Introduction to pseudo-differential operators

Michael Ruzhansky

January 21, 2014

Abstract

The present notes give introduction to the theory of pseudo-differential operators on Euclidean spaces. The first part is devoted to the necessary analysis of functions, such as basics of the Fourier analysis and the theory of distributions and Sobolev spaces. The second part is devoted to pseudo-differential operators and their applications to partial differential equations. We refer to the monograph [1] by Ruzhansky and Turunen for further details on this theory on the Euclidean space, torus, and more general compact Lie groups and homogeneous spaces. This book will be the main source of examples and further details to complement these notes. The course will contain the material from these notes and from [1], but not everything, with the rest of the material provided here for the reader's convenience.

The reader interested in the relations of this topic to the harmonic analysis may consult the monograph [3] by Stein, and those interested in the relations to the spectral theory can consult the monograph [2] by Shubin, to mention only very few related sources of material.

Contents

1 Analysis of functions

1.1 Basic properties of the Fourier transform

1.1.1 L^p–spaces

Let $\Omega \subset \mathbb{R}^n$ be a measurable subset of $\mathbb{R}^n$. For simplicity, we may always think of Ω being open or closed in $\mathbb{R}^n$. In this section we will mostly have $\Omega = \mathbb{R}^n$.

Let $1 \leq p < \infty$. A function $f : \Omega \to \mathbb{C}$ is said to be in $L^p(\Omega)$ if it is measurable and its norm

$$||f||_{L^p(\Omega)} = \left(\int_\Omega |f(x)|^p \, dx \right)^{1/p}$$

is finite. In particular, $L^1(\Omega)$ is the space of absolutely integrable functions on Ω with

$$||f||_{L^1(\Omega)} = \int_\Omega |f(x)| dx.$$

In the case $p = \infty$, it is said that $f \in L^\infty(\Omega)$ if it is measurable and essentially bounded, i.e. if

$$||f||_{L^\infty(\Omega)} = \operatorname{esssup}_{x \in \Omega} |f(x)|$$

is finite. Here $\operatorname{esssup}_{x\in\Omega} |f(x)|$ is defined as the smallest M such that $|f(x)| \leq M$ for almost all $x \in \Omega$. We will often abbreviate the $||f||_{L^p(\Omega)}$ norm by $||f||_{L^p}$, or by $||f||_p$, if the choice of Ω is clear from the context.

1.1.2 Definition of the Fourier transform

For $f \in L^1(\mathbb{R}^n)$ we define its Fourier transform by

$$(\mathcal{F}f)(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} f(x) dx.$$

In fact, other similar definitions are often encountered in the literature. For example, one can use $e^{-ix\cdot\xi}$ instead of $e^{-2\pi i x\cdot\xi}$, multiply the integral by the constant $(2\pi)^{-n/2}$, etc. Changes in definitions may lead to changes in constants in formulas.

1.1.3 Fourier transform in $L^1(\mathbb{R}^n)$

It is easy to check that $\mathcal{F} : L^1(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$ is a bounded linear operator with norm one:

$$||\widehat{f}||_\infty \leq ||f||_1.$$

Moreover, if $f \in L^1(\mathbb{R}^n)$, its Fourier transform $\widehat{f}$ is continuous, which follows from the Lebesgue's dominated convergence theorem. For completeness, let us state it here:

Lebesgue's dominated convergence theorem. *Let $\{f_k\}_{k=1}^\infty$ be a sequence of measurable functions on Ω such that $f_k \to f$ pointwise almost everywhere on Ω as $k \to \infty$. Suppose there is an integrable function $g \in L^1(\Omega)$ such that $|f_k| \leq g$ for all k. Then f is integrable and*

$$\int_\Omega f dx = \lim_{k\to\infty} \int_\Omega f_k\, dx.$$

1.1.4 Riemann-Lebesgue theorem

It is quite difficult to characterise the image of the space $L^1(\mathbb{R}^n)$ under the Fourier transform. But we have the following theorem.

Riemann-Lebesgue Theorem. *Let $f \in L^1(\mathbb{R}^n)$. Then its Fourier transform $\widehat{f}$ is a continuous function on $\mathbb{R}^n$ vanishing at infinity, i.e. $\widehat{f}(\xi) \to 0$ as $\xi \to \infty$.*

Proof. It is enough to make an explicit calculation for f being a characteristic function of a cube and then use a standard limiting argument from the measure theory. Thus, let f be a characteristic function of the unit cube, i.e. $f(x) = 1$ for $x \in [-1, 1]^n$ and $f(x) = 0$ otherwise. Then

$$\begin{aligned}\widehat{f}(\xi) &= \int_{[-1,1]^n} e^{-2\pi i x\cdot\xi} dx = \prod_{j=1}^n \int_{-1}^1 e^{-2\pi i x_j \xi_j} dx_j = \prod_{j=1}^n \frac{1}{-2\pi i \xi_j} e^{-2\pi i x_j \xi_j}|_{-1}^1 \\ &= \left(\frac{i}{2\pi}\right)^n \frac{1}{\xi_1 \cdots \xi_n} \prod_{j=1}^n \left[e^{-2\pi i \xi_j} - e^{2\pi i \xi_j}\right].\end{aligned}$$

The product of exponents is bounded, so the whole expression tends to zero as $\xi \to \infty$ away from coordinate axis. In the case some of ξ_j's are zero, an obvious modification of this argument yields the same result.

1.1.5 Schwartz space

We define the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ of rapidly decreasing functions as follows. First, for multi-indices $\alpha = (\alpha_1, \cdots, \alpha_n)$ and $\beta = (\beta_1, \cdots, \beta_n)$ with integer entries $\alpha_j, \beta_j \geq 0$, we define

$$\partial^\alpha \varphi(x) = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \varphi(x) \text{ and } x^\beta = x_1^{\beta_1} \cdots x_n^{\beta_n}.$$

For such multi-indices we will write $\alpha, \beta \geq 0$. Then we say that $\varphi \in \mathcal{S}(\mathbb{R}^n)$ if φ is smooth on $\mathbb{R}^n$ and

$$\sup_{x\in\mathbb{R}^n} \left|x^\beta \partial^\alpha \varphi(x)\right| < \infty$$

holds for all multi-indices $\alpha, \beta \geq 0$. The length of the multi-index α will be denoted by $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

It is easy to see that a smooth function f is in the Schwartz space if and only if for all $\alpha \geq 0$ and $N \geq 0$ there is a constant $C_{\alpha,N}$ such that

$$|\partial^\alpha \varphi(x)| \leq C_{\alpha,N}(1+|x|)^{-N}$$

holds for all $x \in \mathbb{R}^n$.

The space $\mathcal{S}(\mathbb{R}^n)$ is a topological space. Without going into detail, let us only introduce the convergence of functions in $\mathcal{S}(\mathbb{R}^n)$. We will say that $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ as $j \to \infty$, if $\varphi_j, \varphi \in \mathcal{S}(\mathbb{R}^n)$ and if

$$\sup_{x \in \mathbb{R}^n} |x^\beta \partial^\alpha (\varphi_j - \varphi)(x)| \to 0$$

as $j \to \infty$, for all multi-indices $\alpha, \beta \geq 0$.

If one is familiar with functional analysis, one can take expressions

$$\sup_{x \in \mathbb{R}^n} |x^\beta \partial^\alpha (\varphi_j - \varphi)(x)|$$

as seminorms on the space $\mathcal{S}(\mathbb{R}^n)$, turning it into a locally convex linear topological space.

1.1.6 Differentiation and multiplication

Since the definition of the Fourier transform contains the complex exponential, it is often convenient to use the notation

$$D_j = \frac{1}{2\pi i}\frac{\partial}{\partial x_j} \text{ and } D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}.$$

If D_j is applied to a function of ξ it will obviously mean $\frac{1}{2\pi i}\frac{\partial}{\partial \xi_j}$. However, there should be no confusion with this convention. The following theorem relates multiplication with differentiation with respect to the Fourier transform.

Theorem. *Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then $\widehat{D_j\varphi}(\xi) = \xi_j \widehat{\varphi}(\xi)$ and $\widehat{x_j\varphi}(\xi) = -D_j\widehat{\varphi}(\xi)$.*

Proof. From the definition we readily see that

$$D_j\widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi}(-x_j)\varphi(x)dx.$$

This gives the second formula. Since the integrals converge uniformly, we can integrate by parts with respect to x_j in the following expression to get

$$\xi_j\widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} \left(-D_j e^{-2\pi i x\cdot\xi}\right)\varphi(x)dx = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} D_j\varphi(x)dx.$$

This implies the first formula. Note that we do not get boundary terms when integrating by parts because function φ vanishes at infinity.

This theorem allows one to tackle some differential equations already. For example, let us look at the equation

$$\Delta u = f$$

with the Laplace operator

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}.$$

Taking the Fourier transform and using the theorem we arrive at the equation

$$-4\pi^2|\xi|^2\widehat{u} = \widehat{f}.$$

If we knew how to invert the Fourier transform we could find the solution

$$u = -\mathcal{F}^{-1}\left(\frac{1}{4\pi^2|\xi|^2}\widehat{f}\right).$$

1.1.7 Fourier transform in the Schwartz space

Corollary. *Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then*

$$\xi^\beta D^\alpha \widehat{\varphi}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} D^\beta((-x)^\alpha \varphi(x))dx.$$

Hence also

$$\sup_{\xi\in\mathbb{R}^n} \left|\xi^\beta D^\alpha \widehat{\varphi}(\xi)\right| \leq C \sup_{x\in\mathbb{R}^n} \left((1+|x|)^{n+1}\left|D^\beta(x^\alpha\varphi(x))\right|\right),$$

with $C = \int_{\mathbb{R}^n}(1+|x|)^{-n-1}dx < \infty$.

Here we used the following useful

Integrability criterion. We have

$$\int_{\mathbb{R}^n} \frac{dx}{(1+|x|)^\rho} < \infty \text{ if and only if } \rho > n.$$

We also have

$$\int_{|x|\leq 1} \frac{dx}{|x|^\rho} < \infty \text{ if and only if } \rho < n.$$

Both of these criteria can be easily seen by passing to the polar coordinates.

Corollary 1.1.7 implies that the Fourier transform $\mathcal{F}$ maps $\mathcal{S}(\mathbb{R}^n)$ to itself. In fact, much more is true (see the next paragraph). Let us note here that Corollary 1.1.7 together with the Lebesgue's dominated convergence theorem imply that the Fourier transform $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is continuous, i.e. $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ implies $\widehat{\varphi_j} \to \widehat{\varphi}$ in $\mathcal{S}(\mathbb{R}^n)$.

1.1.8 Fourier inversion formula

Theorem. *The Fourier transform* $\mathcal{F} : \varphi \mapsto \widehat{\varphi}$ *is an isomorphism of* $\mathcal{S}(\mathbb{R}^n)$ *into* $\mathcal{S}(\mathbb{R}^n)$, *whose inverse is given by*

$$\varphi(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} \widehat{\varphi}(\xi) d\xi.$$

This formula is called *the Fourier inversion formula* and *the inverse Fourier transform* is denoted by

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

Thus, we can say that

$$\mathcal{F} \circ \mathcal{F}^{-1} = \mathcal{F}^{-1} \circ \mathcal{F} = identity \quad \text{on} \quad \mathcal{S}(\mathbb{R}^n).$$

The proof of this theorem will rely on several lemmas which have a significance on their own.

1.1.9 Multiplication formula for the Fourier transform

Theorem. *Let* $f, g \in L^1(\mathbb{R}^n)$. *Then* $\int_{\mathbb{R}^n} \widehat{f} g dx = \int_{\mathbb{R}^n} f \widehat{g} dx$.

Proof. We will apply the Fubini formula. Thus,

$$\int_{\mathbb{R}^n} \widehat{f} g dx = \int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} e^{-2\pi i x\cdot y} f(y) dy \right] g(x) dx =$$

$$\int_{\mathbb{R}^n} \left[\int_{\mathbb{R}^n} e^{-2\pi i x\cdot y} g(x) dx \right] f(y) dy = \int_{\mathbb{R}^n} \widehat{g} f dy.$$

1.1.10 Fourier transform of Gaussian distributions

We will show the equality

$$\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} e^{-2\epsilon\pi^2 |x|^2} dx = (2\pi\epsilon)^{-n/2} e^{-|\xi|^2/(2\epsilon)}. \tag{1.1}$$

By the change of $2\pi x \to x$ it is equivalent to

$$\int_{\mathbb{R}^n} e^{-ix\cdot\xi} e^{-\epsilon|x|^2/2} dx = \left(\frac{2\pi}{\epsilon}\right)^{n/2} e^{-|\xi|^2/(2\epsilon)}. \tag{1.2}$$

We will use the standard identities

$$\int_{-\infty}^{\infty} e^{-t^2/2} dt = \sqrt{2\pi} \quad \text{and hence} \quad \int_{\mathbb{R}^n} e^{-|x|^2/2} dx = (2\pi)^{n/2}. \tag{1.3}$$

In fact, (1.2) will follow from the one-dimensional case, where we have

$$\int_{-\infty}^{\infty} e^{-it\tau} e^{-t^2/2} dt = e^{-\tau^2/2} \int_{-\infty}^{\infty} e^{-(t+i\tau)^2/2} dt = e^{-\tau^2/2} \int_{-\infty}^{\infty} e^{-t^2/2} dt = \sqrt{2\pi} e^{-\tau^2/2},$$

where we used the Cauchy theorem about changing the contour of integration for analytic functions. Changing $t \to \sqrt{\epsilon} t$ and $\tau \to \tau/\sqrt{\epsilon}$ gives

$$\sqrt{\epsilon} \int_{-\infty}^{\infty} e^{-it\tau} e^{-\epsilon t^2/2} dt = \sqrt{2\pi} e^{-\tau^2/(2\epsilon)}.$$

Extending this to n dimensions yields (1.2).

1.1.11 Proof of the Fourier inversion formula

Now we can prove the Fourier inversion formula 1.1.8. So, for $\varphi \in \mathcal{S}(\mathbb{R}^n)$, we want to prove that

$$\varphi(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\widehat{\varphi}(\xi)d\xi.$$

By the Lebesgue's dominated convergence theorem we can replace the right hand side by

$$\begin{aligned}
RHS &= \lim_{\epsilon\to 0}\int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\widehat{\varphi}(\xi)e^{-2\epsilon\pi^2|\xi|^2}d\xi \\
&= \lim_{\epsilon\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\xi}\varphi(y)e^{-2\epsilon\pi^2|\xi|^2}dyd\xi \quad (\text{change } y\to y+x) \\
&= \lim_{\epsilon\to 0}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-2\pi i y\cdot\xi}\varphi(y+x)e^{-2\epsilon\pi^2|\xi|^2}dyd\xi \quad (\text{Fubini's theorem}) \\
&= \lim_{\epsilon\to 0}\int_{\mathbb{R}^n}\varphi(y+x)dy\left(\int_{\mathbb{R}^n} e^{-2\pi i y\cdot\xi}e^{-2\epsilon\pi^2|\xi|^2}d\xi\right) \quad (\text{F.T. of Gaussian}) \\
&= \lim_{\epsilon\to 0}\int_{\mathbb{R}^n}\varphi(y+x)(2\pi\epsilon)^{-n/2}e^{-|y|^2/(2\epsilon)}dy \quad (\text{change } y=\sqrt{\epsilon}z) \\
&= (2\pi)^{-n/2}\lim_{\epsilon\to 0}\int_{\mathbb{R}^n}\varphi(\sqrt{\epsilon}z+x)e^{-|z|^2/2}dz \\
&= \varphi(x).
\end{aligned}$$

This finishes the proof.

Remark. In fact, in the proof we implicitly made use of the following useful relation between Fourier transforms and translations of functions. Let $h \in \mathbb{R}^n$ and define $(\tau_h f)(x) = f(x-h)$. Then we also see that

$$\begin{aligned}
\widehat{(\tau_h f)}(\xi) &= \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi}(\tau_h f)(x)dx \\
&= \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi}f(x-h)dx \quad (\text{change } y = x-h) \\
&= \int_{\mathbb{R}^n} e^{-2\pi i(y+h)\cdot\xi}f(y)dy \\
&= e^{-2\pi i h\cdot\xi}\widehat{f}(\xi).
\end{aligned}$$

1.1.12 Convolutions

For functions $f, g \in L^1(\mathbb{R}^n)$, we define their *convolution* by

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

It is easy to see that $f*g \in L^1(\mathbb{R}^n)$ with norm inequality

$$||f*g||_{L^1(\mathbb{R}^n)} \leq ||f||_{L^1(\mathbb{R}^n)}||g||_{L^1(\mathbb{R}^n)}$$

and that

$$f * g = g * f.$$

Also, in particular for $f, g \in \mathcal{S}(\mathbb{R}^n)$, integrals are absolutely convergent and we can differentiate under the integral sign to get

$$\partial^\alpha (f * g) = \partial^\alpha f * g = f * \partial^\alpha g.$$

The following properties relate convolutions with Fourier transforms.

Theorem. *Let $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$. Then we have*

(i) $\int_{\mathbb{R}^n} \varphi \overline{\psi} dx = \int_{\mathbb{R}^n} \widehat{\varphi} \overline{\widehat{\psi}} d\xi$;

(ii) $\widehat{\varphi * \psi}(\xi) = \widehat{\varphi}(\xi)\widehat{\psi}(\xi)$;

(iii) $\widehat{\varphi\psi}(\xi) = (\widehat{\varphi} * \widehat{\psi})(\xi)$.

Proof. (i) Let us denote

$$\chi(\xi) = \overline{\widehat{\psi}}(\xi) = \int_{\mathbb{R}^n} e^{2\pi i x \cdot \xi} \overline{\psi}(x) dx = \mathcal{F}^{-1}(\overline{\psi})(\xi),$$

so that $\widehat{\chi} = \overline{\psi}$. It follows now that

$$\int_{\mathbb{R}^n} \varphi \overline{\psi} = \int_{\mathbb{R}^n} \varphi \widehat{\chi} = \int_{\mathbb{R}^n} \widehat{\varphi} \chi = \int_{\mathbb{R}^n} \widehat{\varphi} \overline{\widehat{\psi}},$$

where we used the multiplication formula for the Fourier transform.
(ii) We can easily calculate

$$\begin{aligned}
\widehat{\varphi * \psi}(\xi) &= \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} (\varphi * \psi)(x) dx = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} \varphi(x - y) \psi(y) dy dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-2\pi i (x-y) \cdot \xi} \varphi(x - y) e^{-2\pi i y \cdot \xi} \psi(y) dy dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-2\pi i z \cdot \xi} \varphi(z) e^{-2\pi i y \cdot \xi} \psi(y) dy dz = \widehat{\varphi}(\xi) \widehat{\psi}(\xi),
\end{aligned}$$

where we used the substitution $z = x - y$. We leave (iii) as an exercise.

1.2 Inequalities

This section will be devoted to several important inequalities which are very useful in Fourier analysis and in many types of analysis involving spaces of functions.

1.2.1 Cauchy's inequality.

For all $a, b \in \mathbb{R}$ we have

$$ab \le \frac{a^2}{2} + \frac{b^2}{2}.$$

Moreover, for any $\epsilon > 0$, we also have

$$ab \le \epsilon a^2 + \frac{b^2}{4\epsilon}.$$

The first inequality follows from $0 \le (a-b)^2 = a^2 - 2ab + b^2$. The second one follows in we apply the first one to $ab = (\sqrt{2\epsilon}a)(b/\sqrt{2\epsilon})$.

As a consequence, we immediately obtain the Cauchy's inequality for functions:

$$\int_\Omega |f(x)g(x)|dx \le \frac{1}{2}\int_\Omega (|f(x)|^2 + |g(x)|^2)dx \quad \text{and} \quad ||fg||_{L^1(\Omega)} \le \frac{1}{2}\Big(||f||^2_{L^2(\Omega)} + ||g||^2_{L^2(\Omega)}\Big).$$

1.2.2 Cauchy–Schwartz's inequality.

Let $x, y \in \mathbb{R}^n$. Then $|x \cdot y| \le |x||y|$.

Proof. For $\epsilon > 0$, we have $0 \le |x \pm \epsilon y|^2 = |x|^2 \pm 2\epsilon x \cdot y + \epsilon^2 |y|^2$. This implies $\pm x \cdot y \le \frac{1}{2\epsilon}|x|^2 + \frac{\epsilon}{2}|y|^2$. Setting $\epsilon = \frac{|x|}{|y|}$, we obtain the required inequality, provided $y \ne 0$ (if $y = 0$ it is trivial).

An alternative proof may be given as follows. We can observe that the inequality

$$0 \le |x + \epsilon y|^2 = |x|^2 + 2\epsilon x \cdot y + \epsilon^2 |y|^2$$

implies that the discriminant of the quadratic (in ϵ) polynomial on the right hand side must be non-positive, which means $|x \cdot y|^2 - |x|^2|y|^2 \le 0$.

1.2.3 Young's inequality.

Let $1 < p, q < \infty$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \quad \textit{for all} \quad a, b > 0.$$

Moreover, if $\epsilon > 0$, we have

$$ab \le \epsilon a^p + C(\epsilon) b^q$$

for all $a, b > 0$, where $C(\epsilon) = (\epsilon p)^{-q/p} q^{-1}$.

To prove the first inequality, we will use the fact that the exponential function $x \mapsto e^x$ is convex (a function $f : \mathbb{R} \to \mathbb{R}$ is called convex if

$$f(\tau x + (1-\tau)y) \le \tau f(x) + (1-\tau) f(y),$$

for all $x, y \in \mathbb{R}$ and all $0 \le \tau \le 1$). This implies

$$ab = e^{\ln a + \ln b} = e^{\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q} \le \frac{1}{p}e^{\ln a^p} + \frac{1}{q}e^{\ln b^q} = \frac{a^p}{p} + \frac{b^q}{q}.$$

The second inequality with ϵ follows if we apply the first one to the product $ab = \big((\epsilon p)^{1/p} a\big)\big(b/(\epsilon p)^{1/p}\big)$.

As a consequence, we immediately obtain that if $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$, then $fg \in L^1(\Omega)$ with

$$||fg||_{L^1} \le \frac{1}{p}||f||^p_{L^p} + \frac{1}{q}||g||^q_{L^q}.$$

1.2.4 Hölder's inequality. *Let* $1 \le p, q \le \infty$ *with* $\frac{1}{p} + \frac{1}{q} = 1$. *Let* $f \in L^p(\Omega)$ *and* $g \in L^q(\Omega)$. *Then* $fg \in L^1(\Omega)$ *and*

$$||fg||_{L^1(\Omega)} = \int_\Omega |fg| dx \le ||f||_{L^p(\Omega)} ||g||_{L^q(\Omega)}.$$

In the formulation we use the standard convention of setting $1/\infty = 0$. In the case of $p = q = 2$ Hölder's inequality is often called the Cauchy–Schwartz' inequality.

Proof. In the case $p = 1$ or $p = \infty$ the inequality is obvious, so let us assume $1 < p < \infty$. Let us first consider the case when $||f||_{L^p} = ||g||_{L^q} = 1$. Then by Young's inequality with $1 < p, q < \infty$, we have

$$||fg||_{L^1} \le \frac{1}{p}||f||_{L^p}^p + \frac{1}{q}||g||_{L^q}^q = \frac{1}{p} + \frac{1}{q} = 1 = ||f||_{L^p}||g||_{L^q},$$

which is the desired inequality. Now, let us consider general f, g. We observe that we may assume that $||f||_{L^p} \neq 0$ and $||g||_{L^q} \neq 0$, since otherwise one of functions is zero almost everywhere in Ω and Hölder's inequality becomes trivial. It follows from the considered case that

$$\int_\Omega \left| \frac{f}{||f||_p} \frac{g}{||g||_q} \right| dx \le 1,$$

which implies the general case by the linearity of the integral.

General Hölder's inequality. *Let* $1 \le p_1, \ldots, p_m \le \infty$ *be such that* $\frac{1}{p_1} + \cdots + \frac{1}{p_m} = 1$. *Let* $f_k \in L^{p_k}(\Omega)$ *for all* $k = 1, \ldots, m$. *Then the product* $f_1 \cdots f_m \in L^1(\Omega)$ *and*

$$||f_1 \cdots f_m||_{L^1(\Omega)} \le \prod_{k=1}^m ||f_k||_{L^{p_k}(\Omega)}.$$

This inequality follows from Hölder's inequality by induction on the number of functions.

1.2.5 Minkowski's inequality. *Let* $1 \le p \le \infty$. *Let* $f, g \in L^p(\Omega)$. *Then*

$$||f + g||_{L^p(\Omega)} \le ||f||_{L^p(\Omega)} + ||g||_{L^p(\Omega)}.$$

In particular, this means that $|| \cdot ||_{L^p}$ satisfies the triangle inequality and is a norm, so $L^p(\Omega)$ is a normed space.

Proof. The cases of $p = 1$ or $p = \infty$ follow from the triangle inequality for complex numbers and are, therefore, trivial. So we may assume $1 < p < \infty$. Then we have

$$\begin{aligned}
||f + g||_{L^p(\Omega)}^p &= \int_\Omega |f + g|^p dx \le \int_\Omega |f + g|^{p-1}(|f| + |g|) dx \\
&= \int_\Omega |f + g|^{p-1}|f| dx + \int_\Omega |f + g|^{p-1}|g| dx \quad (\text{Hölder's ineq } p = p, q = \frac{p}{p-1}) \\
&\le \left(\int_\Omega |f + g|^p dx\right)^{\frac{p-1}{p}} \left[\left(\int_\Omega |f|^p dx\right)^{\frac{1}{p}} + \left(\int_\Omega |g|^p dx\right)^{\frac{1}{p}}\right] \\
&= ||f + g||_{L^p(\Omega)}^{p-1} \big(||f||_{L^p(\Omega)} + ||g||_{L^p(\Omega)}\big),
\end{aligned}$$

which implies the desired inequality.

1.2.6 Interpolation for L^p–norms. *Let $1 \le s \le r \le t \le \infty$ be such that $\frac{1}{r} = \frac{\theta}{s} + \frac{1-\theta}{t}$ for some $0 \le \theta \le 1$. Let $f \in L^s(\Omega) \bigcap L^t(\Omega)$. Then $f \in L^r(\Omega)$ and*

$$||f||_{L^r(\Omega)} \le ||f||^{\theta}_{L^s(\Omega)} ||f||^{1-\theta}_{L^t(\Omega)}.$$

To prove this, we use that $\frac{\theta r}{s} + \frac{(1-\theta)r}{t} = 1$ and so we can apply Hölder's inequality in the following way:

$$\int_\Omega |f|^r dx = \int_\Omega |f|^{\theta r} |f|^{(1-\theta)r} dx \le \left(\int_\Omega |f|^{\theta r \cdot \frac{s}{\theta r}} dx \right)^{\frac{\theta r}{s}} \left(\int_\Omega |f|^{(1-\theta)r \cdot \frac{t}{(1-\theta)r}} dx \right)^{\frac{(1-\theta)r}{t}},$$

which is the desired inequality.

1.3 Fourier transforms of distributions

In this section we will introduce several spaces of distributions and will extend the Fourier transform to more general spaces of functions than $\mathcal{S}(\mathbb{R}^n)$ or $L^1(\mathbb{R}^n)$ in the first section. The main problem with the immediate extension is that the integral in the definition of the Fourier transform may no longer converge if we go beyond the space $L^1(\mathbb{R}^n)$ of integrable functions.

1.3.1 Tempered distributions

We define the *space of tempered distributions* $\mathcal{S}'(\mathbb{R}^n)$ as the space of all continuous linear functionals on $\mathcal{S}(\mathbb{R}^n)$. This means that $u \in \mathcal{S}'(\mathbb{R}^n)$ if it is a functional $u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ such that

1. u is linear, i.e. $u(\alpha\varphi + \beta\psi) = \alpha u(\varphi) + \beta u(\psi)$ for all $\alpha, \beta \in \mathbb{C}$ and all $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n)$;
2. u is continuous, i.e. $u(\varphi_j) \to u(\varphi)$ in $\mathbb{C}$ whenever $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$.

Here one should also recall the definition of the convergence $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ from 1.1.5, which said that $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ as $j \to \infty$, if $\varphi_j, \varphi \in \mathcal{S}(\mathbb{R}^n)$ and if

$$\sup_{x \in \mathbb{R}^n} |x^\beta \partial^\alpha (\varphi_j - \varphi)(x)| \to 0$$

as $j \to \infty$, for all multi-indices $\alpha, \beta \ge 0$.

We can also define the *convergence in the space* $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions. Let $u_j, u \in \mathcal{S}'(\mathbb{R}^n)$. We will say that $u_j \to u$ in $\mathcal{S}'(\mathbb{R}^n)$ as $j \to \infty$ if $u_j(\varphi) \to u(\varphi)$ in $\mathbb{C}$ as $j \to \infty$, for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

Functions in $\mathcal{S}(\mathbb{R}^n)$ are called the *test functions* for tempered distributions in $\mathcal{S}'(\mathbb{R}^n)$. Another notation for $u(\varphi)$ will be $\langle u, \varphi \rangle$.

1.3.2 Fourier transform of tempered distributions

If $u \in \mathcal{S}'(\mathbb{R}^n)$, we can define its (generalised) Fourier transforms by setting

$$\widehat{u}(\varphi) = u(\widehat{\varphi}),$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

Then we can readily see that also $\widehat{u} \in \mathcal{S}'(\mathbb{R}^n)$. Indeed, since $\varphi \in \mathcal{S}$, it follows that $\widehat{\varphi} \in \mathcal{S}$ and so $u(\widehat{\varphi})$ is a well-defined complex number. Moreover, $\widehat{u}$ is linear since both u and the Fourier transform $\mathcal{F}$ are linear. Finally, $\widehat{u}$ is continuous because $\varphi_j \to \varphi$ in $\mathcal{S}$ implies $\widehat{\varphi}_j \to \widehat{\varphi}$ in $\mathcal{S}$, and hence

$$\widehat{u}(\varphi_j) = u(\widehat{\varphi}_j) \to u(\widehat{\varphi}) = \widehat{u}(\varphi)$$

by the continuity of both u from $\mathcal{S}(\mathbb{R}^n)$ to $\mathbb{C}$ and of the Fourier transform $\mathcal{F}$ as a mapping from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$ (see 1.1.7).

Now, it follows that it is also continuous as a mapping from $\mathcal{S}'(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$, i.e. $u_j \to u$ in $\mathcal{S}'(\mathbb{R}^n)$ implies that $\widehat{u_j} \to \widehat{u}$ in $\mathcal{S}'(\mathbb{R}^n)$. Indeed, if $u_j \to u$ in $\mathcal{S}'(\mathbb{R}^n)$, we have

$$\widehat{u_j}(\varphi) = u_j(\widehat{\varphi}) \to u(\widehat{\varphi}) = \widehat{u}(\varphi)$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$, which means that $\widehat{u_j} \to \widehat{u}$ in $\mathcal{S}'(\mathbb{R}^n)$.

1.3.3 Two principles for tempered distributions

Here we give two immediate but important principles for distributions.

Convergence principle. *Let X be a topological subspace in $\mathcal{S}'(\mathbb{R}^n)$ (i.e. convergence in X implies convergence in $\mathcal{S}'(\mathbb{R}^n)$). Suppose that $u_j \to u$ in $\mathcal{S}'(\mathbb{R}^n)$ and that $u_j \to v$ in X. Then $u \in X(\mathbb{R}^n)$ and $u = v$.*

This statement is simply the consequence of the fact that the space $\mathcal{S}'(\mathbb{R}^n)$ is Hausdorff, hence it has the uniqueness of limits property (recall that a topological space is called Hausdorff if any two points have open disjoint neighborhoods, i.e. open disjoint sets containing them). The convergence principle is also related to another principle which we call

Uniqueness principle for distributions. *Let $u, v \in \mathcal{S}'(\mathbb{R}^n)$ and suppose that $u(\varphi) = v(\varphi)$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then $u = v$.*

This can be reformulated by saying that if an element $o \in \mathcal{S}'(\mathbb{R}^n)$ satisfies $o(\varphi) = 0$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then o is the zero element in $\mathcal{S}'(\mathbb{R}^n)$.

1.3.4 Functions as distributions

We can interpret functions in $L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, as tempered distributions. If $f \in L^p(\mathbb{R}^n)$, we define the functional u_f by

$$u_f(\varphi) = \int_{\mathbb{R}^n} f(x)\varphi(x)dx,$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. By Hölder's inequality, we observe that

$$|u_f(\varphi)| \leq ||f||_{L^p} ||\varphi||_{L^q},$$

for $\frac{1}{p} + \frac{1}{q} = 1$, and hence $u_f(\varphi)$ is well-defined in view of the simple inclusion $\mathcal{S}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$, for all $1 \leq q \leq \infty$. It needs to be verified that u_f is a linear continuous functional on $\mathcal{S}(\mathbb{R}^n)$. It is clearly linear in φ, while its continuity follows from inequality

$$|u_f(\varphi_j) - u_f(\varphi)| \leq ||f||_{L^p} ||\varphi_j - \varphi||_{L^q}$$

and

Lemma. *We have $\mathcal{S}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ with continuous embedding, i.e. $\varphi_j \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ implies that $\varphi_j \to \varphi$ in $L^q(\mathbb{R}^n)$.*

To summarise, any function $f \in L^p(\mathbb{R}^n)$ leads to a tempered distribution $u_f \in \mathcal{S}'(\mathbb{R}^n)$ in the canonical way described above. In this way we will view functions in $L^p(\mathbb{R}^n)$ as tempered distributions and continue to simply write f instead of u_f. There should be no confusion with this notation since writing $f(x)$ suggests that f is a function while $f(\varphi)$ suggests that it is applied to test functions and so it is viewed as a distribution u_f.

Consistency of all definitions. With this identification, the definition of the Fourier transform for functions in $L^1(\mathbb{R}^n)$ agrees with the definition of the Fourier transforms of tempered distributions. Indeed, let $f \in L^1(\mathbb{R}^n)$. Then we have two ways of looking at its Fourier transforms:

1. We can use the first definition $\widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i x \cdot \xi} f(x) dx$, and then we know that $\widehat{f} \in L^\infty(\mathbb{R}^n)$. In this way we get $u_{\widehat{f}} \in \mathcal{S}'(\mathbb{R}^n)$.

2. We can immediately think of $f \in L^1(\mathbb{R}^n)$ as of $u_f \in \mathcal{S}'(\mathbb{R}^n)$, and the second definition then produces its Fourier transform $\widehat{u_f} \in \mathcal{S}'(\mathbb{R}^n)$.

Fortunately, these two approaches are consistent and produce the same tempered distribution $u_{\widehat{f}} = \widehat{u_f} \in \mathcal{S}'(\mathbb{R}^n)$. Indeed, we have

$$\widehat{u}(\varphi) = \int_{\mathbb{R}^n} \widehat{u}\varphi \, dx = \int_{\mathbb{R}^n} u\widehat{\varphi} \, dx = u(\widehat{\varphi}).$$

Here we used the multiplication formula for the Fourier transform and the fact that both $u \in L^1(\mathbb{R}^n)$ and $\widehat{u} \in L^\infty(\mathbb{R}^n)$ can be viewed as tempered distributions in the canonical way. It follows that we have $\widehat{u}(\varphi) = u(\widehat{\varphi})$ which was exactly the definition in 1.3.2.

Finally, we note that if $u \in L^1(\mathbb{R}^n)$ and also $\widehat{u} \in L^1(\mathbb{R}^n)$, then the Fourier inversion formula 1.1.8 holds for almost all $x \in \mathbb{R}^n$. A more general Fourier inversion formula for tempered distributions will be given in 1.3.7.

1.3.5 Plancherel's formula

It turns out that the Fourier transform acts especially nicely on one of the spaces $L^p(\mathbb{R}^n)$, namely on the space $L^2(\mathbb{R}^n)$, which is also a Hilbert space. These two facts lead to a very rich Fourier analysis on $L^2(\mathbb{R}^n)$ which we will deal with only briefly.

Theorem. *Let* $u \in L^2(\mathbb{R}^n)$*. Then* $\widehat{u} \in L^2(\mathbb{R}^n)$ *and*

$$||\widehat{u}||_{L^2(\mathbb{R}^n)} = ||u||_{L^2(\mathbb{R}^n)} \qquad \text{(Plancherel's formula)}$$

Moreover, for all $u, v \in L^2(\mathbb{R}^n)$ *we have*

$$\int_{\mathbb{R}^n} u\bar{v}dx = \int_{\mathbb{R}^n} \widehat{u}\bar{\widehat{v}}d\xi \qquad \text{(Parseval's formula)}$$

Proof. We will use the fact (a special case of the fact to be proved later) that $\mathcal{S}(\mathbb{R}^n)$ is sequentially dense in $L^2(\mathbb{R}^n)$, i.e. for every $u \in L^2(\mathbb{R}^n)$ there exists a sequence $u_j \in \mathcal{S}(\mathbb{R}^n)$ such that $u_j \to u$ in $L^2(\mathbb{R}^n)$. Then Theorem 1.1.12, (i), with $\varphi = \psi = u_j - u_k$, implies that

$$||\widehat{u_j} - \widehat{u_k}||^2_{L^2} = ||u_j - u_k||^2_{L^2} \to 0,$$

since u_j is a convergent sequence in $L^2(\mathbb{R}^n)$. Thus, $\widehat{u_j}$ is a Cauchy sequence in the complete (Banach) space $L^2(\mathbb{R}^n)$. It follows that it must converge to some $v \in L^2(\mathbb{R}^n)$. By the continuity of the Fourier transform in $\mathcal{S}'(\mathbb{R}^n)$ we must also have $\widehat{u_j} \to \widehat{u}$ in $\mathcal{S}'(\mathbb{R}^n)$. By the convergence principle for distributions, we get that $\widehat{u} = v \in L^2(\mathbb{R}^n)$. Applying Theorem 1.1.12, (i), again, to $\varphi = \psi = u_j$, we get

$$||\widehat{u_j}||^2_{L^2} = ||u_j||^2_{L^2}.$$

Passing to the limit, we get

$$||\widehat{u}||^2_{L^2} = ||u||^2_{L^2},$$

which is the Plancherel's formula.

Finally, for $u, v \in L^2(\mathbb{R}^n)$, let $u_j, v_j \in \mathcal{S}(\mathbb{R}^n)$ be such that $u_j \to u$ and $v_j \to v$ in $L^2(\mathbb{R}^n)$. Applying Theorem 1.1.12, (i), to $\varphi = u_j, \psi = v_j$, and passing to the limit, we obtain the Parseval's identity.

1.3.6 Operations with distributions

Besides the Fourier transform, there are several other operations than can be extended from functions in $\mathcal{S}(\mathbb{R}^n)$ to tempered distributions in $\mathcal{S}'(\mathbb{R}^n)$.

For example, partial differentiation $\frac{\partial}{\partial x_j}$ can be extended to a continuous operator $\frac{\partial}{\partial x_j} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$. Indeed, for $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, let us define

$$\left(\frac{\partial}{\partial x_j} u\right)(\varphi) = -u\left(\frac{\partial \varphi}{\partial x_j}\right).$$

It is necessary to include the negative sign in this definition. Indeed, if $u \in \mathcal{S}(\mathbb{R}^n)$, then the integration by parts formula and the identification of functions with distributions 1.3.4 yield

$$\left(\frac{\partial}{\partial x_j} u\right)(\varphi) = \int_{\mathbb{R}^n} \left(\frac{\partial u}{\partial x_j}\right)(x)\varphi(x)dx = -\int_{\mathbb{R}^n} u(x)\left(\frac{\partial \varphi}{\partial x_j}\right)(x)dx = -u\left(\frac{\partial \varphi}{\partial x_j}\right),$$

which explains the sign. This also shows the consistence of this definition of the derivative with the usual definition for differentiable functions. More generally, for any multi-index α, one can define

$$(\partial^\alpha u)(\varphi) = (-1)^{|\alpha|} u(\partial^\alpha \varphi).$$

Then $\partial^\alpha : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is continuous. Indeed, if $\varphi_k \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$, then clearly also $\partial^\alpha \varphi_k \to \partial^\alpha \varphi$ in $\mathcal{S}(\mathbb{R}^n)$, and, therefore,

$$(\partial^\alpha u)(\varphi_k) = (-1)^{|\alpha|} u(\partial^\alpha \varphi_k) \to (-1)^{|\alpha|} u(\partial^\alpha \varphi) = (\partial^\alpha u)(\varphi),$$

which means that $\partial^\alpha u \in \mathcal{S}'(\mathbb{R}^n)$. Moreover, let $u_k \to u \in \mathcal{S}'(\mathbb{R}^n)$. Then

$$\partial^\alpha u_k(\varphi) = (-1)^{|\alpha|} u_k(\partial^\alpha \varphi) \to (-1)^{|\alpha|} u(\partial^\alpha \varphi) = \partial^\alpha u(\varphi),$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$, i.e. ∂^α is continuous on $\mathcal{S}'(\mathbb{R}^n)$.

If a smooth function f and all of its derivatives are bounded by some polynomial functions, we can define the multiplication of tempered distributions by f by setting

$$(fu)(\varphi) = u(f\varphi).$$

This is well-defined since $\varphi \in \mathcal{S}(\mathbb{R}^n)$ implies $f\varphi \in \mathcal{S}(\mathbb{R}^n)$ if f is a function as above.

1.3.7 Fourier inversion formula for tempered distributions

As we have seen, statements on $\mathcal{S}(\mathbb{R}^n)$ can usually be extended to corresponding statements on $\mathcal{S}'(\mathbb{R}^n)$. This applies to the Fourier inversion formula as well. Let us define $\mathcal{F}^{-1}$ on $\mathcal{S}'(\mathbb{R}^n)$ by

$$(\mathcal{F}^{-1} u)(\varphi) = u(\mathcal{F}^{-1} \varphi),$$

for $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$. As before, it can be readily checked that $\mathcal{F}^{-1} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is well-defined and continuous.

Theorem. *$\mathcal{F}$ and $\mathcal{F}^{-1}$ are inverse to each other on $\mathcal{S}'(\mathbb{R}^n)$, i.e. $\mathcal{F}\mathcal{F}^{-1} = \mathcal{F}^{-1}\mathcal{F} = id$ on $\mathcal{S}'(\mathbb{R}^n)$.*

To prove this, let $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then by 1.1.8 and definitions, we get

$$(\mathcal{F}\mathcal{F}^{-1} u)(\varphi) = (\mathcal{F}^{-1} u)(\mathcal{F}\varphi) = u(\mathcal{F}^{-1}\mathcal{F}\varphi) = u(\varphi),$$

so $\mathcal{F}\mathcal{F}^{-1} u = u$ by the uniqueness principle for distributions. A similar argument applies to show that $\mathcal{F}^{-1}\mathcal{F} = id$.

To give an example of these operations, let us define the Heaviside function H on $\mathbb{R}$ by setting

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

Clearly $H \in L^\infty(\mathbb{R})$, so in particular, it is a tempered distribution. Let us also define the δ–distribution by setting

$$\delta(\varphi) = \varphi(0)$$

for all $\varphi \in \mathcal{S}(\mathbb{R})$.

We claim first that $H' = \delta$. Indeed, we have

$$H'(\varphi) = -H(\varphi') = -\int_0^\infty \varphi'(x)dx = \varphi(0) = \delta(\varphi),$$

hence $H' = \delta$ by the uniqueness principle for distributions.

Let us now calculate the Fourier transform of δ. According to the definitions, we have

$$\widehat{\delta}(\varphi) = \delta(\widehat{\varphi}) = \widehat{\varphi}(0) = \int_{\mathbb{R}} \varphi(x)dx = 1(\varphi),$$

hence $\widehat{\delta} = 1$. Here we used the fact that the constant one is in $L^\infty(\mathbb{R}^n)$, hence also a tempered distribution. It can be also checked that $\widehat{1} = \delta$.

1.3.8 $C_0^\infty(\Omega)$ and sequential density

For an open set $\Omega \subset \mathbb{R}^n$, the space $C_0^\infty(\Omega)$ is defined as the space of smooth functions $\varphi : \Omega \to \mathbb{C}$ with compact support. Here the support of φ is defined as the closure of the set where φ is non-zero, i.e. by

$$\operatorname{supp} \varphi = \overline{\{x \in \Omega : \varphi(x) \neq 0\}}.$$

It can be seen that this space is non-empty. For example, if we define function $\chi(t)$ by $\chi(t) = e^{-1/t^2}$ for $t > 0$ and by $\chi(t) = 0$ for $t \leq 0$, then $f(t) = \chi(t)\chi(1-t)$ is a smooth compactly supported function on $\mathbb{R}$. Consequently, $\varphi(x) = f(x_1) \cdots f(x_n)$ is a function in $C_0^\infty(\mathbb{R}^n)$, with $\operatorname{supp} \varphi = [0,1]^n$.

Another example is function ψ defined by $\psi(x) = e^{1/(|x|^2-1)}$ for $|x| < 1$ and by $\psi(x) = 0$ for $|x| \geq 1$. We have $\psi \in C_0^\infty(\mathbb{R}^n)$ with $\operatorname{supp} \psi = \{|x| \leq 1\}$.

Although these examples are quite special, products of these functions with any other smooth function as well as their derivatives are all in $C_0^\infty(\mathbb{R}^n)$. On the other hand, $C_0^\infty(\mathbb{R}^n)$ can not contain analytic functions, making it relatively small. Still, it is dense in very large spaces of functions/distributions in their respective topologies.

Theorem. *The space $C_0^\infty(\mathbb{R}^n)$ is sequentially dense in $\mathcal{S}'(\mathbb{R}^n)$, i.e. for every $u \in \mathcal{S}'(\mathbb{R}^n)$ there exists a sequence $u_k \in C_0^\infty(\mathbb{R}^n)$ such that $u_k \to u$ in $\mathcal{S}'(\mathbb{R}^n)$ as $k \to \infty$.*

Lemma. *The space $C_0^\infty(\mathbb{R}^n)$ is sequentially dense in $\mathcal{S}(\mathbb{R}^n)$, i.e. for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$ there exists a sequence $\varphi_k \in C_0^\infty(\mathbb{R}^n)$ such that $\varphi_k \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ as $k \to \infty$.*

Proof of Lemma. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Let us fix some $\psi \in C_0^\infty(\mathbb{R}^n)$ such that $\psi = 1$ in a neighborhood of the origin and let us define $\psi_k(x) = \psi(x/k)$. Then it can be easily checked that $\varphi_k = \psi_k\varphi \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$.

Proof of Theorem. Let $u \in \mathcal{S}'(\mathbb{R}^n)$ and let ψ and ψ_k be as in the proof of the lemma. Then $\psi u \in \mathcal{S}'(\mathbb{R}^n)$ is well-defined by $(\psi u)(\varphi) = u(\psi\varphi)$, for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. We have that $\psi_k u \to u$ in $\mathcal{S}'(\mathbb{R}^n)$. Indeed, we have that

$$(\psi_k u)(\varphi) = u(\psi_k\varphi) \to u(\varphi)$$

by the lemma. Similarly, we have that $\psi_k\widehat{u} \to \widehat{u}$ in $\mathcal{S}'(\mathbb{R}^n)$, and hence also $\mathcal{F}^{-1}(\psi_k\widehat{u}) \to u$ in $\mathcal{S}'(\mathbb{R}^n)$ because of the continuity of the Fourier transform in $\mathcal{S}'(\mathbb{R}^n)$. Consequently, we have

$$u_{kj} = \psi_j(\mathcal{F}^{-1}(\psi_k\widehat{u})) \to u$$

in $\mathcal{S}'(\mathbb{R}^n)$ as $k, j \to \infty$. It remains to show that $u_{kj} \in C_0^\infty(\mathbb{R}^n)$.

In general, let $\chi \in C_0^\infty(\mathbb{R}^n)$ and let $w = \chi\widehat{u}$. We claim that $\mathcal{F}^{-1}w \in C^\infty(\mathbb{R}^n)$. Indeed, we have

$$(\mathcal{F}^{-1}w)(\varphi) = w(\mathcal{F}^{-1}\varphi) = w_\xi(\int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\varphi(x)dx) = \int_{\mathbb{R}^n} w_\xi(e^{2\pi i x\cdot\xi})\varphi(x)dx,$$

where we write w_ξ to emphasize that w acts on the test function as the function of ξ-variable, and where we used the continuity of w and the fact that $w_\xi(e^{2\pi i x\cdot\xi}) = \widehat{u}(\chi e^{2\pi i x\cdot\xi})$ is well-defined. Now, it follows that $\mathcal{F}^{-1}w$ can be identified with the function $(\mathcal{F}^{-1}w)(x) = \widehat{u}_\xi(\chi(\xi)e^{2\pi i x\cdot\xi})$, which is smooth with respect to x. Indeed, we can note first that the right hand side depends continuously on x because of the continuity of $\widehat{u}$ on $\mathcal{S}(\mathbb{R}^n)$. Here we also use that everything is well defined since $\chi \in C_0^\infty(\mathbb{R}^n)$. Moreover, since function $\chi(\xi)e^{2\pi i x\cdot\xi}$ is compactly supported in ξ, so are its derivatives with respect to x, and hence all the derivatives of $(\mathcal{F}^{-1}w)(x)$ are also continuous in x, proving the claim and the theorem.

1.3.9 Distributions (generalised functions)

Since our main interest is in the Fourier analysis, we started with the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions which allows the definition and use of the Fourier transform. However, there is a bigger space of distributions which we will sketch here. It will contain some important classes of functions that $\mathcal{S}'(\mathbb{R}^n)$ does not contain.

Localisations of L^p–spaces. We define local versions of spaces $L^p(\Omega)$ as follows. We will say that $f \in L^p_{loc}(\Omega)$ if $\varphi f \in L^p(\Omega)$ for all $\varphi \in C_0^\infty(\Omega)$. We note that $L^p_{loc}(\mathbb{R}^n)$ are not subspaces of $\mathcal{S}'(\mathbb{R}^n)$ since they do not encode any information on the global behaviour of functions. For example, $e^{|x|^2}$ is smooth, and hence belongs to all $L^p_{loc}(\mathbb{R}^n)$, $1 \le p \le \infty$, but it is not in $\mathcal{S}'(\mathbb{R}^n)$. There is a natural notion of convergence in the localised spaces $f \in L^p_{loc}(\Omega)$. Thus, we will write

$$f_m \to f \text{ in } L^p_{loc}(\Omega) \text{ as } m \to \infty,$$

if f and f_m belong to $L^p(\Omega)_{loc}$ for all m, and if $\varphi f_m \to \varphi f$ in $L^p_{loc}(\Omega)$ as $m \to \infty$, for all $\varphi \in C_0^\infty(\Omega)$.

The difference between the space of distributions $\mathcal{D}'(\mathbb{R}^n)$ and the space of tempered distributions $\mathcal{S}'(\mathbb{R}^n)$ is the choice of the set $C_0^\infty(\mathbb{R}^n)$ rather than $\mathcal{S}(\mathbb{R}^n)$ as the space of test functions. At the same time, choosing $C_0^\infty(\Omega)$ as test functions allows one to obtain the space $\mathcal{D}'(\Omega)$ of distributions in Ω. The definition and facts below are sketched only as they are similar to 1.3.1.

We say that $\varphi_k \to \varphi$ in $C_0^\infty(\Omega)$ if $\varphi_k, \varphi \in C_0^\infty(\Omega)$, if there is a compact set $K \subset \Omega$ such that $\operatorname{supp}\varphi_k \subset K$ for all k, and if

$$\sup_{x\in\Omega} |\partial^\alpha(\varphi_k - \varphi)(x)| \to 0$$

for all multi-indices α. Then $\mathcal{D}'(\Omega)$ is defined as the set of all linear continuous functionals $u : C_0^\infty(\Omega) \to \mathbb{C}$.

It is easy to see that $C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$ and that if $\varphi_k \to \varphi$ in $C_0^\infty(\mathbb{R}^n)$, then $\varphi_k \to \varphi$ in $\mathcal{S}(\mathbb{R}^n)$. Thus, if $u \in \mathcal{S}'(\mathbb{R}^n)$ and if $\varphi_k \to \varphi$ in $C_0^\infty(\mathbb{R}^n)$, we have $u(\varphi_k) \to u(\varphi)$, which means that $u \in \mathcal{D}'(\mathbb{R}^n)$. Thus, we showed that $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$. We say that $u_k \to u \in \mathcal{D}'(\Omega)$ if $u_k, u \in \mathcal{D}'(\Omega)$ and if $u_k(\varphi) \to u(\varphi)$ for all $\varphi \in C_0^\infty(\Omega)$. One readily checks that $u_k \to u$ in $\mathcal{S}'(\mathbb{R}^n)$ implies $u_k \to u$ in $\mathcal{D}'(\mathbb{R}^n)$.

Finally, one can readily see that the canonical identification 1.3.4 yields the inclusions $L^p_{loc}(\Omega) \subset \mathcal{D}'(\Omega)$ for all $1 \leq p \leq \infty$.

1.3.10 Weak derivatives

There is a notion of a weak derivative which is a special case of the distributional derivative from 1.3.6. However, it allows a realisation in an integral form and we mention it here briefly.

Let Ω be an open subset of $\mathbb{R}^n$ and let $u, v \in L^1_{loc}(\Omega)$. We say that v is the α^{th}*–weak partial derivative of* u if

$$\int_\Omega u \partial^\alpha \varphi dx = (-1)^{|\alpha|} \int_\Omega v \varphi dx, \quad \text{for all} \quad \varphi \in C_0^\infty(\Omega).$$

In this case we also write $v = \partial^\alpha u$. The constant $(-1)^{|\alpha|}$ stands for the consistency with the corresponding definition for smooth functions when using the integration by parts in Ω. This is also the reason to include the same constant $(-1)^{|\alpha|}$ in the definition in 1.3.6. The weak derivative defined in this way is uniquely determined:

Lemma. *Let $u \in L^1_{loc}(\Omega)$. If a weak α^{th} derivative of u exists, it is uniquely defined up to a set of measure zero.*

Indeed, assume that there are two functions $v, w \in L^1_{loc}(\Omega)$ such that

$$\int_\Omega u \partial^\alpha \varphi dx = (-1)^{|\alpha|} \int_\Omega v \varphi dx = (-1)^{|\alpha|} \int_\Omega w \varphi dx,$$

for all $\varphi \in C_0^\infty(\Omega)$. Then $\int_\Omega (v - w)\varphi dx = 0$ for all $\varphi \in C_0^\infty(\Omega)$. A standard result from the measure theory now implies that $v = w$ almost everywhere in Ω.

Examples and Exercises.

(1) Let us define $u, v : \mathbb{R} \to \mathbb{R}$ by

$$u(x) = \begin{cases} x, & \text{if} \quad x \leq 1, \\ 1, & \text{if} \quad x > 1, \end{cases} \qquad v(x) = \begin{cases} 1, & \text{if} \quad x \leq 1, \\ 0, & \text{if} \quad x > 1, \end{cases}$$

Prove that $u' = v$ weakly.

(2) Define $u : \mathbb{R} \to \mathbb{R}$ by

$$u(x) = \begin{cases} x, & \text{if} \quad x \leq 1, \\ 2, & \text{if} \quad x > 1. \end{cases}$$

Prove that u has no weak derivative.

(3) Calculate the distributional derivative of u from (2).

(4) Prove that the δ-distribution is not an element of $L^1_{loc}(\mathbb{R}^n)$.

1.3.11 Sobolev spaces

Let $1 \leq p \leq \infty$ and let $k \in \mathbb{N} \cup \{0\}$. The *Sobolev space* $L^p_k(\Omega)$ (or $W^{p,k}(\Omega)$) consists of all $u \in L^1_{loc}(\Omega)$ such that for all multi-indices α with $|\alpha| \leq k$, $\partial^\alpha u$ exists weakly (or distributionally) and $\partial^\alpha u \in L^p(\Omega)$.

Since $p \geq 1$, we know that $L^p_{loc}(\Omega) \subset L^1_{loc}(\Omega)$, so we note that it does not matter whether we take a weak or a distributional derivative.

In the case $p = 2$, one often uses the notation $H^k(\Omega)$ for $L^2_k(\Omega)$, and in the case $p = 2$ and $k = 0$, we get $H^0(\Omega) = L^2(\Omega)$. As usual, we identify functions in $L^p_k(\Omega)$ which are equal almost everywhere.

For $u \in L^p_k(\Omega)$, we define

$$||u||_{L^p_k(\Omega)} = \left(\sum_{|\alpha| \leq k} ||\partial^\alpha u||^p_{L^p} \right)^{1/p} = \left(\sum_{|\alpha| \leq k} \int_\Omega |\partial^\alpha u|^p dx \right)^{1/p},$$

for $1 \leq p < \infty$, and for $p = \infty$ we define

$$||u||_{L^\infty_k(\Omega)} = \sum_{|\alpha| \leq k} \text{esssup}_\Omega |\partial^\alpha u|.$$

It can be readily checked that these expressions are norms. Indeed, we clearly have

$$||\lambda u||_{L^p_k} = |\lambda| ||u||_{L^p_k}$$

for all $\lambda \in \mathbb{C}$, and $||u||_{L^p_k} = 0$ if and only if $u = 0$ almost everywhere. For the triangle inequality, the case $p = \infty$ is straightforward. For $1 \leq p < \infty$ and for $u, v \in L^p_k(\Omega)$, Minkowski's inequality implies

$$\begin{aligned} ||u+v||_{L^p_k} &= \left(\sum_{|\alpha| \leq k} ||\partial^\alpha u + \partial^\alpha v||^p_{L^p} \right)^{1/p} \leq \left(\sum_{|\alpha| \leq k} (||\partial^\alpha u||_{L^p} + ||\partial^\alpha v||_{L^p})^p \right)^{1/p} \\ &\leq \left(\sum_{|\alpha| \leq k} ||\partial^\alpha u||^p_{L^p} \right)^{1/p} + \left(\sum_{|\alpha| \leq k} ||\partial^\alpha v||^p_{L^p} \right)^{1/p} \\ &= ||u||_{L^p_k} + ||v||_{L^p_k}. \end{aligned}$$

Localisations of Sobolev spaces. We define local versions of spaces $L^p_k(\Omega)$ similarly to local versions of L^p–spaces. We will say that $f \in L^p_k(\Omega)_{loc}$ if $\varphi f \in L^p_k(\Omega)$ for all $\varphi \in C^\infty_0(\Omega)$. We will write $f_m \to f$ in $L^p_k(\Omega)_{loc}$ as $m \to \infty$, if f and f_m belong to $L^p_k(\Omega)_{loc}$ for all m, and if $\varphi f_m \to \varphi f$ in $L^p_k(\Omega)$ as $m \to \infty$, for all $\varphi \in C^\infty_0(\Omega)$.

1.3.12 Example of a point singularity

An often encountered example of a function with a point singularity is the following $u(x) = |x|^{-a}$ defined for $x \in \Omega = B(0,1) \subset \mathbb{R}^n$, $x \neq 0$. We may ask a question: for which $a > 0$ do we have $u \in L^p_1(\Omega)$?

First we observe that away from the origin, u is a smooth function and can be differentiated pointwise with $\partial_{x_j} u = -ax_j|x|^{-a-2}$ and hence also $|\nabla u(x)| = |a||x|^{-a-1}$, $x \neq 0$. In particular, $|\nabla u| \in L^1(\Omega)$ for $a+1 < n$. We also have $|\nabla u| \in L^p(\Omega)$ for $(a+1)p < n$ So we must assume $a+1 < n$ and $(a+1)p < n$. Let us now calculate the weak (distributional) derivative of u in Ω. Let $\varphi \in C_0^\infty(\Omega)$. Let $\epsilon > 0$. On $\Omega \backslash B(0,\epsilon)$ we can integrate by parts to get

$$\int_{\Omega \backslash B(0,\epsilon)} u \partial_{x_j} \varphi dx = -\int_{\Omega \backslash B(0,\epsilon)} \partial_{x_j} u \varphi dx + \int_{\partial B(0,\epsilon)} u \varphi \nu^j d\sigma,$$

where $d\sigma$ is the surface measure on the sphere $\partial B(0,\epsilon)$ and $\nu = (\nu^1, \ldots, \nu^n)$ is the inward pointing normal on $\partial B(0,\epsilon)$. Now, since $u = |\epsilon|^{-a}$ on $\partial B(0,\epsilon)$, we can estimate

$$\left| \int_{\partial B(0,\epsilon)} u \varphi \nu^j d\sigma \right| \leq ||\varphi||_{L^\infty} \int_{\partial B(0,\epsilon)} \epsilon^{-a} d\sigma \leq C \epsilon^{n-1-a} \to 0$$

as $\epsilon \to 0$, since $a+1 < n$. Passing to the limit in the integration by parts formula, we get

$$\int_\Omega u \partial_{x_j} \varphi dx = -\int_\Omega \partial_{x_j} u \varphi dx,$$

which means that $\partial_{x_j} u$ is also the weak derivative of u. So, $u \in L_1^p(\Omega)$ if $u, |\nabla u| \in L^p(\Omega)$, which holds for $(a+1)p < n$, i.e. for $a < (n-p)/p$.

We leave it as an exercise to find conditions on a for which $u \in L_k^p(\Omega)$.

1.3.13 Some properties of Sobolev spaces

Since $L^p(\Omega) \subset \mathcal{D}'(\Omega)$, we can work with $u \in L^p(\Omega)$ as with function or as with distributions. In particular, we can differentiate them distributionally, etc. Moreover, as we have already seen, the equality of objects (be it functions, functionals, distributions, etc.) depends on the spaces in which the equality is considered. In Sobolev spaces we can use tools from the measure theory so we work with functions defines almost everywhere. Thus, in equality in Sobolev spaces (as in the following theorem) means pointwise equality almost everywhere.

Theorem. *Let $u, v \in L_k^p(\Omega)$, and let α be a multi-index with $|\alpha| \leq k$. Then*

(i) $\partial^\alpha u \in L^p_{k-|\alpha|}(\Omega)$, *and* $\partial^\alpha(\partial^\beta u) = \partial^\beta(\partial^\alpha u) = \partial^{\alpha+\beta} u$, *for all multi-indices* α, β *such that* $|\alpha| + |\beta| \leq k$.

(ii) *For all* $\lambda, \mu \in \mathbb{C}$ *we have* $\lambda u + \mu v \in L_k^p(\Omega)$ *and* $\partial^\alpha(\lambda u + \mu v) = \lambda \partial^\alpha u + \mu \partial^\alpha v$.

(iii) *If* $\widetilde{\Omega}$ *is an open subset of* Ω, *then* $u \in L_k^p(\widetilde{\Omega})$.

(iv) *If* $\chi \in C_0^\infty(\Omega)$, *then* $\chi u \in L_k^p(\Omega)$ *and we have the Leibnitz formula*

$$\partial^\alpha(\chi u) = \sum_{\beta \leq \alpha} C_\alpha^\beta (\partial^\beta \chi)(\partial^{\alpha-\beta} u),$$

where the binomial coefficient is $C_\alpha^\beta = \frac{\alpha!}{\beta!(\alpha-\beta)!}$.

(v) $L^p_k(\Omega)$ is a Banach space.

Proof. Statements (i), (ii), and (iii) are easy. For example, if $\varphi \in C_0^\infty(\Omega)$ then also $\partial^\beta \varphi \in C_0^\infty(\Omega)$, and (i) follows from

$$\int_\Omega \partial^\alpha u \, \partial^\beta \varphi dx = (-1)^{|\alpha|} \int_\Omega u \, \partial^{\alpha+\beta} \varphi dx = (-1)^{|\alpha|+|\alpha+\beta|} \int_\Omega \partial^{\alpha+\beta} u \, \varphi dx,$$

since $(-1)^{|\alpha|+|\alpha+\beta|} = (-1)^{|\beta|}$.

Let us now show (iv). The proof will be carried out by induction on $|\alpha|$. For $|\alpha| = 1$, writing $\langle u, \varphi \rangle$ for $u(\varphi) = \int_\Omega u\varphi dx$, we get

$$\langle \partial^\alpha(\chi u), \varphi \rangle = (-1)^{|\alpha|} \langle u, \chi \partial^\alpha \varphi \rangle = -\langle u, \partial^\alpha(\chi\varphi) - (\partial^\alpha \chi)\varphi \rangle = \langle \chi \partial^\alpha u, \varphi \rangle + \langle (\partial^\alpha \chi) u, \varphi \rangle,$$

which is what was required. Now, suppose that the Leibnitz formula is valid for all $|\beta| \leq l$, and let us take α with $|\alpha| = l + 1$. Then we can write $\alpha = \beta + \gamma$ with some $|\beta| = l$ and $|\gamma| = 1$. We get

$$\begin{aligned}
\langle \chi u, \partial^\alpha \varphi \rangle &= \big\langle \chi u, \partial^\beta (\partial^\gamma \varphi) \big\rangle \\
&= (-1)^{|\beta|} \big\langle \partial^\beta(\chi u), \partial^\gamma \varphi \big\rangle \quad \text{(by induction hypothesis)} \\
&= (-1)^{|\beta|} \left\langle \sum_{\sigma \leq \beta} C_\beta^\sigma \partial^\sigma \chi \partial^{\beta-\sigma} u, \partial^\gamma \varphi \right\rangle \quad \text{(by definition)} \\
&= (-1)^{|\beta|+|\gamma|} \left\langle \sum_{\sigma \leq \beta} C_\beta^\sigma \partial^\gamma (\partial^\sigma \chi \partial^{\beta-\sigma} u), \varphi \right\rangle \quad (\text{set } \ \rho = \sigma + \gamma) \\
&= (-1)^{|\alpha|} \left\langle \sum_{\sigma \leq \beta} C_\beta^\sigma \big(\partial^\rho \chi \partial^{\alpha-\rho} u + \partial^\sigma \chi \partial^{\alpha-\sigma} u \big), \varphi \right\rangle \\
&= (-1)^{|\alpha|} \left\langle \sum_{\rho \leq \alpha} C_\alpha^\rho \partial^\rho \chi \partial^{\alpha-\rho} u, \varphi \right\rangle,
\end{aligned}$$

where we used that $C_\beta^\sigma + C_\beta^\rho = C_{\alpha-\gamma}^{\rho-\gamma} + C_{\alpha-\gamma}^\rho = C_\alpha^\rho$.

Now let us prove (v). We have already shown in 1.3.11 that $L^p_k(\Omega)$ is a normed space. Let us show now that the completeness of $L^p(\Omega)$ implies the completeness of $L^p_k(\Omega)$. Let u_m be a Cauchy sequence in $L^p_k(\Omega)$. Then $\partial^\alpha u_m$ is a Cauchy sequence in $L^p(\Omega)$ for any $|\alpha| \leq k$. Since $L^p(\Omega)$ is complete, there exists some $u_\alpha \in L^p(\Omega)$ such that $\partial^\alpha u_m \to u_\alpha$ in $L^p(\Omega)$. Let $u = u_{(0,\cdots,0)}$, so in particular, we have $u_m \to u$ in $L^p(\Omega)$. Let us now show that in fact $u \in L^p_k(\Omega)$ and $\partial^\alpha u = u_\alpha$ for all $|\alpha| \leq k$. Let $\varphi \in C_0^\infty(\Omega)$. Then

$$\begin{aligned}
\langle \partial^\alpha u, \varphi \rangle = (-1)^{|\alpha|} \langle u, \partial^\alpha \varphi \rangle &= (-1)^{|\alpha|} \lim_{m \to \infty} \langle u_m, \partial^\alpha \varphi \rangle \\
&= \lim_{m \to \infty} \langle \partial^\alpha u_m, \varphi \rangle = \langle u_\alpha, \varphi \rangle,
\end{aligned}$$

which implies $u \in L^p_k(\Omega)$ and $\partial^\alpha u = u_\alpha$. Moreover, we have $\partial^\alpha u_m \to \partial^\alpha u$ in $L^p(\Omega)$ for all $|\alpha| \leq k$, which means that $u_m \to u$ in $L^p_k(\Omega)$ and hence $L^p_k(\Omega)$ is complete.

1.3.14 Mollifiers

In 1.3.8 we saw that we can approximate irregular function or (tempered) distributions by much more regular functions. The argument relied on the use of the Fourier analysis and worked well on $\mathbb{R}^n$. Such technique is very powerful, as could have been seen from the proof of Plancherel's formula. On the other hand, when working in subsets of $\mathbb{R}^n$ we may be unable to use the Fourier transform (since for its definition we used the whole space $\mathbb{R}^n$). We want to be able to approximate functions (or distributions) by smooth functions without using the Fourier techniques. This turns out to be possible using the so-called *mollification* of functions.

Assume for a moment that we are in $\mathbb{R}^n$ again and let us first argue very informally. Let us first look at the Fourier transform of the convolution with a δ-distribution. Thus, for a function f we must have $\widehat{\delta * f} = \widehat{\delta}\widehat{f} = \widehat{f}$, if we use that $\widehat{\delta} = 1$. Taking the inverse Fourier transform we obtain the important identity

$$\delta * f = f,$$

which will be justified formally later. Now, if we take a sequence of smooth functions η_ϵ approximating δ-distribution, i.e. if $\eta_\epsilon \to \delta$ in some sense, and if this convergence is preserved by the convolution, we should get

$$\eta_\epsilon * f \to \delta * f = f \quad \text{as } \epsilon \to 0.$$

Now, the convolution $\eta_\epsilon * f$ may be defined locally in $\mathbb{R}^n$, and functions $\eta_\epsilon * f$ will be smooth if η_ϵ are, thus giving us a way to approximate f. We will now make this argument precise. For this, we will deal in straightforward manner by looking at the limit of $\eta_\epsilon * f$ for a suitably chosen sequence of functions η_ϵ, without referring neither to δ-distribution nor to the Fourier transform.

For an open set $\Omega \subset \mathbb{R}^n$ and $\epsilon > 0$ we define $\Omega_\epsilon = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \epsilon\}$. Let us define $\eta \in C_0^\infty(\mathbb{R}^n)$ by

$$\eta(x) = \begin{cases} Ce^{\frac{1}{|x|^2-1}}, & \text{if} \quad |x| < 1, \\ 0, & \text{if} \quad |x| \geq 1, \end{cases}$$

where constant C is chosen so that $\int_{\mathbb{R}^n} \eta dx = 1$. Function η is called a *mollifier*. For $\epsilon > 0$, we define

$$\eta_\epsilon(x) = \frac{1}{\epsilon^n}\eta\left(\frac{x}{\epsilon}\right),$$

so that $\operatorname{supp}\eta_\epsilon \subset B(0,\epsilon)$ and $\int_{\mathbb{R}^n} \eta_\epsilon dx = 1$.

Let $f \in L^1_{loc}(\Omega)$. A mollification of f corresponding to η is a family $f^\epsilon = \eta_\epsilon * f$ in Ω_ϵ, i.e.

$$f^\epsilon(x) = \int_\Omega \eta_\epsilon(x-y)f(y)dy = \int_{B(0,\epsilon)} \eta_\epsilon(y)f(x-y)dy, \quad \text{for } x \in \Omega_\epsilon.$$

Theorem. *Let $f \in L^1_{loc}(\Omega)$. Then we have the following properties.*

(i) $f^\epsilon \in C^\infty(\Omega_\epsilon)$.

(ii) $f^\epsilon \to f$ *almost everywhere as* $\epsilon \to 0$.

(iii) *If* $f \in C(\Omega)$, *then* $f^\epsilon \to f$ *uniformly on compact subsets of* Ω.

(iv) $f^\epsilon \to f$ *in* $L^p_{loc}(\Omega)$ *for all* $1 \le p < \infty$.

Proof. To show (i), we can differentiate $f^\epsilon(x) = \int_\Omega \eta_\epsilon(x-y)f(y)dy$ and use the fact that $f \in L^1_{loc}(\Omega)$. The proof of (ii) will rely on the following

Lebesgue's differentiation theorem. *Let* $f \in L^1_{loc}(\Omega)$. *Then*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y)-f(x)|dy = 0 \quad \text{for a.e. } x \in \Omega.$$

Now, for all x for which the statement of the Lebesgue's differentiation theorem is true, we can estimate

$$\begin{aligned} |f^\epsilon(x)-f(x)| &= \left|\int_{B(x,\epsilon)} \eta_\epsilon(x-y)(f(y)-f(x))dy\right| \\ &\le \epsilon^{-n}\int_{B(x,\epsilon)} \eta\left(\frac{x-y}{\epsilon}\right)|f(y)-f(x)|dy \\ &\le C\frac{1}{|B(x,\epsilon)|}\int_{B(x,\epsilon)} |f(y)-f(x)|dy, \end{aligned}$$

where the last expression goes to zero as $\epsilon \to 0$. For (iii), let K be a compact subset of Ω. Let $K_0 \subset \Omega$ be another compact set such that K is contained in the interior of K_0. Then f is uniformly continuous on K_0 and the limit in the Lebesgue's differentiation theorem holds uniformly for $x \in K$. The same argument as in (ii) then shows that $f^\epsilon \to f$ uniformly on K.

Finally, to show (iv), let us choose open sets $U \subset V \subset \Omega$ such that $U \subset V_\delta$ and $V \subset \Omega_\delta$ for some small $\delta > 0$. Let us show first that $||f^\epsilon||_{L^p(U)} \le ||f||_{L^p(V)}$ for all sufficiently small $\epsilon > 0$. Indeed, for all $x \in U$, we can estimate

$$\begin{aligned} |f^\epsilon(x)| &= \left|\int_{B(x,\epsilon)} \eta_\epsilon(x-y)f(y)dy\right| \\ &\le \int_{B(x,\epsilon)} \eta_\epsilon^{1-1/p}(x-y)\eta_\epsilon^{1/p}(x-y)|f(y)|dy \quad \text{(Hölder's inequality)} \\ &\le \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)dy\right)^{1-1/p}\left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|^p dy\right)^{1/p}. \end{aligned}$$

Since $\int_{B(x,\epsilon)} \eta_\epsilon(x-y)dy = 1$, we get

$$\begin{aligned} \int_U |f^\epsilon(x)|^p dx &\le \int_U \left(\int_{B(x,\epsilon)} \eta_\epsilon(x-y)|f(y)|^p dy\right)dx \\ &\le \int_V \left(\int_{B(y,\epsilon)} \eta_\epsilon(x-y)dx\right)|f(y)|^p dy \\ &= \int_V |f(y)|^p dy. \end{aligned}$$

Now, let $\delta > 0$ and let us choose $g \in C(V)$ such that $||f - g||_{L^p(V)} < \delta$. (Here we used the fact that $C(V)$ is (sequentially) dense in $L^p(V)$). Then

$$\begin{aligned} ||f^\epsilon - f||_{L^p(U)} &\leq ||f^\epsilon - g^\epsilon||_{L^p(U)} + ||g^\epsilon - g||_{L^p(U)} + ||g - f||_{L^p(U)} \\ &\leq 2||f - g||_{L^p(V)} + ||g^\epsilon - g||_{L^p(U)} \\ &< 2\delta + ||g^\epsilon - g||_{L^p(U)}. \end{aligned}$$

Since $g^\epsilon \to g$ uniformly on the closure of V by (iii), it follows that $||f^\epsilon - f||_{L^p(U)} \leq 3\delta$ for small enough $\epsilon > 0$, completing the proof of (iv).

The is a simple but useful corollary of the Lebesgue's differentiation theorem, partly explaining its name.

Corollary of the Lebesgue's differentiation theorem. *Let $f \in L^1_{loc}(\Omega)$. Then*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)dy = f(x) \quad \text{for a.e. } x \in \Omega.$$

1.3.15 Approximation of Sobolev space functions

With the use of mollifications we can approximate functions in Sobolev spaces by smooth functions. We have a local approximation in localised spaces $L^p_k(\Omega)_{loc}$, a global approximation in $L^p_k(\Omega)$, and further approximations dependent on the regularity of the boundary of Ω. Although the set Ω is bounded, we still say that an approximation in $L^p_k(\Omega)$ is global if it works *up to the boundary.*

Local approximation by smooth functions. *Assume that $\Omega \subset \mathbb{R}^n$ is open. Let $f \in L^p_k(\Omega)$ for $1 \leq p < \infty$ and $k \in \mathbb{N} \cup \{0\}$. Let $f^\epsilon = \eta_\epsilon * f$ in Ω_ϵ be the mollification of f, $\epsilon > 0$. Then $f^\epsilon \in C^\infty(\Omega_\epsilon)$ and $f^\epsilon \to f$ in $L^p_k(\Omega)_{loc}$ as $\epsilon \to 0$, i.e. $f^\epsilon \to f$ in $L^p_k(K)$ as $\epsilon \to 0$ for all compact $K \subset \Omega$.*

Proof. It was already proved in Theorem 1.3.14, (i), that $f^\epsilon \in C^\infty(\Omega_\epsilon)$. Since f is locally integrable, we can differentiate the convolution under the integral sign to get $\partial^\alpha f^\epsilon = \eta_\epsilon * \partial^\alpha u$ in Ω_ϵ. Now, let U be an open and bounded subset of Ω containing K. Then by Theorem 1.3.14, (iv), we get $\partial^\alpha f^\epsilon \to \partial^\alpha f$ in $L^p(U)$ as $\epsilon \to 0$, for all $|\alpha| \leq k$. Hence

$$||f^\epsilon - f||^p_{L^p_k(U)} = \sum_{|\alpha|\leq k} ||\partial^\alpha f^\epsilon - \partial^\alpha f||^p_{L^p(U)} \to 0$$

as $\epsilon \to 0$, proving the statement.

Global approximation by smooth functions. *Assume that $\Omega \subset \mathbb{R}^n$ is open and bounded. Let $f \in L^p_k(\Omega)$ for $1 \leq p < \infty$ and $k \in \mathbb{N} \cup \{0\}$. Then there is a sequence $f_m \in C^\infty(\Omega) \bigcap L^p_k(\Omega)$ such that $f_m \to f$ in $L^p_k(\Omega)$.*

Proof. Let us write $\Omega = \bigcup_{j=1}^\infty \Omega_j$, where

$$\Omega_j = \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > 1/j\}.$$

Let $V_j = \Omega_{j+3} \backslash \overline{\Omega}_{j+1}$ (this definition will be very important). Take also any open V_0 with $\overline{V_0} \subset \Omega$ so that $\Omega = \bigcup_{j=0}^\infty V_j$. Let χ_j be a partition of unity subordinate to V_j, i.e. a family $\chi_j \in C_0^\infty(V_j)$ such that $0 \leq \chi_j \leq 1$ and $\sum_{j=0}^\infty \chi_j = 1$ in Ω. Then $\chi_j f \in L_k^p(\Omega)$ and $\text{supp}(\chi_j f) \subset V_j$. Let us fix some $\delta > 0$ and choose $\epsilon_j > 0$ so small that function $f^j = \eta_{\epsilon_j} * (\chi_j f)$ is supported in $W_j = \Omega_{j+4} \backslash \overline{\Omega}_j$ and satisfies

$$||f^j - \chi_j f||_{L_k^p(\Omega)} \leq \delta 2^{-j-1}$$

for all j. Let now $g = \sum_{j=0}^\infty f^j$. Then $g \in C^\infty(\Omega)$ since in any open set U in Ω there are only finitely many non-zero terms in the sum. Moreover, since $f = \sum_{j=0}^\infty \chi_j f$, for each such U we have

$$||g - f||_{L_k^p(U)} \leq \sum_{j=0}^\infty ||f^j - \chi_j f||_{L_k^p(\Omega)} \leq \delta \sum_{j=0}^\infty \frac{1}{2^{j+1}} = \delta.$$

Taking the supremum over all open subsets U of Ω, we obtain $||g - f||_{L_k^p(\Omega)} \leq \delta$, completing the proof.

In general, there are many versions of these results depending on the set Ω, in particular on the regularity of its boundary. For example, we give here without proof the following

Further result. *Let Ω be a bounded subset of $\mathbb{R}^n$ with C^1 boundary. Let $f \in L_k^p(\Omega)$ for $1 \leq p < \infty$ and $k \in \mathbb{N} \cup \{0\}$. Then there is a sequence $f_m \in C^\infty(\overline{\Omega})$ such that $f_m \to f$ in $L_k^p(\Omega)$.*

2 Pseudo-differential operators on $\mathbb{R}^n$

2.1 Analysis of operators

2.1.1 Motivation and definition

We will start with an informal observation that if T is a translation invariant linear operator on some space of functions on $\mathbb{R}^n$, then we can write

$$T(e^{2\pi i x \cdot \xi}) = a(\xi) e^{2\pi i x \cdot \xi} \quad \text{for all} \quad \xi \in \mathbb{R}^n.$$

Indeed, if T acts on functions of variable y, we can denote $f(x, \xi) = T(e^{2\pi i y \cdot \xi})(x)$. To avoid unclarity here and in the sequel, denoting $e_\xi(x) = e^{2\pi i x \cdot \xi}$, we have $f(x, \xi) = (Te_\xi)(x)$.

Let $(\tau_h f)(x) = f(x - h)$ be the translation operator by $h \in \mathbb{R}^n$. We say that T is *translation invariant* if $T\tau_h = \tau_h T$ for all h. By our assumptions on T we get

$$f(x + h, \xi) = T(e^{2\pi i (y+h) \cdot \xi}) = e^{2\pi i h \cdot \xi} T(e^{2\pi i y \cdot \xi}) = e^{2\pi i h \cdot \xi} f(x, \xi).$$

Now, setting $x = 0$, we get $f(h, \xi) = e^{2\pi i h \cdot \xi} f(0, \xi)$, so we obtain the desired formula with $a(\xi) = f(0, \xi)$. If we now formally apply T to the Fourier inversion formula

$$f(x) = \int_{\mathbb{R}^n} e^{2\pi i x \cdot \xi} \widehat{f}(\xi) d\xi$$

and use the linearity of T, we obtain

$$Tf(x) = \int_{\mathbb{R}^n} T(e^{2\pi i x \cdot \xi}) \widehat{f}(\xi) d\xi = \int_{\mathbb{R}^n} e^{2\pi i x \cdot \xi} a(\xi) \widehat{f}(\xi) d\xi.$$

This formula allows one to reduce certain properties of operator T to properties of the multiplication by the corresponding function $a(\xi)$, called the symbol of T. For example, continuity of T on L^2 would reduce to the boundedness of $a(\xi)$, composition of two operators $T_1 \circ T_2$ would reduce to the multiplication of their symbols $a_1(\xi)a_2(\xi)$, etc. Pseudo-differential operators will extend this construction to functions which are not necessarily translation invariant. In fact, as we saw above we can always denote

$$a(x, \xi) := e^{-2\pi i x\cdot\xi} T(e^{2\pi i x\cdot\xi}),$$

so that we would have

$$T(e^{2\pi i x\cdot\xi}) = e^{2\pi i x\cdot\xi} a(x, \xi).$$

Consequently, reasoning as above, we could arrive at the formula

$$Tf(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} a(x, \xi) \widehat{f}(\xi) d\xi. \tag{2.1}$$

Now, in order to avoid several rather informal conclusions in the arguments above, one usually takes the opposite route and adopts the last formula as the definition of the pseudo-differential operator with symbol $a(x, \xi)$. Such operators are then often denoted by T_a or by $a(x, D)$.

The simplest and perhaps most useful class of symbols allowing this approach to work well is the following class denoted by S^m, or by $S^m_{1,0}$. We will say that $a \in S^m$ if $a = a(x, \xi)$ is smooth on $\mathbb{R}^n \times \mathbb{R}^n$ and if

$$|\partial_x^\beta \partial_\xi^\alpha a(x, \xi)| \leq A_{\alpha\beta}(1 + |\xi|)^{m-|\alpha|} \quad \text{for all} \quad \alpha, \beta \quad \text{and all} \quad x, \xi \in \mathbb{R}^n.$$

Note that for partial differential operators symbols are just the characteristic polynomials. One can readily see that the symbol of the differential operator

$$L = \sum_{|\alpha|\leq m} a_\alpha(x) \partial_x^\alpha$$

is given by

$$a(x, \xi) = \sum_{|\alpha|\leq m} a_\alpha(x)(2\pi i \xi)^\alpha,$$

and $a \in S^m$ if coefficients a_α and all of their derivatives are smooth and bounded on $\mathbb{R}^n$.

2.1.2 Freezing principle for PDEs

Suppose we want to solve the following equation for an unknown function $u = u(x)$:

$$(Lu)(x) = \sum_{1\leq i,j\leq n} a_{ij}(x) \frac{\partial^2 u}{\partial x_i \partial x_j}(x) = f(x),$$

where the matrix $\{a_{ij}(x)\}$ is a real-valued, smooth, symmetric and positive definite. If we want to proceed in analogy to the Laplace equation in 1.1.6, we should look for the inverse of the operator L. In the case of an operator with variable coefficients this may turn out to be difficult, so we may look for an approximate inverse P such that

$$LP = I + E,$$

where the error E is small in some sense. To argue similar to 1.1.6, we "freeze" L at x_0 to get the constant coefficients operator

$$L_{x_0} = \sum_{1\le i,j\le n} a_{ij}(x_0)\frac{\partial^2}{\partial x_i \partial x_j}.$$

Now, L_{x_0} has the exact inverse which is an operator of multiplication by

$$\left(-4\pi^2 \sum_{1\le i,j\le n} a_{ij}(x_0)\xi_i\xi_j\right)^{-1}$$

on the Fourier transform side. To avoid a singularity at the origin, we introduce a cut-off function $\chi \in C^\infty$ which is 0 near the origin and 1 for large ξ. Then we define

$$(P_{x_0}f)(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\left(-4\pi^2 \sum_{1\le i,j\le n} a_{ij}(x_0)\xi_i\xi_j\right)^{-1} \chi(\xi)\widehat{f}(\xi)d\xi.$$

Then we can readily see that

$$(L_{x_0}P_{x_0}f)(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\chi(\xi)\widehat{f}(\xi)d\xi = f(x) + \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}(\chi(\xi)-1)\widehat{f}(\xi)d\xi.$$

It follows that

$$L_{x_0}P_{x_0} = I + E_{x_0},$$

where

$$(E_{x_0}f)(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}(\chi(\xi)-1)\widehat{f}(\xi)d\xi$$

is an operator of multiplication by a compactly supported function on the Fourier transform side. Writing it as a convolution with a smooth test function we can readily see that it is smoothing.

Now, we can "unfreeze" the point x_0 expecting that the inverse P will be close to P_{x_0} for x close to x_0, and define

$$(Pf)(x) = (P_xf)(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi}\left(-4\pi^2 \sum_{1\le i,j\le n} a_{ij}(x)\xi_i\xi_j\right)^{-1} \chi(\xi)\widehat{f}(\xi)d\xi.$$

It will be clear from a later discussion that we still have

$$LP = I + E_1$$

with error E_1 being "smoothing or order one". We can then set up an iterative procedure to improve the approximation of the inverse operator relying on the calculus of the appearing operators.

2.1.3 Pseudo-differential operators on the Schwartz space

It can be shown that if $a \in S^m$ and $f \in \mathcal{S}(\mathbb{R}^n)$, then $T_a f \in \mathcal{S}(\mathbb{R}^n)$. First we observe that the integral in (2.1) converges absolutely. The same is true for all of its derivatives with respect to x by the Lebesgue's dominated convergence theorem, which implies that $T_a f \in C^\infty(\mathbb{R}^n)$. Let us show now that in fact $T_a f \in \mathcal{S}(\mathbb{R}^n)$. Introducing operator

$$L_\xi = (1 + 4\pi^2 |x|^2)^{-1}(I - \Delta_\xi)$$

with the property $L_\xi e^{2\pi i x\cdot\xi} = e^{2\pi i x\cdot\xi}$, integrating (2.1) by parts N times yields

$$(T_a f)(x) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} (L_\xi)^N [a(x,\xi)\widehat{f}(\xi)] d\xi.$$

From this we get

$$|(T_a f)(x)| \leq C_N (1 + |x|)^{-2N}$$

for all N, so $T_a f$ is rapidly decreasing. The same argument applies to derivatives of $T_a f$ to show that $T_a f \in \mathcal{S}(\mathbb{R}^n)$.

The following convergence criterion will be useful in the sequel. It follows directly from the Lebesgue's dominated convergence theorem.

Convergence criterion for pseudo-differential operators. Suppose we have a sequence of symbols $a_k \in S^m$ which satisfies uniform symbolic estimates

$$|\partial_x^\beta \partial_\xi^\alpha a_k(x,\xi)| \leq A_{\alpha\beta}(1 + |\xi|)^{m-|\alpha|},$$

for all α, β, all $x, \xi \in \mathbb{R}^n$, and all k. Suppose that $a \in S^m$ is such that $a_k(x,\xi)$ and all of its derivatives converge to $a(x,\xi)$ and its derivatives, respectively, pointwise as $k \to \infty$. Then $T_{a_k} f \to T_a f$ in $\mathcal{S}(\mathbb{R}^n)$ for any $f \in \mathcal{S}(\mathbb{R}^n)$.

2.1.4 Alternative definition of pseudo-differential operators

If we writing out the Fourier transform in (2.1), we obtain

$$(T_a f)(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i (x-y)\cdot\xi} a(x,\xi) f(y) dy d\xi. \tag{2.2}$$

However, a problem with this argument and with this formula is that the ξ-integral does not converge absolutely even for $f \in \mathcal{S}(\mathbb{R}^n)$. To overcome this difficulty, one uses the idea to approximate $a(x,\xi)$ by symbols with compact support. To this end, let us fix some $\gamma \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ such that $\gamma = 1$ near the origin. Let us now define

$$a_\epsilon(x,\xi) = a(x,\xi)\gamma(\epsilon x, \epsilon\xi).$$

Then one can readily check that $a_\epsilon \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ and

- if $a \in S^m$, then $a_\epsilon \in S^m$ uniformly in $0 < \epsilon \leq 1$ (this means that constants in symbolic inequalities may be chosen independent of $0 < \epsilon \leq 1$);
- $a_\epsilon \to a$ pointwise as $\epsilon \to 0$, uniformly in $0 < \epsilon \leq 1$. The same is true for derivatives of a_ϵ and a.

It follows now from the convergence criterion of 2.1.3 that

$$T_{a_\epsilon} f \to T_a f \text{ in } \mathcal{S}(\mathbb{R}^n) \text{ as } \epsilon \to 0,$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$. Here $T_a f$ is defined as in (2.1). Now, formula (2.2) does make sense for $a_\epsilon \in C_0^\infty$, so we may define the double integral in (2.2) as the limit in $\mathcal{S}(\mathbb{R}^n)$ of $T_{a_\epsilon} f$, i.e. take

$$(T_a f)(x) = \lim_{\epsilon \to 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i (x-y)\cdot\xi} a_\epsilon(x,\xi) f(y) dy d\xi, \ f \in \mathcal{S}(\mathbb{R}^n).$$

2.1.5 Pseudo-differential operators on $\mathcal{S}'(\mathbb{R}^n)$

Recall that we can define the L^2-adjoint T_a^* of operator T_a by the formula

$$(T_a f, g)_{L^2} = (f, T_a^* g)_{L^2}, \quad f, g \in \mathcal{S}(\mathbb{R}^n),$$

where

$$(u, v)_{L^2} = \int_{\mathbb{R}^n} u(x)\overline{v(x)} dx$$

is the usual L^2–inner product. From this formula we can readily calculate that

$$(T_a^* g)(y) = \lim_{\epsilon \to 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i (y-x)\cdot\xi} \overline{a_\epsilon(x,\xi)} g(x) dx d\xi, \quad g \in \mathcal{S}(\mathbb{R}^n).$$

With the same understanding of divergent integrals as in (2.2) and replacing x by z to eliminate any confusion, we can write

$$(T_a^* g)(y) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i (y-z)\cdot\xi} \overline{a(z,\xi)} g(z) dz d\xi, \quad g \in \mathcal{S}(\mathbb{R}^n).$$

As before, by integration by parts, we can check that $T_a^* : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$. Let $u \in \mathcal{S}'(\mathbb{R}^n)$. We can now define $T_a u$ by the formula

$$(T_a u)(\varphi) = u(\overline{T_a^* \overline{\varphi}}) \quad \text{for all } \varphi \in \mathcal{S}(\mathbb{R}^n).$$

We clearly have

$$\overline{T_a^* \overline{\varphi}}(y) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i (z-y)\cdot\xi} a(z,\xi) \varphi(z) dz d\xi,$$

so if $u, \varphi \in \mathcal{S}(\mathbb{R}^n)$, we have the consistency in

$$(T_a u)(\varphi) = \int_{\mathbb{R}^n} T_a u(x)\varphi(x) dx = (T_a u, \overline{\varphi})_{L^2} = (u, T_a^* \overline{\varphi})_{L^2} = \int_{\mathbb{R}^n} u(x) \overline{T_a^* \overline{\varphi}(x)} dx = u(\overline{T_a^* \overline{\varphi}}).$$

One can readily check that $T_a u \in \mathcal{S}'(\mathbb{R}^n)$ and that $T_a : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is continuous. Indeed, let $u_k \to u$ in $\mathcal{S}'(\mathbb{R}^n)$. Then we have

$$(T_a u_k)(\varphi) = u_k(\overline{T_a^* \overline{\varphi}}) \to u(\overline{T_a^* \overline{\varphi}}) = (T_a u)(\varphi),$$

so $T_a u_k \to T_a u$ in $\mathcal{S}'(\mathbb{R}^n)$ and, therefore, $T_a : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is continuous.

2.1.6 Kernel representation of pseudo-differential operators

Summarising, we can now write pseudo-differential operators in different ways:

$$\begin{aligned}
T_a f(x) &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} a(x,\xi)\widehat{f}(\xi)d\xi \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i (x-y)\cdot\xi} a(x,\xi) f(y) dy d\xi \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i z\cdot\xi} a(x,\xi) f(x-z) dz d\xi \\
&= \int_{\mathbb{R}^n} k(x,z) f(x-z) dz \\
&= \int_{\mathbb{R}^n} K(x,y) f(y) dy,
\end{aligned}$$

with kernels

$$K(x,y) = k(x,x-y), \quad k(x,z) = \int_{\mathbb{R}^n} e^{2\pi i z\cdot\xi} a(x,\xi) d\xi.$$

Theorem. *Let $a \in S^m$. Then the kernel $K(x,y)$ satisfies*

$$|\partial^\beta_{x,y} K(x,y)| \le C|x-y|^{-N}$$

for $N > m+n+|\beta|$ and $x \neq y$. Thus, for $x \neq y$, the kernel $K(x,y)$ is a smooth function, rapidly decreasing as $|x-y| \to \infty$.

Proof. We notice that $k(x,\cdot)$ is the inverse Fourier transform of $a(x,\cdot)$. It follows then that $(-2\pi i z)^\alpha \partial_z^\beta k(x,z)$ is the inverse Fourier transform with respect to ξ of the derivative $\partial_\xi^\alpha\big[(2\pi i\xi)^\beta a(x,\xi)\big]$, i.e.

$$(-2\pi i z)^\alpha \partial_z^\beta k(x,z) = \mathcal{F}_\xi^{-1}\left(\partial_\xi^\alpha\big[(2\pi i\xi)^\beta a(x,\xi)\big]\right)(z).$$

Since $(2\pi i\xi)^\beta a(x,\xi) \in S^{m+|\beta|}$ is a symbol of order $m+|\beta|$, we have that

$$\left|\partial_\xi^\alpha\big[(2\pi i\xi)^\beta a(x,\xi)\big]\right| \le C_{\alpha\beta}\langle\xi\rangle^{m+|\beta|-|\alpha|}.$$

Therefore, $\partial_\xi^\alpha\big[(2\pi i\xi)^\beta a(x,\xi)\big]$ is in $L^1(\mathbb{R}^n_\xi)$ with respect to ξ, if $|\alpha| > m+n+|\beta|$. Consequently, its inverse Fourier transform is bounded:

$$(-2\pi i z)^\alpha \partial_z^\beta k(x,z) \in L^\infty(\mathbb{R}^n_z) \text{ for } |\alpha| > m+n+|\beta|.$$

Since taking derivatives of $k(x,z)$ with respect to x do not change the argument, this implies the statement of the theorem.

2.1.7 Smoothing operators

We can define symbols of order $-\infty$ by setting $S^{-\infty} = \bigcap_{m\in\mathbb{R}} S^m$, so that $a \in S^{-\infty}$ if $a \in C^\infty$ and if

$$|\partial_x^\beta \partial_\xi^\alpha a(x,\xi)| \le A_{\alpha\beta N}(1+|\xi|)^{-N},$$

for all N.

Proposition. *Let $a \in S^{-\infty}$. Then the integral kernel K of T_a is smooth on $\mathbb{R}^n \times \mathbb{R}^n$.*

Proof. Since $a(x,\cdot) \in L^1(\mathbb{R}^n)$, we immediately get $k \in L^\infty(\mathbb{R}^n)$. Moreover,

$$\partial_x^\beta \partial_z^\alpha k(x,z) = \int_{\mathbb{R}^n} e^{2\pi i z\cdot\xi}(2\pi i\xi)^\alpha \partial_x^\beta a(x,\xi)d\xi.$$

Since $(2\pi i\xi)^\alpha \partial_x^\beta a(x,\xi)$ is absolutely integrable, it follows that from the Lebesgue's dominated convergence theorem that $\partial_x^\beta \partial_z^\alpha k$ is continuous. This is true for all α, β, hence k, and then also K, are smooth. Let us write $k_x(\cdot) = k(x,\cdot)$.

Corollary. *Let $a \in S^{-\infty}$. Then $k_x \in \mathcal{S}(\mathbb{R}^n)$. We have*

$$T_a f(x) = (k_x * f)(x)$$

and, consequently, $T_a f \in C^\infty(\mathbb{R}^n)$ for all $f \in \mathcal{S}'(\mathbb{R}^n)$.

We note that the convolution in the corollary is understood in the sense of distributions. We will prove it and discuss this notion in detail in the next section.

2.1.8 Convolution of distributions

We can write the convolution of two functions $f, g \in \mathcal{S}(\mathbb{R}^n)$ in the following way

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

where $(Rg)(x) = g(-x)$ and $(\tau_h g)(x) = g(x-h)$, so that

$$(\tau_x Rg)(z) = (Rg)(z-x) = g(x-z).$$

Recalling our identification of functions with distributions, we can write

$$(f * g)(x) = f(\tau_x Rg).$$

This can now be extended to distributions. Indeed, for $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, we can define

$$(u * \varphi)(x) = u(\tau_x R\varphi),$$

which makes sense since $\tau_x R\varphi \in \mathcal{S}(\mathbb{R}^n)$ and since $\tau_x, R : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ are continuous.

Lemma. *Let $u \in \mathcal{S}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then $u * \varphi \in C^\infty(\mathbb{R}^n)$.*

Proof. We can observe that $(u * \varphi)(x) = u(\tau_x R\varphi)$ is continuous in x since $\tau_x : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ and $u : \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ are continuous. The same applies when we look at derivatives in x, implying that $u * \varphi$ is smooth. Here we note that we are allowed to pass the limit through u since it is a continuous functional.

Now Corollary 2.1.7 follows from the fact that for $a \in S^{-\infty}$ we can write $T_a f(x) = (k_x * f)(x)$ with $k_x(\cdot) = k(x,\cdot) \in \mathcal{S}(\mathbb{R}^n)$. So

$$(T_a f)(x) = f(\tau_x R k_x).$$

If now $f \in \mathcal{S}'(\mathbb{R}^n)$, it follows that $T_a f \in C^\infty$ because of the continuity of $f(\tau_x R k_x)$ and all of its derivatives with respect to x.

2.1.9 L^2–boundedness of pseudo-differential operators

Theorem. *Let $a \in S^0$. Then $T_a : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a bounded linear operator.*

Proof. First of all, we note that by a standard functional analytic argument it is sufficient to show the boundedness inequality

$$||T_a f||_{L^2(\mathbb{R}^n)} \leq C||f||_{L^2(\mathbb{R}^n)} \tag{2.3}$$

only for $f \in \mathcal{S}(\mathbb{R}^n)$, with constant C independent of the choice of f. Indeed, let $f \in L^2(\mathbb{R}^n)$ and let $f_k \in \mathcal{S}(\mathbb{R}^n)$ be a sequence of rapidly decreasing functions such that $f_k \to f$ in $L^2(\mathbb{R}^n)$. Then by (2.3) we have

$$||T_a(f_k - f_m)||_{L^2(\mathbb{R}^n)} \leq C||f_k - f_m||_{L^2(\mathbb{R}^n)},$$

so $T_a f_k$ is a Cauchy sequence in $L^2(\mathbb{R}^n)$. By completeness there is some $g \in L^2(\mathbb{R}^n)$ such that $T_a f_k \to g$ in $L^2(\mathbb{R}^n)$. On the other hand $T_a f_k \to T_a f$ in $\mathcal{S}'(\mathbb{R}^n)$. By the uniqueness principle we have $T_a f = g \in L^2(\mathbb{R}^n)$. Passing to the limit in (2.3) applied to f_k, we get $||T_a f||_{L^2(\mathbb{R}^n)} \leq C||f||_{L^2(\mathbb{R}^n)}$, with the same constant C.

The proof of (2.3) will consist of two parts, and we follow [3] in this. In the first part we establish it for compactly supported (with respect to x) symbols and in the second part we will extend it to the general case of $a \in S^0$.

So, let us first assume that $a(x, \xi)$ has compact support with respect to x. This will allow us to use the Fourier transform with respect to x, in particular the formulae

$$a(x, \xi) = \int_{\mathbb{R}^n} e^{2\pi i x\cdot\lambda} \widehat{a}(\lambda, \xi) d\lambda, \quad \widehat{a}(\lambda, \xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\lambda} a(x, \xi) dx,$$

with absolutely convergent integrals. We will use the fact that $a(\cdot, \xi) \in C_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, so that $a(\cdot, \xi)$ is in the Schwartz space in the first variable. Consequently, we have $\widehat{a}(\cdot, \xi) \in \mathcal{S}(\mathbb{R}^n)$ uniformly in ξ. To see the uniformity, we can notice that

$$(2\pi i\lambda)^\alpha \widehat{a}(\lambda, \xi) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot\lambda} \partial_x^\alpha a(x, \xi) dx,$$

and hence $|(2\pi i\lambda)^\alpha \widehat{a}(\lambda, \xi)| \leq C_\alpha$ for all $\xi \in \mathbb{R}^n$. It follows that

$$\sup_{\xi \in \mathbb{R}^n} |\widehat{a}(\lambda, \xi)| \leq C_N (1 + |\lambda|)^{-N}$$

for all N. Now we can write

$$\begin{aligned} T_a f(x) &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} a(x, \xi) \widehat{f}(\xi) d\xi \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i x\cdot\xi} e^{2\pi i x\cdot\lambda} \widehat{a}(\lambda, \xi) \widehat{f}(\xi) d\lambda d\xi \\ &= \int_{\mathbb{R}^n} (Sf)(\lambda, x) d\lambda, \end{aligned}$$

where

$$(Sf)(\lambda, x) = e^{2\pi i x\cdot\lambda} (T_{\widehat{a}(\lambda, \xi)} f)(x).$$

Here $T_{\widehat{a}(\lambda,\xi)}f$ is just a Fourier multiplier with symbol independent of x, so by Plancherel's identity we get

$$||T_{\widehat{a}(\lambda,\xi)}f||_{L^2} = ||\widehat{T_{\widehat{a}(\lambda,\xi)}f}||_{L^2} = ||\widehat{a}(\lambda,\cdot)\widehat{f}||_{L^2} \leq \sup_{\xi\in\mathbb{R}^n} |\widehat{a}(\lambda,\xi)|||\widehat{f}||_{L^2} \leq C_N(1+|\lambda|)^{-N}||f||_{L^2},$$

for all $N \geq 0$. Hence we get

$$||T_a f||_{L^2} \leq \int_{\mathbb{R}^n} ||Sf(\lambda,\cdot)||_{L^2} d\lambda \leq C_N \int_{\mathbb{R}^n} (1+|\lambda|)^{-N}||f||_{L^2} d\lambda \leq C||f||_{L^2},$$

if we take $N > n$.

Now, to pass to symbols which are not necessarily compactly supported with respect to x, we will use the following inequality

$$\int_{|x-x_0|\leq 1} |T_a f(x)|^2 dx \leq C_N \int_{\mathbb{R}^n} \frac{|f(x)|^2 dx}{(1+|x-x_0|)^N}, \tag{2.4}$$

which holds for every for every $x_0 \in \mathbb{R}^n$ and for every $N \geq 0$, with C_N independent of x_0 and dependent only on constants in the symbolic inequalities for a. Let us show first that (2.4) implies (2.3). Writing $\chi_{|x-x_0|\leq 1}$ for the characteristic function of the set $|x-x_0| \leq 1$ and integrating (2.4) with respect to x_0 yields

$$\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \chi_{|x-x_0|\leq 1} |T_a f(x)|^2 dx \right) dx_0 \leq C_N \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \frac{|f(x)|^2 dx}{(1+|x-x_0|)^N} \right) dx_0.$$

Changing the order of integration, we arrive at

$$\mathrm{vol}(B(1)) \int_{\mathbb{R}^n} |T_a f(x)|^2 dx \leq \widetilde{C}_N \int_{\mathbb{R}^n} |f(x)|^2 dx,$$

which is (2.3). Let us now prove (2.4).

Let us prove it for $x_0 = 0$ first. We can write $f = f_1 + f_2$, where f_1 and f_2 are smooth functions such that $|f_1| \leq |f|$, $|f_2| \leq |f|$, and $\operatorname{supp} f_1 \subset \{|x| \leq 3\}$, $\operatorname{supp} f_2 \subset \{|x| \geq 2\}$. We will do the estimate for f_1 first. Let us fix $\eta \in C_0^\infty(\mathbb{R}^n)$ such that $\eta(x) = 1$ for $|x| \leq 1$. Then $\eta T_a = T_{\eta a}$ is a pseudo-differential operator with a compactly supported symbol, thus by the first part we have

$$\int_{\{|x|\leq 1\}} |T_a f_1(x)|^2 dx = \int_{\mathbb{R}^n} |T_{\eta a} f_1(x)|^2 dx \leq C \int_{\mathbb{R}^n} |f_1(x)|^2 dx \leq C \int_{\{|x|\leq 3\}} |f(x)|^2 dx,$$

which is the required estimate for f_1. Let us now do the estimate for f_2. If $|x| \leq 1$, then $x \notin \operatorname{supp} f_2$, so we can write

$$T_a f_2(x) = \int_{\{|x|\geq 2\}} k(x, x-y) f_2(y) dy,$$

where k is the kernel of T_a. Since $|x| \leq 1$ and $|y| \geq 2$, we have $|x-y| \geq 1$ and hence by Theorem 2.1.6 we can estimate

$$|k(x, x-z)| \leq C_1 |x-y|^{-N} \leq C_2 |y|^{-N}$$

for all $N \geq 0$. Thus we can estimate

$$|T_a f_2(x)| \leq C_1 \int_{\{|y|\geq 2\}} \frac{|f_2(y)|}{|y|^N} dy \leq C_2 \int_{\mathbb{R}^n} \frac{|f(y)|}{(1+|y|)^N} dy \leq C_3 \Big(\int_{\mathbb{R}^n} \frac{|f(y)|^2}{(1+|y|)^N} dy \Big)^{1/2},$$

where by the Cauchy-Schwartz' inequality we have $C_3 = C_2 \Big(\int_{\mathbb{R}^n} \frac{1}{(1+|y|)^N} dy \Big)^{1/2} < \infty$ for $N > n$. This in turn implies

$$\int_{\{|x|\leq 1\}} |T_a f_2(x)|^2 dx \leq C \Big(\int_{\mathbb{R}^n} \frac{|f(y)|^2}{(1+|y|)^N} dy \Big)^{1/2},$$

which is the required estimate for f_2. These estimates for f_1 and f_2 imply (2.4) with $x_0 = 0$. We note that constant C_0 depends only on the dimension and on the constants in symbolic inequalities for a.

Let us now show (2.4) with an arbitrary $x_0 \in \mathbb{R}^n$. Let us define

$$a_{x_0}(x, \xi) = a(x - x_0, \xi).$$

Then we immediately see that estimate (2.4) for T_a in the ball $\{|x - x_0| \leq 1\}$ is equivalent to the same estimate for $T_{a_{x_0}}$ in the ball $\{|x| \leq 1\}$. Finally we note that since constants in symbolic inequalities for a and a_{x_0} are the same, we obtain (2.4) with constant C_N independent of x_0. This completes the proof of Theorem 2.1.9.

2.1.10 Compositions of pseudo-differential operators

Theorem. *Let $a \in S^{m_1}$ and $b \in S^{m_2}$. Then there exists some $c \in S^{m_1+m_2}$ such that $T_c = T_a \circ T_b$. Moreover, we have*

$$c \sim \sum_{\alpha} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} (\partial_\xi^\alpha a)(\partial_x^\alpha b),$$

which means that for all $N > 0$ we have

$$c - \sum_{|\alpha|<N} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} (\partial_\xi^\alpha a)(\partial_x^\alpha b) \in S^{m_1+m_2-N}.$$

Proof. Let us assume first that all symbols are compactly supported, so we can change the order of integration freely. Indeed, we can think of, say, symbol $a(x, \xi)$ as of

$$a_\epsilon(x, \xi) = a(x, \xi)\gamma(\epsilon x, \epsilon \xi),$$

make all the calculations uniformly in $0 < \epsilon \leq 1$, use that $a_\epsilon \in S^{m_1}$ uniformly in ϵ, and then take a limit as $\epsilon \to 0$. Let us now plug in

$$(T_b f)(y) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i (y-z)\cdot\xi} b(y, \xi) f(z)\, dz d\xi$$

into

$$(T_a g)(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\eta} a(x,\eta) g(y)\, dy d\eta$$

to get

$$\begin{aligned} T_a(T_b f)(x) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\eta} a(x,\eta) e^{2\pi i(y-z)\cdot\xi} b(y,\xi) f(z) dz d\xi dy d\eta \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-z)\cdot\xi} c(x,\xi) f(z)\, dz d\xi, \end{aligned}$$

with

$$\begin{aligned} c(x,\xi) &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot(\eta-\xi)} a(x,\eta) b(y,\xi)\, dy d\eta \\ &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot(\eta-\xi)} a(x,\eta)\widehat{b}(\eta-\xi,\xi)\, d\eta \\ &= \int_{\mathbb{R}^n} e^{2\pi i x\cdot\eta} a(x,\xi+\eta)\widehat{b}(\eta,\xi)\, d\eta. \end{aligned}$$

Here $\widehat{b}$ is the Fourier transform with respect to the first variable, and we used that

$$(x-y)\cdot\eta + (y-z)\cdot\xi - (x-z)\cdot\xi = (x-y)\cdot(\eta-\xi).$$

Asymptotic formula. Let us assume first that $b(x,\xi)$ has compact support in x (although if we think of b as b_ϵ, it will have a compact support, but the size of the support is not uniform in ϵ; so, we can approach the case of general symbols by treating the compactly supported symbols in a way of getting estimates that are uniform with respect to the size of the support). Since b is compactly supported in x, its Fourier transform with respect to x is rapidly decaying, so we can estimate

$$|\widehat{b}(\eta,\xi)| \le C_M(1+|\eta|)^{-M}(1+|\xi|)^{m_2},$$

for all $M \ge 0$. The Taylor's formula for $a(x,\xi+\eta)$ in the second variable gives

$$a(x,\xi+\eta) = \sum_{|\alpha|<N} \frac{1}{\alpha!}\partial_\xi^\alpha a(x,\xi)\eta^\alpha + R_N(x,\xi,\eta),$$

with remainder R_N that we will analyse later. Plugging this formula in to our expression for c and looking at terms in the sum, we get

$$\int_{\mathbb{R}^n} e^{2\pi i x\cdot\eta}\big[\partial_\xi^\alpha a(x,\xi)\eta^\alpha\big]\widehat{b}(\eta,\xi)\, d\eta = (2\pi i)^{-|\alpha|}\partial_\xi^\alpha a(x,\xi)\partial_x^\alpha b(x,\xi),$$

so that we have

$$c(x,\xi) = \frac{(2\pi i)^{-|\alpha|}}{\alpha!}\sum_{|\alpha|<N}\partial_\xi^\alpha a(x,\xi)\partial_x^\alpha b(x,\xi) + \int_{\mathbb{R}^n} e^{2\pi i x\cdot\eta} R_N(x,\xi,\eta)\widehat{b}(\eta,\xi)\, d\eta.$$

Remainder. For the remainder in the Taylor's series we have the estimate

$$|R_N(x,\xi,\eta)| \le C_N|\eta|^N \max\big\{|\partial_\xi^\alpha a(x,\zeta)| :\ |\alpha| = N, \zeta \text{ on the line between } \xi \text{ and } \xi+\eta\big\}.$$

Now, if $|\eta| \leq |\xi|/2$, then points on the line between ξ and $\xi+\eta$ are proportional to ξ, so we can estimate

$$|R_N(x,\xi,\eta)| \leq C_N|\eta|^N(1+|\xi|)^{m_1-N} \quad (|\eta| \leq |\xi|/2).$$

On the other hand, if $N \geq m_1$, we get the following estimate for all ξ and η:

$$|R_N(x,\xi,\eta)| \leq C_N|\eta|^N \quad (N \geq m_1).$$

Using the estimate for $\widehat{b}(\eta,\xi)$ and these two estimates, we get

$$\left|\int_{\mathbb{R}^n} e^{2\pi i x\cdot\eta} R_N(x,\xi,\eta)\widehat{b}(\eta,\xi)\,d\eta\right| \leq C_{N,M}(1+|\xi|)^{m_1+m_2-N}\int_{\mathbb{R}^n}(1+|\eta|)^{-M}|\eta|^N\,d\eta$$
$$+(1+|\xi|)^{m_2}\int_{|\eta|\geq|\xi|/2}(1+|\eta|)^{-M}|\eta|^N\,d\eta.$$

Taking M large enough, we can estimate both terms on the right hand side by $C(1+|\xi|)^{m_1+m_2-N}$. Making an estimate for $\partial_x^\alpha\partial_\xi^\beta R_N$ in a similar way, we can estimate

$$\left|\int_{\mathbb{R}^n} e^{2\pi i x\cdot\eta}\partial_x^\alpha\partial_\xi^\beta R_N(x,\xi,\eta)\widehat{b}(\eta,\xi)\,d\eta\right| \leq C(1+|\xi|)^{m_1+m_2-N-|\beta|},$$

implying the statement of the theorem for b compactly supported with respect to x.

General symbols. We note that it is sufficient to have the asymptotic formula and an estimate for the remainder for x near some fixed point x_0, uniformly in x_0. Let $\chi \in C_0^\infty(\mathbb{R}^n)$ be such that $\operatorname{supp}\chi \subset \{x : |x-x_0| \leq 2\}$ and such that $\chi(x)=1$ for $|x-x_0| \leq 1$. Let us decompose

$$b = \chi b + (1-\chi)b = b_1 + b_2.$$

Since symbol $b_1 = \chi b$ is compactly supported with respect to x, the composition formula for $T_a \circ T_{b_1}$ is given by the theorem, and it is equal to the claimed series for $T_a \circ T_b$ near x_0. We now have to show that the operator $T_a \circ T_{b_2}$ is smoothing, i.e. its symbol $c_2(x,\xi)$ is of order $-\infty$, and does not change the asymptotic formula for the composition. Indeed, we already know that the symbol of operator $T_a \circ T_{b_2}$ is given by

$$c_2(x,\xi) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot(\eta-\xi)}a(x,\eta)b_2(y,\xi)\,dyd\eta,$$

and we claim that

$$|c_2(x,\xi)| \leq C_N(1+|\xi|)^{m_1+m_2-N} \quad \text{for all } N \geq 0,\ |x-x_0| \leq \frac{1}{2}.$$

In the above integral for c_2 we can integrate by part to derive various properties of its decay. For example, we can integrate by parts with respect to η using the identity

$$\Delta_\eta^{N_1} e^{2\pi i(x-y)\cdot(\eta-\xi)} = (-4\pi^2)^{N_1}|x-y|^{2N_1}e^{2\pi i(x-y)\cdot(\eta-\xi)},$$

to see that in the integral we can replace $a(x,\eta)$ by $\frac{\Delta_\eta^{N_1} a(x,\eta)}{(-4\pi^2)^{N_1}|x-y|^{2N_1}}$, i.e.

$$c_2(x,\xi) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot(\eta-\xi)} \frac{\Delta_\eta^{N_1} a(x,\eta)}{(-4\pi^2)^{N_1}|x-y|^{2N_1}} b_2(y,\xi)\, dyd\eta,$$

Here $|x-x_0| \leq 1/2$ and $y \in \text{supp}(1-\chi)$ implies $|x-y| \geq 1/2$, so the integration by parts is well defined. We can also integrate by parts with respect to y using the identity

$$(1-\Delta_y)^{N_2} e^{2\pi i(x-y)\cdot(\eta-\xi)} = (1+4\pi^2|\xi-\eta|^2)^{N_2} e^{2\pi i(x-y)\cdot(\eta-\xi)}.$$

Moreover, we can use that $1+|\xi| \leq 1+|\xi-\eta|+|\eta| \leq (1+|\xi-\eta|)(1+|\eta|)$, and hence

$$\frac{1}{(1+|\xi-\eta|)^{2N_2}} \leq \frac{(1+|\eta|)^{2N_2}}{(1+|\xi|)^{2N_2}}.$$

Thus, integrating by parts with respect to y and using this estimate together with symbolic estimates for a and b_2, we get

$$\begin{aligned}|c_2(x,\xi)| &\leq C\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{(1+|\eta|)^{m_1-2N_1}}{(1+|x-y|)^{2N_1}} \frac{(1+|\xi|)^{m_2}}{(1+|\xi-\eta|)^{2N_2}}\, d\eta \\ &\leq C\int_{\mathbb{R}^n} (1+|\eta|)^{m_1-2N_1+2N_2}(1+|\xi|)^{m_2-2N_2}\, dyd\eta \\ &\leq C(1+|\xi|)^{m_1+m_2-N}\int_{\mathbb{R}^n}(1+|\eta|)^{N-2N_1}\, d\eta,\end{aligned}$$

if we take $N = m_1 - 2N_2$. Now, taking large N_2 and $2N_1 > N+n$, we obtain the desired estimate for $c_2(x,\xi)$. Similar estimates can be done for $\partial_x^\alpha\partial_\xi^\beta c_2(x,\xi)$ simply using symbolic inequalities for a and b_2, so we obtain that $c_2 \in S^{-\infty}$ for $|x-x_0| \leq 1/2$.

Finally, we notice that (in the general case) we have never used any information on the size of the support of symbols, so all constants depend only on the constants in symbolic inequalities. Thus, they remain uniformly bounded in $0 < \epsilon \leq 1$ for the composition $T_{c_\epsilon} = T_{a_\epsilon} \circ T_{b_\epsilon}$, and we have $c_\epsilon \in S^{m_1+m_2}$ with symbolic constants uniform in ϵ. Moreover, the asymptotic formula is satisfied uniformly in ϵ. Since $c_\epsilon \to c$ pointwise as $\epsilon \to 0$, we can conclude that $c \in S^{m_1+m_2}$, that $T_c = T_a \circ T_b$, and that the asymptotic formula of the theorem holds.

2.1.11 Amplitudes

We have already seen that when taking the adjoint of a pseudo-differential operator, we get the symbol which depends on the "wrong" set of variables: (y,ξ) instead of (x,ξ). Nevertheless, we want the adjoint to be a pseudo-differential operator as well. For this we need to study operators with symbols depending on all combinations of variables. This leads to the definition of an amplitude. We will write $c = c(x,y,\xi) \in \widetilde{S^m}$ if $c \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n)$ and if

$$\left|\partial_y^\gamma \partial_x^\beta \partial_\xi^\alpha c(x,y,\xi)\right| \leq C_{\alpha,\beta,\gamma}(1+|\xi|)^{m-|\alpha|}$$

holds for all $x,y,\xi \in \mathbb{R}^n$ and all multi-indices α,β,γ. The corresponding operator $T_{[c]}$ is defined by

$$(T_{[c]}f)(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\xi} c(x,y,\xi) f(y)\, dyd\xi.$$

As in 2.1.4, we can justify this formula by considering $c_\epsilon(x,y,\xi) = c(x,y,\xi)\gamma(\epsilon y, \epsilon\xi)$, with γ as in 2.1.4. Then $c_\epsilon \to c$ pointwise (also with the pointwise convergence of derivatives), uniformly in $\widetilde{S^m}$ for $0 < \epsilon \leq 1$, so $T_{[c_\epsilon]}f \to T_{[c]}f$ in $\mathcal{S}(\mathbb{R}^n)$ for all $f \in \mathcal{S}(\mathbb{R}^n)$. Thus, $T_{[c]}$ is well defined and continuous as an operator from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}(\mathbb{R}^n)$.

2.1.12 Symbols of pseudo-differential operators

Proposition. *A pseudo-differential operator $T \in \Psi^m$ defines its symbol $a \in S^m$ uniquely, so that $T = T_a$. The symbol $a(x,\xi)$ is defined by the formula*

$$a(x,\xi) = e^{-2\pi i x\cdot\xi} T(e^{2\pi i x\cdot\xi}).$$

The notation used in the statement is a useful abbreviation for

$$e^{-2\pi i x\cdot\xi} T(e^{2\pi i x\cdot\xi}) = e^{-2\pi i x\cdot\xi} \left(T(e^{2\pi i y\cdot\xi})\right)(x) = e^{-2\pi i x\cdot\xi} \left(T(e^{2\pi i\langle\cdot,\xi\rangle})\right)(x).$$

The formula for the symbol can be justified either in $\mathcal{S}'(\mathbb{R}^n)$ or as the limit in the expression

$$e^{-2\pi i x\cdot\xi} T(e^{2\pi i x\cdot\xi}\varphi_\epsilon) \to a(x,\xi),$$

as $\varphi_\epsilon \to 1$ for a family of $\varphi_\epsilon \in C_0^\infty(\mathbb{R}^n)$.

2.1.13 Symbols of amplitude operators

Proposition. *Let $c \in \widetilde{S^m}$ be an amplitude. Then there exists a symbol $a \in S^m$ such that $T_a = T_{[c]}$. Moreover, the asymptotic expansion for a is given by*

$$a(x,\xi) - \sum_{|\alpha|<N} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} \partial_\xi^\alpha \partial_y^\beta c(x,y,\xi)|_{y=x} \in S^{m-N}, \quad \forall N \geq 0.$$

Proof. The proof is similar to the proof of the composition formula 2.1.10. As in that proof, we first formally conclude that we must have

$$\begin{aligned} a(x,\xi) &= e^{-2\pi i x\cdot\xi} T(e^{2\pi i x\cdot\xi}) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-2\pi i x\cdot\xi} e^{2\pi i(x-y)\cdot\eta} c(x,y,\eta) e^{2\pi i y\cdot\xi} dy d\eta \\ &= \int_{\mathbb{R}^n} c(x,y,\eta) e^{2\pi i(x-y)\cdot(\eta-\xi)} dy d\eta \\ &= \int_{\mathbb{R}^n} \widehat{c}(x,\eta,\xi+\eta) e^{2\pi i x\cdot\eta} d\eta, \end{aligned}$$

where $\widehat{c} = \mathcal{F}_y c$ is the Fourier transform of c with respect to y and where we used the change of variables $\eta \mapsto \eta + \xi$. By the usual justification, we may work with amplitudes compactly supported in y, and make sure that all our arguments do not depend on the size of the support of c in y. Taking the Taylor's expansion of $\widehat{c}(x,\eta,\xi+\eta)$ at ξ, we obtain

$$\widehat{c}(x,\eta,\xi+\eta)=\sum_{|\alpha|<N}\frac{1}{\alpha!}\partial_\xi^\alpha\widehat{c}(x,\eta,\xi)\eta^\alpha+R_N(x,\eta,\xi),$$

where, as before, we can show that the remainder $R_N(x,\eta,\xi)$ satisfies estimates

$$|R_N(x,\eta,\xi)|\leq A|\eta|^N(1+|\eta|)^{-M}(1+|\xi|)^{m-N}\ \text{ for large } M \text{ and for } 2|\eta|\leq|\xi|,$$

and

$$|R_N(x,\eta,\xi)|\leq A|\eta|^N(1+|\eta|)^{-M}\ \text{ for all large } M \text{ and } N.$$

The last estimate is used in the region $2|\eta|\geq|\xi|$ and, similar to the proof of the composition formula in Section 2.1.10, we can complete the proof in the case of $c(x,y,\xi)$ compactly supported in y.

To treat the general case where $c(x,y,\xi)$ is not necessarily compactly supported in y we reduce the analysis to the case when $c(x,y,\xi)$ vanishes for y away from $y=x$. For this, we can use the same argument as in the proof of the composition formula in Section 2.1.10, where the estimate for the remainder by the integration by parts argument becomes

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{(1+|\eta|)^{m-2N_1}}{(1+|\eta-\xi|)^{2N_2}}\frac{1}{(1+|x-y|)^{2N_1}}dyd\eta=O((1+|\xi|)^{m-N}),$$

for large N_1,N_2. The proof is complete.

Remark. Note that we can also justify the following alternative argument. We know by Proposition 2.1.12 that for $T\in\Psi^m$ its symbols can be defined by the formula

$$a(x,\xi)=e^{-2\pi ix\cdot\xi}\left(T(e^{2\pi iy\cdot\xi})\right)(x).$$

Let now $T=T_{[c]}$ be the operator with amplitude c. Then Proposition 2.1.13 says that we can write

$$(T_{[c]}u)(x)=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{2\pi i(x-y)\cdot\xi}a(x,\xi)u(y)dyd\xi,$$

for all $u\in C_0^\infty(\mathbb{R}^n)$. To show this, we write the Fourier inversion formula

$$u(y)=\int_{\mathbb{R}^n}e^{2\pi iy\cdot\xi}\widehat{u}(\xi)d\xi,$$

and, justifying the application of T to it and using formula for the symbol from Proposition 2.1.12, we get

$$(Tu)(x)=\int_{\mathbb{R}^n}T(e^{2\pi iy\cdot\xi})(x)\widehat{u}(\xi)d\xi=\int_{\mathbb{R}^n}e^{2\pi ix\cdot\xi}a(x,\xi)\widehat{u}(\xi)d\xi.$$

For the asymptotic expansion, we write

$$\begin{aligned}(Te^{2\pi i\langle\cdot,\xi\rangle})(x)&=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{2\pi i(x-y)\cdot\eta}c(x,y,\eta)e^{2\pi iy\cdot\xi}dyd\eta\\&=e^{2\pi ix\cdot\xi}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}e^{-2\pi iy\cdot\eta}c(x,x+y,\xi+\eta)dyd\eta,\end{aligned}$$

where we changed variables $y \mapsto y + x$ and $\eta \mapsto \eta + \xi$, as well as recalculated the phase in new variables as

$$(x - (y + x)) \cdot (\eta + \xi) + (y + x) \cdot \xi = -y \cdot \eta + x \cdot \xi.$$

Now, using the Taylor's expansion and estimates for the remainder, we can obtain Proposition 2.1.13 again.

2.1.14 Adjoint operators

Theorem. *Let $a \in S^m$. Then there exists a symbol $a^* \in S^m$ such that $T_a^* = T_{a^*}$, where T_a^* is the L^2–adjoint operator of T_a. Moreover, we have the asymptotic expansion*

$$a^*(x, \xi) - \sum_{|\alpha|<N} \frac{(2\pi i)^{-|\alpha|}}{\alpha!} \partial_\xi^\alpha \partial_x^\alpha \overline{a}(x, \xi) \in S^{n-N}, \quad \text{for all } N \geq 0,$$

where $\overline{a}(x, \xi)$ is the complex conjugate of $a(x, \xi)$.

Proof. As we have already calculated, we have $T_{a^*} = T_{[c]}$, which is an operator with amplitude $c(x, y, \xi) = \overline{a}(y, \xi)$. Applying Proposition 2.1.13, we obtain the statement of Theorem 2.1.14.

2.1.15 Changes of variables

It is clear from the definition of the symbol class S^m that it is locally invariant under smooth changes of variables, i.e. if we take a local change of variable x in the symbol from S^m, it will still belong to the same symbol class S^m. We will now investigate what happens with pseudo-differential operators when we make a change of variable in space.

Let $U, V \subset \mathbb{R}^n$ be open bounded sets and let $\tau : V \to U$ be a surjective diffeomorphism. For $f \in C_0^\infty(V)$, we define its pullback by the change of variables τ by

$$(\tau f)(x) = f(\tau^{-1}(x)).$$

It easily follows that the new function satisfies $\tau f \in C_0^\infty(U)$. Let now $a \in S^m$ be a symbol of order m with compact support with respect to x, and assume that this support is contained in U. Then we will show that there exists a symbol $b \in S^m$ with compact support with respect to x which is contained in V such that $\tau^{-1} T_a \tau = T_b$. In other words, we have

$$\tau^{-1} T_a \tau f = T_b f \quad \text{for all } f \in C_0^\infty(V).$$

More precisely, we have the following expression for the "main part" of b:

Proposition. *We have $b(x, \xi) = a(\tau(x), \left[\frac{\partial \tau}{\partial x}\right]' \xi)$ modulo S^{m-1}, where $\left[\frac{\partial \tau}{\partial x}\right]' = ((D\tau)^T)^{-1}$.*

Proof. We can write

$$\begin{aligned}(\tau^{-1} T_a \tau f)(x) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i(\tau(x)-y)\cdot\xi} a(\tau(x), \xi) f(\tau^{-1}(y)) dy d\xi \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{2\pi i(\tau(x)-\tau(y))\cdot\xi} a(\tau(x), \xi) f(y) |\det \frac{\partial \tau}{\partial y}| dy d\xi, \qquad (2.5)\end{aligned}$$

where we changed variables $y \mapsto \tau(y)$ in to get the last equality. Now we will argue that the main contribution in this integral comes from variables y close to x. Indeed, let us insert a cut-off function $\chi(x, y)$ in the integral, where χ is a smooth function supported in the set where $|y-x|$ is small and where $\chi(x, x) = 1$ for all x. The remaining integral

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(\tau(x)-\tau(y))\cdot\xi}(1-\chi(x,y))a(\tau(x),\xi)f(y)|\det\frac{\partial\tau}{\partial y}|dyd\xi$$

has a smooth kernel since we can integrate by parts any number of times with respect to ξ to see that the symbol of this operator belongs to symbol classes S^{-N} for all $N \geq 0$. Let us now analyse what happens for y close to x. By the mean value theorem, for y sufficiently close to x, we have

$$\tau(x) - \tau(y) = L_{x,y}(x-y), \tag{2.6}$$

where $L_{x,y}$ is an invertible linear mapping which is smooth in x and y, and satisfies

$$L_{x,x} = \frac{\partial\tau}{\partial x}(x).$$

Using (2.6) in (2.5), we get

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i L_{x,y}(x-y)\cdot\xi}\chi(x,y)a(\tau(x),\xi)|\det\frac{\partial\tau}{\partial y}|f(y)dyd\xi =$$
$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\xi}\chi(x,y)a(\tau(x),L'_{x,y}\xi)|\det\frac{\partial\tau}{\partial y}||\det L^{-1}_{x,y}|f(y)dyd\xi,$$

where we changed variables $L^T_{x,y}\xi \mapsto \xi$ and where $L^T_{x,y}$ is the transpose matrix of $L_{x,y}$ and $L'_{x,y} = (L^T_{x,y})^{-1}$. Thus, we get an operator with the amplitude c defined by

$$c(x,y,\xi) = \chi(x,y)a(\tau(x),L'_{x,y}\xi)|\det\frac{\partial\tau}{\partial y}||\det L^{-1}_{x,y}|.$$

Applying the asymptotic expansion in Proposition 2.1.13, we see that the first term of this expansion is given by

$$c(x,y,\xi)|_{y=x} = \chi(x,x)a(\tau(x),L'_{x,x}\xi)|\det\frac{\partial\tau}{\partial x}||\det L^{-1}_{x,x}| = a(\tau(x),\left[\frac{\partial\tau}{\partial x}\right]'\xi),$$

completing the proof.

2.1.16 Principal symbol and classical symbols

We see in Proposition 2.1.15 that the equivalence class modulo S^{m-1} has some meaning for the changes of variables. In fact, we may notice that the transformation

$$(x,\xi) \mapsto (\tau(x),\left[\frac{\partial\tau}{\partial x}\right]'\xi)$$

is the same as the change of variables in the cotangent bundle $T^*\mathbb{R}^n$ of $\mathbb{R}^n$ which is induced by the change of variables $x \mapsto \tau(x)$ in $\mathbb{R}^n$. This observation allows one to make an invariant geometric interpretation of the class of symbols in S^m modulo terms of order S^{m-1}. We note that we can use the asymptotic expansion for amplitudes in the proof of Proposition 2.1.15 to find also lower order terms in the asymptotic expansion of $b(x,\xi)$, but most of these terms will not have such nice invariant interpretation. Without going into detail, let us just mention that Proposition 2.1.15 allows to introduce a notion of a *principal symbol* of a pseudo-differential operator with symbol in S^m as the equivalence class of this symbol modulo the subclass S^{m-1}, and this principal symbol becomes a function on the cotangent bundle of $\mathbb{R}^n$. This construction can be further carried out on manifolds leading to many remarkable applications, in particular to those in geometry and index theory.

We will not pursue this path further in this chapter, but we will clarify the notion of such equivalent classes for the so-called *classical symbols* which is a class of symbols that plays a very important role in applications to partial differential equations. First, we define homogeneous functions/symbols.

Definition. *We will say that a symbol $a_k = a_k(x,\xi) \in S^k$ is* positively homogeneous *(or simply* homogeneous*) of order k if for all $x \in \mathbb{R}^n$ we have*

$$a_k(x,\lambda\xi) = \lambda^k a_k(x,\xi) \quad \text{for all } \lambda > 0 \text{ and all } |\xi| > 1.$$

We note that we exclude small ξ from this definition because if we assumed property $a_k(x,\lambda\xi) = \lambda^k a_k(x,\xi)$ for all $\xi \in \mathbb{R}^n$, such function a_k would not be in general smooth at $\xi = 0$.

Definition. *We will say that a symbol $a \in S^m_{cl}$ is a* classical symbol of order m *if $a \in S^m$ and if there is an asymptotic expansion $a \sim \sum_{k=0}^{\infty} a_{m-k}$, where each a_{m-k} is homogeneous of order $m-k$, and if $a - \sum_{k=0}^{N} a_{m-k} \in S^{m-N-1}$, for all $N \geq 0$.*

Now, for a classical symbol $a \in S^m_{cl}$ the principal symbol, i.e. its equivalence class modulo S^{m-1} can be easily found. This is simply the first term a_m in the asymptotic expansion in its definition. We will discuss asymptotic sums in more detail in the next section, namely we will show that if we start with the asymptotic sum $\sum_{k=0}^{\infty} a_{m-k}$, we can in turn interpret it as a symbol from S^m.

2.1.17 Asymptotic sums

Our next objective is to show how the constructed calculus of pseudo-differential operators can be applied to "easily" solve so-called elliptic partial differential equations. However, in order to carry out this application we will need another very useful construction.

Proposition. *Let $a_j \in S^{m_j}$, $j = 0,1,2,\ldots$, where the sequence m_j of orders satisfies $m_0 > m_1 > m_2 > \cdots$ and $m_j \to -\infty$ as $j \to \infty$. Then there exists a symbol $a \in S^{m_0}$ such that*

$$a \sim \sum_{j=0}^{\infty} a_j = a_0 + a_1 + a_2 + \ldots,$$

which means that we have

$$a - \sum_{j=0}^{k} a_j \in S^{m_k}, k = 1,2,\ldots$$

We can also note that in this notation $a \sim 0$ if and only if $a \in S^{-\infty} = \cap_{m\in\mathbb{R}} S^m$.

Proof. Let us fix a function $\chi \in C^\infty(\mathbb{R}^n)$ such that $\chi(\xi) = 1$ for all $|\xi| \geq 1$ and such that $\chi(\xi) = 0$ for all $|\xi| \leq 1/2$. Then, for some sequence τ_j increasing sufficiently fast and to be chosen later, we define

$$a(x,\xi) = \sum_{j=0}^{\infty} a_j(x,\xi)\chi\left(\frac{\xi}{\tau_j}\right).$$

We note that this sum is well-defined pointwise because it is in fact locally finite since $\chi\left(\frac{\xi}{\tau_j}\right) = 0$ for $|\xi| < \tau_j/2$. In order to show that $a \in S^{m_0}$ we first take a sequence τ_j such that the inequality

$$\left|\partial_x^\beta \partial_\xi^\alpha \left[a_j(x,\xi)\chi\left(\frac{\xi}{\tau_j}\right)\right]\right| \leq 2^{-j}(1+|\xi|)^{m_{j-1}-1-|\alpha|} \tag{2.7}$$

is satisfied for all $|\alpha|, |\beta| \leq j$. We note that because the sequence m_j is strictly decreasing we can also replace the right hand side by $2^{-j}(1+|\xi|)^{m_j+1-|\alpha|}$. We first show that function $\xi^\alpha \partial_\xi^\alpha \chi\left(\frac{\xi}{\tau_j}\right)$ is uniformly bounded in ξ for each j. Indeed, we have

$$\xi^\alpha \partial_\xi^\alpha \chi\left(\frac{\xi}{\tau_j}\right) = \begin{cases} 0, & |\xi| < \tau_j/2, \\ \text{bounded by } C\left|\left(\frac{\xi}{\tau_j}\right)^\alpha\right|, & \tau_j/2 \leq |\xi| \leq |\tau|, \\ 0, & \tau_j < |\xi|, \end{cases}$$

so that $\left|\xi^\alpha \partial_\xi^\alpha \chi\left(\frac{\xi}{\tau_j}\right)\right| \leq C$ is uniformly bounded for all ξ, for any given j. Using this fact, we can also estimate

$$\begin{aligned}
\left|\partial_x^\beta \partial_\xi^\alpha \left[a_j(x,\xi)\chi\left(\frac{\xi}{\tau_j}\right)\right]\right| &= \left|\sum_{\alpha_1+\alpha_2=\alpha} c_{\alpha_1\alpha_2} \partial_x^\beta \partial_\xi^{\alpha_1} a_j(x,\xi) \partial_\xi^{\alpha_2} \chi\left(\frac{\xi}{\tau_j}\right)\right| \\
&\leq \sum_{\alpha_1+\alpha_2=\alpha} |c_{\alpha_1\alpha_2}|(1+|\xi|)^{m_j-|\alpha_1|}(1+|\xi|)^{-|\alpha_2|} \\
&\leq C(1+|\xi|)^{m_j-|\alpha|} \\
&= \left[C(1+|\xi|)^{-1}\right](1+|\xi|)^{m_j+1-|\alpha|}.
\end{aligned}$$

Now, the left hand side in estimate (2.7) is zero for $|\xi| < \tau_j/2$, so we may assume that $|\xi| \geq \tau_j/2$. Hence we can have

$$C(1+|\xi|)^{-1} \leq C(1+|\tau_j/2|)^{-1} < 2^{-j}$$

if we take τ_j sufficiently large. This implies that we can take the sum of $\partial_x^\beta \partial_\xi^\alpha$–derivatives in the definition of $a(x,\xi)$ and (2.7) implies that $a \in S^{m_0}$. Finally, to show the asymptotic formula, we can write

$$a - \sum_{j=0}^{k} a_j = \sum_{j=k+1}^{\infty} a_j(x,\xi)\chi\left(\frac{\xi}{\tau_j}\right),$$

and so

$$\left|\partial_x^\beta \partial_\xi^\alpha \left[a - \sum_{j=0}^{k} a_j\right]\right| \leq C(1 + |\xi|)^{m_k - |\alpha|}.$$

In this arguments we fix α and β first, and then use the required estimates for all $j \geq |\alpha|, |\beta|$. This shows that $a - \sum_{j=0}^{k} a_j \in S^{m_k}$ finishing the proof.

2.2 Applications to partial differential equations

2.2.1 Solving partial differential equations

The main question in the theory of partial differential equations is how to solve the equation

$$Lu = f$$

for a given partial differential operator L and a given function f. In other words, how to find the *inverse of* L, i.e. an operator L^{-1} such that

$$L \circ L^{-1} = L^{-1} \circ L = I \tag{2.8}$$

is the identity operator (on some space of functions where everything is well-defined). In this case function $u = L^{-1} f$ gives a solution to the partial differential equation $Lu = f$.

First of all we can observe that if operator L is an operator with variables coefficients in most cases it is impossible or is very hard to find an explicit formula for its inverse L^{-1} (even when it exists). However, in many questions in the theory of partial differential equations one is actually not so much interested in having a precise explicit formula for L^{-1}. Indeed, in reality one is mostly interested not in knowing the solution u to the equation $Lu = f$ explicitly but rather in knowing some fundamental properties of u. Among the most important properties are the position and the strength of singularities of u. Thus, the question becomes whether we can say something about singularities of u knowing singularities of $f = Lu$. In this case we do not need to solve equation $Lu = f$ exactly but it is sufficient to know its solution modulo the class of smooth functions. Namely, instead of L^{-1} in (2.8) one is interested in finding an "approximate" inverse of L modulo smooth functions, i.e. an operator P such that

$$u = Pf$$

solves equation $Lu = f$ modulo smooth functions, i.e. if $(PL - I)f$ and $(LP - I)f$ are smooth for all functions f from some class. Recalling that operators in $\Psi^{-\infty}$ have such property, we have following definition.

Definition. *Operator P is called the right parametrix of L if $LP - I \in \Psi^{-\infty}$. Operator S is called the left parametrix of L if $SL - I \in \Psi^{-\infty}$.*

In fact, the left and right parametrix are closely related. Indeed, by definition we have $LP - I = R_1$ and $SL - I = R_2$ with some $R_1, R_2 \in \Psi^{-\infty}$. Then we have

$$S = S(LP - R_1) = (SL)P - SR_1 = P + R_2 P - SR_1.$$

If L, S, P are pseudo-differential operators of finite orders, the composition formulae imply that $R_2P, SR_1 \in \Psi^{-1}$, i.e. $S - P$ is a smoothing operator. Thus, we will be mainly interested in the right parametrix P because $u = Pf$ immediately solves equation $Lu = f$ modulo smooth functions.

We also note that since we work here modulo smoothing operators (i.e. operators in $\Psi^{-\infty}$), parametrices are obviously not unique – funding one of them is already very good because any two parametrices differ by a smoothing operator.

2.2.2 Elliptic equations

We will now show how we can use the calculus to "solve" elliptic partial differential equations. First, we recall the notion of ellipticity.

Definition. *A symbol $a \in S^m$ is called elliptic if for some $A > 0$ it satisfies*

$$|a(x, \xi)| \geq A|\xi|^m$$

for all $|\xi| \geq 1$ and all $x \in \mathbb{R}^n$. We also say that symbol a is elliptic in $U \subset \mathbb{R}^n$ if the above estimate holds for all $x \in U$. Pseudo-differential operators with elliptic symbols are also called elliptic.

Now, let $L = T_a$ be an elliptic pseudo-differential operator with symbol $a \in S^m$ (which is then also elliptic by definition). Let is introduce a cut-off function $\chi \in C^\infty(\mathbb{R}^n)$ such that $\chi(\xi) = 0$ for small ξ, e.g. for $|\xi| \leq 1$, and such that $\chi(\xi) = 1$ for large ξ, e.g. for $|\xi| > 2$. The ellipticity of $a(x, \xi)$ assures that it can be inverted pointwise for $|\xi| \geq 1$, so we can define symbol

$$b(x, \xi) = \chi(\xi) \left[a(x, \xi)\right]^{-1}.$$

Since $a \in S^m$ is elliptic, we easily see that $b \in S^{-m}$. If we take $P_0 = T_b$ then by composition theorems we obtain

$$LP_0 = I + E_1, \quad PL = I + E_2,$$

for some $E_1, E_2 \in \Psi^{-1}$. Thus, we may view P_0 as a good first approximation for a parametrix of L. In order to find a parametrix of L, we need to modify P_0 in such a way that E_1 and E_2 would be in $\Psi^{-\infty}$. This constructions can be carried out in an iterative way. We now give a slightly more general statement which is useful for other purposes as well.

Theorem. *Let $a \in S^m$ be elliptic on an open set $U \subset \mathbb{R}^n$, i.e. there exists some $A > 0$ such that*

$$|a(x, \xi)| \geq A|\xi|^m$$

for all $x \in U$ and all $|\xi| \geq 1$. Let $c \in S^l$ be a symbol of order l whose support with respect to x is a compact subset of U. Then there exists a symbol $b \in S^{l-m}$ such that

$$T_b T_a = T_c - T_e$$

for some symbol $e \in S^{-\infty}$.

Proof. In order to simplify the notation, in this proof we will write $a_1 \circ a_2 = a_3$ to express the relation $T_{a_1}T_{a_2} = T_{a_3}$. We will construct the symbol b as an asymptotic sum

$$b \sim b_0 + b_1 + b_2 + \cdots$$

and then use Proposition 2.1.17 to justify this infinite sum. We will also work with $|\xi| \geq 1$ since small ξ are not relevant for the symbolic constructions.

First, we take $b_0 = c/a$ which is well defined for $|\xi| \geq 1$ in view of the ellipticity of a. Then we have

$$b_0 \circ a = c - e_0, \quad b_0 \in S^{l-m}, \text{ with } e_0 = c - b_0 \circ a \in S^{l-1}.$$

Then we take $b_1 = e_0/a \in S^{l-m-1}$ so that we have

$$(b_0 + b_1) \circ a = c - e_0 + b_1 \circ a = c - e_1, \quad \text{ with } e_1 = e_0 - b_1 \circ a \in S^{l-2}.$$

Inductively, we define $b_j = e_{j-1}/a \in S^{l-m-j}$ and we have

$$(b_0 + b_1 + \cdots + b_j) \circ a = c - e_j, \quad \text{ with } e_j = e_{j-1} - b_j \circ a \in S^{l-j-1}.$$

Now, Proposition 2.1.17 shows that $b \in S^{l-m}$ and it satisfies $b \circ a = c - e$ with $e \in S^{-\infty}$ by its construction, completing the proof.

2.2.3 Parametrix for elliptic operators and estimates

Now, a parametrix for an elliptic partial differential operator L is given by Theorem 2.2.2 with $T_c = I$ being the identity operator. More precisely, we also have the following local version of this.

Corollary. *Let $L = \sum_{|\alpha| \leq m} a_\alpha(x)\partial_x^\alpha$ be an elliptic partial differential operator in an open set $U \subset \mathbb{R}^n$. Let $\chi_1, \chi_2, \chi_3 \in C_0^\infty(\mathbb{R}^n)$ be such that $\chi_2 = 1$ on the support of χ_1 and $\chi_3 = 1$ on the support of χ_2. Then there is an operator $P \in S^{-m}$ such that*

$$P(\chi_2 L) = \chi_1 I + E\chi_3,$$

for some $E \in \Psi^{-\infty}$.

Proof. We take $T_a = \chi_2 L$ and $T_c = \chi_1 I$. Then T_a is elliptic on the support of χ_2 and we can take $P = T_b$ with $b \in S^{-m}$ from Theorem 2.2.2.

We will now apply this result to obtain a statement on the regularity of solution to elliptic partial differential equations. We assume that the order m below is an integer which is certainly true when L is a partial differential operator. However, if we take into account the discussion from the next section, we will see that the statements below are still true for any $m \in \mathbb{R}$.

Theorem. *Let $L \in \Psi^m$ be an elliptic pseudo-differential operator in an open set $U \subset \mathbb{R}^n$ and let $Lu = f$ in U. Assume that $f \in (L^2_k)_{loc}$. Then $u \in (L^2_{m+k})_{loc}$.*

This theorem shows that if u is a solution of an elliptic partial differential equation $Lu = f$ then there is local gain of m derivatives for u compared to f, where m is the order of the operator L.

Proof. Let $\chi_1, \chi_2, \chi_3 \in C_0^\infty(U)$ be non-zero functions such that $\chi_2 = 1$ on the support of χ_1 and $\chi_3 = 1$ on the support of χ_2. Then, similar to the proof of Corollary 2.2.3 we have

$$P(\chi_2 L) = \chi_1 I + E\chi_3,$$

with some $P \in \Psi^{-m}$. Since $f \in (L_k^2)_{loc}$ we have $P(\chi_2 f) \in (L_{m+k}^2)_{loc}$. Also, $E(\chi_3 u) \in C^\infty$ so that $||\chi_3 E(\chi_3 u)||_{L_k^2} \le ||\chi_3 u||_{L^2}$ for any k. Summarising and using

$$P(\chi_2 f) = \chi_1 u + E(\chi_3 u),$$

we obtain

$$||\chi_1 u||_{L_{k+m}^2} \le C\left(||\chi_2 f||_{L_k^2} + ||\chi_3 u||_{L^2}\right),$$

which implies that $u \in (L_{m+k}^2)_{loc}$.

Remark. We can observe from the proof that properties of solution u by the calculus and the existence of a parametrix are reduced to the fact that pseudo-differential operators in Ψ^{-m} map L_k^2 to L_{k+m}^2. In fact, in this way many properties of solutions to partial differential equations are reduced to questions about general pseudo-differential operators. In particular, let us assume without proof the following statement where for now we will take μ and k to be integers or zeros. However, if we adopt the definition of Sobolev spaces from the next section, this statement is valid for all $\mu, k \in \mathbb{R}$.

Statement on L^p–continuity. *Let $T \in S^{-\mu}$ be a pseudo-differential operator of order μ and let $1 < p < \infty$. Then T is bounded from the Sobolev space $L_k^p(\mathbb{R}^n)$ to the Sobolev space $L_{k+\mu}^p(\mathbb{R}^n)$.*

If μ and k are integers or zero, we will prove this statement in the next section. As an immediate consequence, by the same argument as in the proof of the theorem, we also obtain

Corollary. *Let $L \in \Psi^m$ be an elliptic pseudo-differential operator in an open set $U \subset \mathbb{R}^n$, let $1 < p < \infty$, and let $Lu = f$ in U. Assume that $f \in (L_k^p)_{loc}$. Then $u \in (L_{m+k}^p)_{loc}$.*

2.2.4 Sobolev spaces revisited

Up to now we defined Sobolev spaces L_k^p assuming that the index k is an integer. In fact, using the calculus of pseudo-differential operators we can show that these spaces can be defined for all $k \in \mathbb{R}$ thus allowing one to measure the regularity of functions much more precisely. In the following discussion we assume the statement on the L^p–continuity of pseudo-differential operators from Section 2.2.3.

We recall from Section 1.3.11 that for an integer $k \in \mathbb{N}$ we defined the Sobolev space $L_k^p(\mathbb{R}^n)$ as the space of all $f \in L^p(\mathbb{R}^n)$ such that their distributional derivatives satisfy $\partial_x^\alpha f \in L^p(\mathbb{R}^n)$, for all $0 \le |\alpha| \le k$. This space is equipped with a norm

$$||f||_{L_k^p} = \sum_{|\alpha| \le k} ||\partial_x^\alpha f||_{L^p}$$

(or with any equivalent norm).

Let now $s \in \mathbb{R}$ be a real number and let us consider operators $(I-\Delta)^{s/2} \in \Psi^s$ which are pseudo-differential operators with symbols $a(x,\xi) = (1+4\pi^2|\xi|^2)^{s/2}$. We will say that

$$f \in L^p_s(\mathbb{R}^n) \text{ if } (I-\Delta)^{s/2} f \in L^p(\mathbb{R}^n).$$

We equip this space with the norm

$$||f||_{L^p_s} = ||(I-\Delta)^{s/2} f||_{L^p}.$$

Proposition. *If $s \in \mathbb{N}$ is an integer, the space $L^p_s(\mathbb{R}^n)$ coincides with the space L^p_k with $k = s$, with equivalence of norms.*

Proof. We will use the index k for both spaces. Since operator $(I-\Delta)^{k/2}$ is a pseudo-differential operator of order k, by the statement on the L^p–continuity in Section 2.2.3 we get that it is bounded from L^p_k to L^p, i.e. we have

$$||(I-\Delta)^{k/2} f||_{L^p} \leq C \sum_{|\alpha| \leq k} ||\partial_x^\alpha f||_{L^p}.$$

Conversely, let P_α be a pseudo-differential operator defined by $P_\alpha = \partial_x^\alpha (I-\Delta)^{-k/2}$, i.e. a pseudo-differential operator with symbol $p_\alpha(x,\xi) = (2\pi i \xi)^\alpha (1+4\pi^2|\xi|^2)^{-k/2}$, independent of x. If $|\alpha| \leq k$, we get that $p_\alpha \in S^{|\alpha|-k} \subset S^0$, so that $P_\alpha \in S^0$ for all $|\alpha| \leq k$. By the statement on the L^p–continuity in Section 2.2.3 operators P_α are bounded on $L^p(\mathbb{R}^n)$. Therefore, we obtain

$$\sum_{|\alpha| \leq k} ||\partial_x^\alpha f||_{L^p} = \sum_{|\alpha| \leq k} ||P_\alpha (I-\Delta)^{k/2} f||_{L^p} \leq C ||(I-\Delta)^{k/2} f||_{L^p},$$

completing the proof.

2.2.5 Proof of the statement on the L^p–continuity

Now, let us prove the statement on the L^p–continuity from Section 2.2.3. However, we will assume without proof that pseudo-differential operators of order zero are bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$. This falls outside the scope of this course and can be proved for example by the so-called Calderon–Zygmund theory of singular integral operators which include pseudo-differential operators from this course, if we view them as integral operators with singular kernels. We will give some indications to this end in the last part of these notes.

Proof of the statement on L^p–continuity. Let $f \in L^p_s(\mathbb{R}^n)$. By definition this means that $(I-\Delta)^{s/2} f \in L^p(\mathbb{R}^n)$. Then we can write using the calculus of pseudo-differential operators:

$$(I-\Delta)^{(s-\mu)/2} T f = (I-\Delta)^{(s-\mu)/2} T (I-\Delta)^{-s/2} (I-\Delta)^{s/2} f \in L^p(\mathbb{R}^n)$$

since operator $(I-\Delta)^{(s-\mu)/2} T (I-\Delta)^{-s/2}$ is a pseudo-differential operator of order zero and is, therefore, bounded on $L^p(\mathbb{R}^n)$.

2.2.6 Calculus proof of L^2–boundedness

In this section we will give a short proof of the fact that pseudo-differential operators of order zero are bounded on $L^2(\mathbb{R}^n)$ which was also proved in Section 2.1.9. The proof will rely on the following lemma that we will give without proof here, and which is sometimes called Schur's lemma.

Lemma. *Let T be an integral operator of the form*

$$Tu(x) = \int_{\mathbb{R}^n} K(x,y)u(y)dy$$

with kernel $K \in L^1_{loc}(\mathbb{R}^n \times \mathbb{R}^n)$ satisfying

$$\sup_{x\in\mathbb{R}^n} \int_{\mathbb{R}^n} |K(x,y)|dy \le C, \qquad \sup_{y\in\mathbb{R}^n} \int_{\mathbb{R}^n} |K(x,y)|dx \le C.$$

Then T is bounded on $L^2(\mathbb{R}^n)$.

Calculus proof of Theorem 2.1.9. Let $T \in \Psi^0$ be an pseudo-differential operator of order zero with symbol $a \in S^0$ and principal symbol σ_a. Then its adjoint satisfies $T^* \in \Psi^0$ and hence also the composition $T^*T \in \Psi^0$. Operator T^*T has bounded principal symbol $|\sigma_a(x,\xi)|^2$ by the composition formula, so that $|\sigma_a(x,\xi)|^2 < M$ for some constant $0 < M < \infty$. Then function

$$p(x,\xi) = \sqrt{M - |\sigma_a(x,\xi)|^2}$$

is well-defined and it is easy to check that $p \in S^0$. Let $P \in \Psi^0$ be the pseudo-differential operator with symbol p. By the calculus again we have the identity

$$T^*T = M - P^*P + R,$$

for some pseudo-differential operator $R \in \Psi^{-1}$. Then

$$||Tu||^2_{L^2} = \langle Tu, Tu\rangle = \langle T^*Tu, u\rangle = M||u||^2_{L^2} - ||Pu||^2_{L^2} + \langle Ru, u\rangle \le M||u||^2_{L^2} + \langle Ru, u\rangle,$$

so that T is bounded on $L^2(\mathbb{R}^n)$ if R is. The boundedness of R on $L^2(\mathbb{R}^n)$ can be proved by induction. Indeed, using the estimate

$$||Ru||^2_{L^2} = \langle Ru, Ru\rangle = \langle R^*Ru, u\rangle \le C||R^*Ru||_{L^2}||u||_{L^2}$$

we see that $R \in \Psi^{-1}$ is bounded on $L^2(\mathbb{R}^n)$ if $R^*R \in \Psi^{-2}$ is bounded on $L^2(\mathbb{R}^n)$. Continuing this argument we can reduce the question of L^2–boundedness to the boundedness of pseudo-differential operators $S \in \Psi^m$ for some sufficiently negative $m < 0$. We can now use Theorem 2.1.6 with $N = 0$ for x close to y to show that the integral kernel $K(x,y)$ of S is bounded for $|y - x| \le 1$ while it is decreases fast for $|x - y|$ large if we take sufficiently large N. Therefore, we can use Lemma 2.2.6 to conclude that S must be bounded on $L^2(\mathbb{R}^n)$ thus completing the proof.

3 Appendix

3.1 Interpolation

3.1.1 Distribution functions

Let μ be the Lebesgue measure on $\mathbb{R}^n$ (a volume element if you are not familiar with the measure theory).

Definition. *For a function $f : \mathbb{R}^n \to \mathbb{C}$ we define its distribution function $\mu_f(\lambda)$ by*

$$\mu_f(\lambda) = \mu\{x \in \mathbb{R}^n : |f(x)| \geq \lambda\}.$$

We have the following useful relation between the L^p–norm and the distribution of a function.

Theorem. *Let $f \in L^p(\mathbb{R}^n)$. Then we have the identity*

$$\int_{\mathbb{R}^n} |f(x)|^p dx = p \int_0^\infty \mu_f(\lambda)\lambda^{p-1} d\lambda.$$

Proof. Let us define a measure on $\mathbb{R}$ by setting

$$\nu((a, b]) = \mu_f(b) - \mu_f(a) = -\mu\{x \in \mathbb{R}^n : a < |f(x)| \leq b\} = -\mu(|f|^{-1}((a, b])).$$

By the standard extension property of measures we can then extend ν to all Borel sets $E \subset (0, \infty)$ by setting $\nu(E) = -\mu(|f|^{-1}(E))$. We note that this definition is well-defined since $|f|$ is measurable if f is measurable. Then we claim that we have the following property for, say, integrable functions $\phi : [0, \infty) \to \mathbb{R}$:

$$\int_{\mathbb{R}^n} \varphi \circ |f| d\mu = -\int_0^\infty \varphi(\alpha) d\mu_f(\alpha). \tag{3.1}$$

Indeed, if $\varphi = \chi_{[a,b]}$ is a characteristic function of a set $[a, b]$, i.e. equal to one on $[a, b]$ and zero on its complement, then the definition of ν implies

$$\int_{\mathbb{R}^n} \chi_{[a,b]} \circ |f| d\mu = \int_{a<|f(x)|\leq b} d\mu = -\int_a^b d\mu_f = -\int_0^\infty \chi_{[a,b]} d\mu_f,$$

which verifies (3.1) for characteristic functions. By the linearity of integrals, we then have (3.1) for finite linear combinations of characteristic functions and, consequently, for all integrable functions by the Lebesgue's monotone convergence theorem. Now, taking $\varphi(\alpha) = \alpha^p$ in (3.1), we get

$$\int_{\mathbb{R}^n} |f|^p d\mu = -\int_0^\infty \alpha^p d\mu_f(\alpha) = p \int_0^\infty \alpha^{p-1} \mu_f(\alpha) d\alpha,$$

where we integrated by parts in the last equality. The proof is complete.

3.1.2 Weak type (p,p)

Definition. *We say that operator T is of weak type (p,p) if there is a constant $C>0$ such that for every $\lambda>0$ we have*

$$\mu\{x\in\mathbb{R}^n : |Tu(x)|>\lambda\} \leq C\frac{||u||_{L^p}^p}{\lambda^p}.$$

Proposition. *If T is bounded from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ then T is also of weak type (p,p).*

Proof. If $v\in L^1(\mathbb{R}^n)$ then for all $\rho>0$ we have a simple estimate

$$\rho\mu\{x\in\mathbb{R}^n : |v(x)|>\rho\} \leq \int_{|v(x)|>\rho}|v(x)|d\mu(x) \leq ||v||_{L^1}.$$

Now, if we take $v(x)=|Tu(x)|^p$ and $\rho=\lambda^p$, this readily implies that T is of weak type (p,p).

3.1.3 Marcinkiewicz interpolation theorem

The following theorem is extremely valuable in proving L^p–continuity of operators since it reduces the analysis to a weaker type continuity only for two values of indices.

Theorem. *Let $r<q$ and assume that operator T is of weak types (r,r) and (q,q). Then T is bounded from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ for all $r<p<q$.*

Proof. Let $u\in L^p(\mathbb{R}^n)$. For each $\lambda>0$ we can define functions u_1 and u_2 by $u_1(x)=u(x)$ for $|u(x)|>\lambda$ and by $u_2(x)=u(x)$ for $|u(x)|\leq\lambda$, and to be zero otherwise. Then we have the identity $u=u_1+u_2$ and estimates $|u_1|,|u_2|\leq|u|$. It follows that

$$\mu_{Tu}(2\lambda)\leq\mu_{Tu_1}(\lambda)+\mu_{Tu_2}(\lambda)\leq C_1\frac{||u_1||_{L^r}^r}{\lambda^r}+C_2\frac{||u_2||_{L^q}^q}{\lambda^q},$$

since T is of weak types (r,r) and (q,q). Therefore, we can estimate

$$\int_{\mathbb{R}^n}|Tu(x)|^pdx = p\int_0^\infty\lambda^{p-1}\mu_{Tu}(\lambda)d\lambda \leq$$
$$C_1p\int_0^\infty\lambda^{p-1-r}\left(\int_{|u|>\lambda}|u(x)|^rdx\right)d\lambda + C_2p\int_0^\infty\lambda^{p-1-q}\left(\int_{|u|\leq\lambda}|u(x)|^qdx\right)d\lambda.$$

Using Fubini's theorem, the first term on the right hand side can be rewritten as

$$\begin{aligned}\int_0^\infty\lambda^{p-1-r}\left(\int_{|u|>\lambda}|u(x)|^rdx\right)d\lambda &= \int_0^\infty\lambda^{p-1-r}\left(\int_{\mathbb{R}^n}\chi_{|u|>\lambda}|u(x)|^rdx\right)d\lambda\\ &= \int_{\mathbb{R}^n}|u(x)|^r\left(\int_0^\infty\lambda^{p-1-r}\chi_{|u|>\lambda}d\lambda\right)dx\\ &= \int_{\mathbb{R}^n}|u(x)|^r\left(\int_0^{|u(x)|}\lambda^{p-1-r}d\lambda\right)dx\\ &= \frac{1}{p-r}\int_{\mathbb{R}^n}|u(x)|^r|u(x)|^{p-r}dx\\ &= \frac{1}{p-r}\int_{\mathbb{R}^n}|u(x)|^pdx,\end{aligned}$$

where $\chi_{|u|>\lambda}$ is the characteristic function of the set $\{x \in \mathbb{R}^n : |u(x)| > \lambda\}$. Similarly, we have

$$\int_0^\infty \lambda^{p-1-q} \left(\int_{|u|\leq\lambda} |u(x)|^q dx \right) d\lambda = \frac{1}{q-p} \int_{\mathbb{R}^n} |u(x)|^p dx,$$

completing the proof.

3.1.4 Calderon–Zygmund covering lemma

As an important tool (which will not be used here so it is given just for the information) for proving various results of boundedness in $L^1(\mathbb{R}^n)$ or of weak type $(1,1)$, we have the following fundamental decomposition of integrable functions.

Theorem. *Let $u \in L^1(\mathbb{R}^n)$ and $\lambda > 0$. Then there exist $v, w_k \in L^1(\mathbb{R}^n)$ and there exists a collection of disjoint cubes Q_k, $k \in \mathbb{N}$, centred at some points x_k, such that the following properties are satisfied*

$$u = v + \sum_{k=1}^\infty w_k, \quad ||v||_{L^1} + \sum_{k=1}^\infty ||w_k||_{L^1} \leq 3||u||_{L^1},$$
$$\operatorname{supp} w_k \subset Q_k, \quad \int_{Q_k} w_k(x)dx = 0,$$
$$\sum_{k=1}^\infty \mu(Q_k) \leq \lambda^{-1}||u||_{L^1}.$$

3.1.5 Remarks on L^p–continuity of pseudo-differential operators

Let $a \in S^0$. Then by Theorem 2.1.6 the integral kernel $K(x,y)$ of pseudo-differential operator T_a satisfies estimates

$$|\partial_x^\alpha \partial_y^\beta K(x,y)| \leq A_{\alpha\beta}|x-y|^{-n-|\alpha|-|\beta|}$$

for all $x \neq y$. In particular, for $\alpha = \beta = 0$ this gives

$$|K(x,y)| \leq A|x-y|^{-n} \text{ for all } x \neq y. \tag{3.2}$$

Moreover, if we use it for $\alpha = 0$ and $|\beta| = 1$, we get

$$\int_{|x-z|\geq 2\delta} |K(x,y) - K(x,z)|dx \leq A \text{ if } |x-z| \leq \delta, \text{ for all } \delta > 0. \tag{3.3}$$

Now, if we take a general integral operator T of the form

$$Tu(x) = \int_{\mathbb{R}^n} K(x,y)u(y)dy,$$

properties (3.2) and (3.3) of the kernel are the starting point of the so-called *Calderon–Zygmund theory of singular integral operators.* In particular, one can conclude that such operators are of weak type $(1,1)$, i.e. they satisfy the estimate

$$\mu\{x \in \mathbb{R}^n : |Tu(x)| > \lambda\} \leq \frac{||u||_{L^1}}{\lambda}.$$

Since we also know from Section 2.1.9 that $T_a \in \Psi^0$ are bounded on $L^2(\mathbb{R}^n)$ and since we also know from Section 3.1.2 that this implies that T_a is of weak type $(2,2)$, we get that pseudo-differential operators of order zero are of weak types $(1,1)$ and $(2,2)$. The, by Marcinkiewicz interpolation theorem, we conclude that T_a is bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < 2$. By the standard duality argument, this implies that T_a is bounded on $L^p(\mathbb{R}^n)$ also for all $2 < p < \infty$. Since we also have the boundedness of $L^2(\mathbb{R}^n)$, we obtain

Theorem. *Let $T \in \Psi^0$. Then T extends to a bounded operator from $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$, for all $1 < p < \infty$.*

We note that there exist different proofs of this theorem. An alternative method is to reduce the L^p–boundedness to the question of unform boundedness of Fourier multipliers in $L^p(\mathbb{R}^n)$ which then follows from Hörmander's theorem on Fourier multipliers.

References

[1] M. Ruzhansky and V. Turunen, *Pseudo-differential operators and symmetries. Background analysis and advanced topics.* Birkhäuser, 2010.

[2] M.A. Shubin, *Pseudodifferential operators and spectral theory*, Nauka, Moscow, 1978.

[3] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton University Press, 1993.