate A. We will use the obvious inequality

$$A(\xi)\ \leq\ \|\hat{f}\|_{L^1} \sup_{|\eta|<c|\xi|} |g(\xi-\eta)|\ =\ \|\hat{f}\|_{L^1} \sup_{|\eta-\xi|<c|\xi|} |g(\eta)|\ .$$

Let $\mathcal{C}_1$ be an open cone whose closure is included in $\mathcal{C}$. We claim that there is then a c, $0<c<1$, independent of f, such that

$$\xi\in\mathcal{C}_1 \text{ and } |\xi-\eta|<c|\xi| \quad\Rightarrow\quad \eta\in\mathcal{C}\ .$$

To see this, first look at $\mathcal{C}_1$ intersected with the unit sphere; this is a compact contained in $\mathcal{C}$, so c certainly exists for such ξ. That particular c works for arbitrary $|\xi|$, by homogeneity. But then

$$\forall\xi\in\mathcal{C}_1\quad A(\xi)\ \leq\ \|\hat{f}\|_{L^1}\ C_N \sup_{|\xi-\eta|<c|\xi|} (1+|\eta|)^{-N}\ ,$$

for all $N>0$. Since

$$|\xi|-|\eta|\ \leq\ |\eta-\xi|,$$

$$|\xi-\eta|<c|\xi|\ \Rightarrow\ |\eta|>(1-c)|\xi|.$$

Since $c<1$, this concludes the proof. Notice by the way that c is independent of f.

□

Corollary 2.2.3 *Let $\pi:\mathcal{U}\times(\mathbb{R}^n\setminus\{0\})\to\mathcal{U}$ the projection onto the first factor. Then*

$$\pi(WF(u))\ =\ \textit{singular support}(u).$$

PROOF Suppose x_0 is not in the singular support of u. We can then find a cut-off function ρ such that $\rho(x_0) \neq 0$ and ρu is a smooth function. Therefore the Fourier transform of ρu is rapidly decreasing, and condition (b) of the previous Lemma is satisfied $\forall \xi_0$ with $\mathcal{C} = \mathbb{R}^n \setminus \{0\}$. Therefore x_0 is not in the projection of the wave-front set of u.

Conversely, suppose $x_0 \notin \pi(WF(u))$, that is: $\forall \xi \in \mathbb{R}^n \setminus \{0\}$ (x_0, ξ) is not in $WF(u)$. By definition, this means that $\forall \xi$ $\exists U_\xi$ neighborhood of x_0 and $\exists \mathcal{C}_\xi$ conic neighborhood of ξ such that $\forall \rho \in C_0^\infty(U_\xi)$ the Fourier transform of ρu is rapidly decreasing in the cone $\mathcal{C}_\xi$. By compactness of the unit sphere in $\mathbb{R}^n$, there exist $\xi_1, \ldots, \xi_M$ such that

$$\cup_{j=1}^M \mathcal{C}_{\xi_j} = \mathbb{R}^n \setminus \{0\}.$$

Let

$$U = \cap_{j=1}^M U_{\xi_j}.$$

Then $\forall \rho \in C_0^\infty(U)$, the Fourier transform of ρu is rapidly decreasing in $\mathbb{R}^n$ and therefore ρu is smooth. This easily implies that u is smooth on U.

$\square$

Exercise 2.2.1 Let u be the distribution on $\mathbb{R}$ which is the Fourier transform of the Heaviside function. Compute the wave-front set of u.

2.3 The Method of Stationary Phase

As motivation consider the following problem. Let $S \subset \mathbb{R}^3$ be a surface, and let δ_S be the distribution

$$< \delta_S \, , \, f > = \int_S f \, dA$$

where dA is the surface area element. What is the wave-front set of δ_S? The question reduces to the following: estimate for large τ the expression

$$I(\tau) \; = < \delta_S \, , \, \rho \, e^{i\tau x \cdot \xi_0} > = \int_S \rho(x) \, e^{-i\tau x \cdot \xi_0} \, dA_x,$$

where ρ is a smooth cut-off with small support. Introducing a parametrization $x = x(u, v)$ of a neighborhood of the support of ρ, the above integral takes the form:

$$I(\tau) = \int\!\!\int e^{-i\tau x(u,v)\cdot\xi_0} \, \rho(x(u, v)) \, J(u, v) \, du \, dv. \tag{2.1}$$

We face the problem of estimating such an integral as $\tau \to \infty$. The method of stationary phase is a method for estimating integrals of the general form:

$$I(\tau) = \int_{\mathbb{R}^n} e^{i\tau\phi(x)} \, a(x) \, dx \tag{2.2}$$

where ϕ and a are smooth and a is compactly supported. We'll restrict our attention to real-valued ϕ, for simplicity. Notice the obvious bound

$$|I(\tau)| \leq \|a\|_{L^1}. \tag{2.3}$$

The large τ behaviour of (2.2) depends on the nature of the phase, ϕ. Let's look at two extreme cases.

1. Suppose ϕ is constant: $\phi(x) = c$. Then
$$I(\tau) = e^{i\tau c} A,$$
where A is a constant (i.e. independent of τ). Therefore $I(\tau)$ oscillates but $|I(\tau)|$ is a constant.

2. Suppose ϕ is a coordinate function, say $\phi(x) = x_1$. Then
$$I(\tau) = \hat{a}(-\tau, 0, \ldots, 0).$$
Since a is smooth compactly supported $\hat{a}$ is Schwartz, and therefore in this case
$$\forall N > 0 \qquad I(\tau) = O(\tau^{-N}).$$

Intuitively, in case 2 we obtain rapid decrease because for large τ the exponential $\exp(i\tau\phi(x))$ oscillates many times between x and $x + dx$, while the amplitude $a(x)$ remains relatively constant (since a is smooth). Therefore in the sum (integral) many cancellations take place. This is clearly the opposite of what happens in case 1. If ϕ has a critical point x_0, i.e. if $d\phi_{x_0} = 0$, then between x_0 and $x_0 + dx$ the phase ϕ doesn't change very much, so infinitesimally we are in case 1. Therefore we expect that in general the main contributions to $I(\tau)$ should arise from the critical points of ϕ, that is points where ϕ is "stationary". This is indeed the case.

2.3.1 Non-stationary phases.

We first prove that indeed if we neglect rapidly decreasing functions of τ the only contributions to $I(\tau)$ come from the singularities of the phase, ϕ.

Proposition 2.3.1 *Let* $I(\tau) = \int_{\mathbb{R}^n} e^{i\tau\phi(x)}\, a(x)\, dx$ *where* $a \in C_o^\infty(\mathbb{R}^n)$ *and* ϕ *is smooth. Assume that* ϕ *doesn't have any critical points in a neighborhood of the support of* a*. Then*
$$\forall N > 0 \qquad I(\tau) = O(\tau^{-N}).$$

PROOF The idea is to use the identity:

$$\nabla\phi\,(e^{i\tau\phi(x)}) \;=\; i\tau\|\nabla\phi\|^2 e^{i\tau\phi(x)}$$

and to integrate by parts. Here on the left-hand side $\nabla\phi$ is considered as the differential operator

$$\nabla\phi \;=\; \sum_{j=1}^{n} \frac{\partial\phi}{\partial x_j}\frac{\partial}{\partial x_j}.$$

Let χ be a smooth cut-off function identically one on the support of a and such that $\nabla\phi$ does not vanish on its support. Therefore $a\chi = a$, and $\chi/\|\nabla\phi\|^2$ is a smooth function. Let

$$L \;:=\; \frac{\chi}{\|\nabla\phi\|^2}\sum_{j=1}^{n} \frac{\partial\phi}{\partial x_j}\frac{\partial}{\partial x_j}.$$

This is a smooth vector field satisfying

$$L\,e^{i\tau\phi(x)} \;=\; i\tau\,\chi\,e^{i\tau\phi(x)},$$

and therefore

$$a\,e^{i\tau\phi(x)} \;=\; a\,\chi\,e^{i\tau\phi(x)} \;=\; \frac{a}{i\tau}\,L\,e^{i\tau\phi(x)}.$$

Therefore

$$I(\tau) \;=\; \int \frac{a}{i\tau}\,L\,e^{i\tau\phi(x)}\,dx \;=\; \frac{i}{\tau}\int e^{i\tau\phi(x)}\,L^t a\,dx. \tag{2.4}$$

Here L^t is the transpose of L. Applying the bound (2.3), we get:

$$|I(\tau)| \leq \frac{1}{|\tau|}\,\|L^t a\|_{L^1}$$

Notice however that the right-hand side of (2.4) is of the same form as $I(\tau)$ itself (except for the factor of i/τ). Therefore the argument can be repeated any number of times.

$\square$

In applications it will be important to let both the phase and the amplitude depend on parameters, that is consider

$$I(\tau, w) \;=\; \int e^{i\tau\phi(x,w)}\,a(x,w)\,dx \tag{2.5}$$

where w ranges over some compact domain Ω in $\mathbb{R}^m$. (For example, in (2.1) we have the very important parameter ξ_0.) Of course we can repeat e.g. the analysis above for each fixed w; the only question is that of the uniformity of the estimates with respect to ω.

Proposition 2.3.2 *Assume a and ϕ are smooth. Let $K \subset \mathbb{R}^n$ be a fixed compact set such that*

$$\forall w \in \Omega,\ \forall x \notin K \qquad a(x,w) \,=\, 0$$

(i.e. for each w the support of $a(\cdot,w)$ is contained in K). Assume that

$$\{(x,w) \in K \times \Omega\,;\, \nabla_x \phi(x,w) = \vec{0}\} \,=\, \emptyset.$$

Then

$$\forall N \qquad |I(\tau,w)| = O(\tau^{-N})$$

uniformly on $w \in \Omega$.

2.3.2 Quadratic phases.

2.3.3 Non-degenerate critical points.

2.4 Operations with distributions.

2.4.1 Push-forwards.

Throughout this section we let let $F : \mathcal{U} \to \mathcal{V}$ be a smooth map between open subsets of Euclidean spaces. Composition with F defines a map,

$$F^* \,:\, C^\infty(\mathcal{V}) \to C^\infty(\mathcal{U}),$$

namely:

$$F^*(g) \,:=\, g \circ F.$$

It is easy to check that this map is continuous in the C^∞ topology, and therefore by duality we can define:

Definition 2.4.1 *If u is a compactly supported distribution on $\mathcal{U}$, the push-forward $F_*(u)$ is the distribution on $\mathcal{V}$ defined by:*

$$< F_*(u)\,,\, g > := < u\,,\, g \circ F > .$$

The behaviour of the wave-front set under this operation is described in Theorem (2.4.3). We will first prove this Theorem in case F is a diffeomorphism:

Theorem 2.4.2 *Assume $F : \mathcal{U} \to \mathcal{V}$ is a diffeomorphism between open subsets of $\mathbb{R}^n$. Then for any $u \in \mathcal{D}'(\mathcal{U})$*

$$WF(F_*(u)) \,=\, F_*(\,WF(u))\,, \quad \textit{where}$$

$$F_*(\,WF(u)) \,:=\, \{\,(y,\eta) \in T^*\mathcal{V}\backslash\{0\}\,;\, \exists(x,\xi) \in WF(u) \textit{ s. t. } y = F(x),\ \xi = dF_x^T(\eta)\,\}. \tag{2.6}$$

PROOF Let $(y_0, \eta_0) \notin F_*(\mathrm{WF}(u))$. We must estimate the Fourier transform of $\rho F_*(u)$, where ρ is a suitably-chosen cut-off function near y_0. Let $x_0 = F^{-1}(y_0)$, and $\xi_0 = dF^T_{x_0}(\eta_0)$. By assumption $(x_0, \xi_0) \notin \mathrm{WF}(u)$, therefore there exist a cut-off $\chi(x)$ and a conic neighborhood of ξ_0, $\mathcal{C}$, such that $\chi(x_0) = 1$ and $\widehat{\chi u}(\xi)$ is rapidly decreasing for $\xi \in \mathcal{C}$. By shrinking the support of χ if necessary we can assume that there exists another cut-off, f, such that $f = 1$ on the support of χ and $\widehat{fu}(\xi)$ is rapidly decreasing for $\xi \in \mathcal{C}$. Notice that $f\chi = \chi$.

Let $\rho = \chi \circ F^{-1}$, and $g = f \circ F^{-1}$. By the Fourier inversion formula we can write:

$$\begin{aligned} (\rho F_*(u))\hat{}(\tau\eta) &= (\rho\, g\, F_*(u))\hat{}(\tau\eta) \;=< gF_*(u)\,,\, \rho\, e^{-i\tau y\cdot\eta} >= \\ & (2\pi)^{-n} \int \widehat{fu}(\xi) \int \chi(x)\, e^{-i\tau F(x)\cdot\eta}\, e^{ix\cdot\xi}\, dx\, d\xi, \end{aligned}$$

where this is to be understood as an iterated integral. Making the change of variables $\xi = \tau w$, we can write this as:

$$(\frac{\tau}{2\pi})^n \int \widehat{fu}(\tau w) \int \chi(x)\, e^{i\tau[x\cdot w - F(x)\cdot\eta]}\, dx\, dw.$$

$\square$

As a corollary of this we get that the wave-front set is a well-defined subset of the punctured cotangent bundle of $\mathcal{U}$, $T^*\mathcal{U} \setminus \{0\}$. Although of course we will identify

$$T^*\mathcal{U} \setminus \{0\} \;\cong\; \mathcal{U} \times (\mathbb{R}^n \setminus \{0\})\ ,$$

this statement means that if we change variables according to a diffeomorphism, $F : \mathcal{U} \to \mathcal{V}$, the wave-front set of a distribution transforms as subsets of $T^*\mathcal{U}$, $T^*\mathcal{V}$.

Theorem 2.4.3 *Let $F : \mathcal{U} \to \mathcal{V}$ be a smooth map, where $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{V} \subset \mathbb{R}^m$ are open subsets. Then for any $u \in \mathcal{D}'(\mathcal{U})$*

$$WF(F_*(u)) \;\subset\; F_*(\,WF(u)),$$

where the right-hand side is defined by (2.6).

2.5 The non-characteristic regularity theorem.

The purpose of this section is to prove the following Theorem, relating the wave-front set and the principal symbol:

Theorem 2.5.1 *Let P be a linear differential operator of order m with smooth coefficients on an open set $\mathcal{U} \subset \mathbb{R}^n$. Then $\forall u \in \mathcal{D}'(\mathcal{U})$*

$$WF(u) \subset Char\ (P) \cup WF(Pu),$$

where

$$Char\ (P) := \{ (x,\xi) \in \mathcal{U} \times (\mathbb{R}^n \setminus \{0\}) \ ; \ \sigma_P(x,\xi) = 0 \}$$

is the characteristic variety of P*.*

This Theorem gives information on the wave-front set of solutions to PDEs: If u solves $Pu = f$ with f given then the Theorem gives us an upper bound on $\mathrm{WF}(u)$ in terms of data associated to P and f. There are much more refined theorems of this sort, for special P's, but this result is the most general one.

We note an important corollary:

Definition 2.5.2 *An operator* P *is called elliptic iff Char* $(P) = \emptyset$*.*

Corollary 2.5.3 *If* P *is elliptic then for all distributions* u*,* $WF(Pu) = WF(u)$ *In particular if* $Pu \in C^\infty(\mathcal{U})$*, then* $u \in C^\infty(\mathcal{U})$*.*

Proof Just apply the Theorem and the fact that for all differential operators P $\mathrm{WF}(Pu) \subset \mathrm{WF}(u)$.

$\square$

In the proof of Theorem 2.5.1 we will need the following Lemma:

Lemma 2.5.4 *Let* $v \in \mathcal{E}'(\mathcal{U})$ *be a compactly supported distribution and let* $(x_0,\xi_0) \in \mathcal{U} \times (\mathbb{R}^n \setminus \{0\}) \setminus WF(v)$*. Then there exists a neighborhood* U *of* x_0 *and a conic neighborhood* V *of* ξ_0 *such that:* $\forall N > 0\ \exists C_N > 0$ *such that for all* $h \in C_0^\infty(U)$ *and all* $\xi \in V$ *one has:*

$$|\widehat{hv}(\xi)| \leq C_N \left(\sum_{|\alpha| \leq N+M} \sup |D^\alpha h| \right) (1 + |\xi|)^{-N},$$

where M *is the order of* v*.*

Proof Since $(x_0,\xi_0) \notin \mathrm{WF}(v)$, there exist ρ and V be such that $\rho(x_0) = 1$ and $\widehat{\rho v}$ is rapidly decreasing on V. Now choose U small enough so that $\rho|_U \geq 1/2$. Now any $h \in C_0^\infty(U)$ is of the form $h = f\rho$, with $f \in C_0^\infty$. We must estimate

$$\widehat{hv} = \hat{f} \star \hat{g}, \qquad \text{where } g := \widehat{\rho v}.$$

This is *exactly* the situation in the proof of (2) $\Rightarrow$ (1) of Lemma (2.2.2), with identical notation. Keeping track of the estimates in that proof, we see that the constants that arise that depend on f are:

$$C_{f,0} := \|\hat{f}\|_{L^1} \leq (2\pi)^{-n} \|f\|_{L^\infty}$$

(the inequality follows from the Fourier inversion formula) and $E_{f,K}$ such that

$$\forall \eta \qquad |\hat{f}(\eta)| \leq E_{f,K}(1+|\eta|)^{-K},$$

the latter providing control in terms of $(1+|\xi|)^{M-K}$. Recalling how one proves that $\hat{f}$ is rapidly decreasing one sees that $E_{f,K}$ is dominated by universal constants times the C^K norm of f.

□

The point of this Lemma is the explicit dependence on f of the constants; C_N is independent of f.

PROOF OF THEOREM 2.5.1. Let $(x_0, \xi_0) \notin \text{Char}\,(P) \cup \text{WF}(Pu)$, that is $(x_0, \xi_0) \notin \text{Char}\,(P)$ and $(x_0, \xi_0) \notin \text{WF}(Pu)$. We need to show that $(x_0, \xi_0) \notin \text{WF}(u)$. Thus we'll need to estimate, as usual,

$$\widehat{\rho u}(\xi) \; = < u\,,\, \rho(x)\, e^{-ix\cdot\xi} > \tag{2.7}$$

where ρ is a cutoff function such that $\rho(x_0) \neq 0$ and the estimate is for ξ large in some conic neighborhood of ξ_0. The idea of the proof is as follows. Let P^T be the transpose of P; thus

$$P \;=\; \sum_\alpha a_\alpha(x)\; D^\alpha \;\Rightarrow\; P^T \;=\; \sum_\alpha (-1)^{|\alpha|}\, D^\alpha\; a_\alpha(x)$$

and

$$< u\,,\, P^T f > = < Pu, f > .$$

(This identity is true by definition of Pu.) Keeping in mind that we know that $(x_0, \xi_0) \notin \text{WF}(Pu)$, it is natural to try to use this identity to estimate (2.7). For this we would need to find f such that

$$P^T(f)(x) \;=\; \rho(x)\, e^{-ix\cdot\xi}, \tag{2.8}$$

and f itself should have the form

$$f(x) \;=\; v(x)\, e^{-ix\cdot\xi} \tag{2.9}$$

with $v(x_0) \neq 0$. It should be clear that if such an f exists we are done. By Lemma 2.5.4, it should also be clear that we could let v depend on ξ as well, provided that it does so in a "gentle" fashion.

Now in general it is impossible to find f satisfying these equations exactly, so we proceed by approximations. We will be able to construct an approximate f because $(x_0, \xi_0) \notin \text{Char}\,(P)$. In constructing the approximation, keep in mind that we are interested in estimating (2.7) *for large* ξ. *In this regime we will be asymptotically* *inverting* P^T *on functions like the right-hand side of (2.8).* In this sense one can say that *an operator is microlocally invertible away from its characteristic variety.*

Let us now show that at least formally one can indeed solve (2.8) asymptotically. Leibnitz' rule implies that applying P^T to f given by (2.9) yields

$$P^T(v\, e^{-ix\cdot\xi})(x) \;=\; [\sigma_P(x,\xi)\, v(x) + \cdots]\, e^{-ix\cdot\xi}$$

where the dots denote lower order monomials in ξ. Therefore a first approximation to f satisfying (2.8) is

$$f_1(x) \;=\; \frac{\rho(x)}{\sigma_P(x,\xi)}\, e^{-ix\cdot\xi}.$$

This is well-defined provided (x,ξ) stays away from the characteristic variety of P. We henceforth choose ρ and a conic neighborhood V of ξ_0 such that:

1. $\rho(x_0) = 1$,
2. $\operatorname{supp}\rho \times V \cap \mathrm{WF}(Pu) = \emptyset$, and
3. For some $C > 0$, $\forall (x,\xi) \in \operatorname{supp}\rho \times V$ $|\sigma_P(x,\xi) \geq C|\xi|^m$, where m is the order of P.

The calculation above suggests defining an operator R by the formula:

$$P^T\left(\frac{w}{\sigma_P(x,\xi)}\, e^{-ix\cdot\xi}\right)(x) \;=\; (w(x) - R(w)(x))\; e^{-ix\cdot\xi}. \tag{2.10}$$

It should be understood that R depends on ξ. Indeed R has the form

$$R \;=\; R_1 + \cdots + R_m,$$

where $|\xi|^j R_j$ is a differential operator in x of order at most j with smooth coefficients which are homogeneous of degree zero with respect to ξ. [1] Notice that there is not an R_0 term and therefore the coefficients of R are $O(1/|xi|)$.

In terms of R we would like to solve the equation

$$w - Rw \;=\; \rho,$$

[1] *Example:* Let's take $P = \Delta$. A simple calculation shows that for any two functions a, b one has $\Delta(ab) = a\Delta b + b\Delta a - 2\nabla a\cdot\nabla b$. using this formula one easily checks that in this case

$$Rw \;=\; -\frac{\Delta w}{|\xi|^2} + \frac{i}{|\xi|^2}\,\xi\cdot\nabla w.$$

and therefore w will depend on ξ as well. A formal solution to the above equation is given by the Neumann series:

$$w = \sum_{l=0}^{\infty} R^l(\rho).$$

Of course this series need not converge, but as it turns out its partial sums provide the desired approximate solution. Thus let

$$w_N := \sum_{l=0}^{N-1} R^l(\rho).$$

Then we have exactly that $w_N - Rw_N = \rho - R^N(\rho)$. By definition of R, this means that

$$P^T \left(\frac{w_N}{\sigma_P(x,\xi)} e^{-ix\cdot\xi} \right)(x) = \left(\rho - R^N(\rho) \right) e^{-ix\cdot\xi}.$$

Pairing the above with u one gets

$$\widehat{\rho u}(\xi) = < Pu \, , \frac{w_N}{\sigma_P(x,\xi)} e^{-ix\cdot\xi} > + < u \, , R^N(\rho) e^{-ix\cdot\xi} > .$$

We now show that each term is rapidly decreasing on $\xi \in V$.

We can write

$$< Pu \, , \frac{w_N}{\sigma_P(x,\xi)} e^{-ix\cdot\xi} > = \widehat{fPu}$$

where $f = \dfrac{w_N}{\sigma_P(x,\xi)}$. Now apply Lemma (2.5.4). Since the C^K norms of this f are polynomial in ξ, this term is rapidly decreasing.

Finally, if M is the order of u, we have:

$$| < u \, , R^N(\rho) e^{-ix\cdot\xi} > | \leq C \sum_{|\alpha| \leq M} \sup |D^\alpha R^N(\rho) e^{-ix\cdot\xi}|.$$

But $\sup |D^\alpha R^N(\rho) e^{-ix\cdot\xi}| = O(|\xi|^{|\alpha|-N})$, so

$$| < u \, , R^N(\rho) e^{-ix\cdot\xi} > | \leq C_N (1 + |\xi|)^{M_N}$$

Since we can take N arbitrarily large this term is rapidly decreasing as well.

$\square$