Microlocal Analysis and Boundary Value Problems

Michael Dreher
School of Mathematical and Computer Sciences
Heriot-Watt University Edinburgh

30th April 2016
— Final Version —

Acknowledgements: These are the notes of an intensive course in the Doctoral School *Dissipation and Dispersion in Nonlinear Differential Equations* at the TU Wien and the University of Vienna, whose support is gratefully acknowledged. And special thanks go to the students Anita Gerstenmayer, Florian Huber, Oliver Leingang, Hung Luong for the active contributions to the course.

Contents

1 Motivation **5**

1.1 Differential Equations and their Solutions . . . 5

1.1.1 Boundary Value Problems for Ordinary Differential Equations . . . 5

1.1.2 Problems in Full Space and Half Space . . . 5

1.1.3 Problems in Bounded Domains . . . 8

1.2 Some Key Ideas of Microlocal Analysis . . . 8

1.3 Typical Results. Challenges . . . 10

2 Tools from Other Parts of Analysis **13**

2.1 The Fourier Transform . . . 13

2.1.1 The Schwartz Function Space $\mathcal{S}(\mathbb{R}^n)$ . . . 13

2.1.2 The Fourier Transform on $L^1(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$ . . . 13

2.2 Distribution Theory . . . 16

2.2.1 Purpose . . . 16

2.2.2 Test Functions and Distributions . . . 17

2.2.3 Operations for Distributions . . . 18

2.2.4 Fourier Transformation of Distributions . . . 20

2.2.5 Support of Distributions . . . 22

2.2.6 The Schwartz Kernel Theorem . . . 23

2.2.7 Smoothing Operators . . . 24

2.3 Integral Operators . . . 26

2.3.1 Convolutions . . . 26

2.3.2 Integral Operators on Domains . . . 29

3 Oscillating Integrals **31**

3.1 Definition . . . 31

3.2 Regularity Properties . . . 34

3.3 Fourier Integral Operators . . . 37

4 Symbol Calculus **43**

4.1 Constructing a Symbol from an Amplitude . . . 43

4.2 Transposed Operators, Compositions, and Adjoints . . . 47

4.3 Properly supported operators, asympt. expansions, algebras . . . 50

4.4 Classical Symbols . . . 54

4.5 Elliptic ΨDOs . . . 58

5 Mapping Properties **61**
5.1 The Case of Ψ^m_{cl} on $\mathbb{R}^n$ 61
5.2 The Case of $\Psi^m_{0,0}$ on $\mathbb{R}^n$ 63
5.3 Summary 68
5.3.1 What has been Obtained So Far 68
5.3.2 What does "Micro-Local" Mean ? 68
5.4 Further Ideas 72
5.4.1 How About L^p Based Sobolev Spaces ? Other Spaces ? 72
5.4.2 What are the Classes $S^m_{\varrho,\delta}$ Good For ? 72
5.4.3 What Remains True for FIOs ? 73

6 Boundary Value Problems **75**
6.1 General Principles 75
6.2 Necessity 78
6.3 Sufficiency 81
6.4 Indications on a General Microlocal Theory of BVPs 84

7 Applications **87**
7.1 General Principles 87
7.2 Population Models 90
7.3 Quantum Hydrodynamics 92

A Facts from Topology and Functional Analysis **99**

B Exercises **103**

Chapter 1

Motivation

1.1 Differential Equations and their Solutions

We list some results that can be found in [11], [18] and [25].

1.1.1 Boundary Value Problems for Ordinary Differential Equations

Consider the boundary value problem

$$\begin{cases} u''(x) = f(x), & 0 < x < 1, \\ u(0) = 0, \\ u(1) = 0, \end{cases}$$

for a given function f. Then its solution u can be written [29] as

$$u(x) = \int_{y=0}^{1} K(x,y) f(y) \, \mathrm{d}y,$$

with a kernel function K given by

$$K(x,y) = \begin{cases} y \cdot (x-1) & : 0 \le y \le x \le 1, \\ x \cdot (y-1) & : 0 \le x \le y \le 1. \end{cases}$$

This kernel function K has a singularity for $x = y$, in the sense that there the first derivative K_x has a jump.

1.1.2 Problems in Full Space and Half Space

To simplify notations, we put $D = \frac{1}{\mathrm{i}} \nabla$ with $\mathrm{i}^2 = -1$ and $D_x^\alpha = D_{x_1}^{\alpha_1} \cdot \ldots \cdot D_{x_n}^{\alpha_n}$ for a multi-index $\alpha \in \mathbb{N}^n$ with $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Now let $u = u(x)$ be a function on $\mathbb{R}^n$ that is sufficiently smooth and decays at infinity sufficiently fast (we make this more precise below). Then we can define the value $\hat{u}(\xi)$ of the Fourier transform of u as

$$\hat{u}(\xi) = (\mathcal{F}_{x \to \xi} u)(\xi) = \int_{\mathbb{R}^n_x} e^{-\mathrm{i} x \cdot \xi} u(x) \, \mathrm{d}x.$$

Then we can justify the formula $(D_x^\alpha u)\hat{}(\xi) = \xi^\alpha \hat{u}(\xi)$, with ξ^α being defined similarly as D^α, namely $\xi^\alpha = \xi_1^{\alpha_1} \cdot \ldots \cdot \xi_n^{\alpha_n}$.

Claim: *Let f be a function for which the Fourier transform makes sense as above. Then there is exactly one solution u with the same smoothness and decay properties as f that solves $(1 - \triangle) u = f$ in $\mathbb{R}^n$.*

Justification. If u is such a solution, then we have

$$\hat{f}(\xi) = (1 + |\xi|^2)\hat{u}(\xi), \qquad \forall \xi \in \mathbb{R}^n.$$

This implies uniqueness of u. And indeed, the function

$$u = u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n_\xi} e^{+\mathrm{i}x\cdot\xi} \frac{1}{1+|\xi|^2} \hat{f}(\xi)\,\mathrm{d}\xi$$

will turn out to be a solution with the desired properties. □

From now on, we will always assume that all appearing functions are differentiable as often as we need it, and that they decay at infinity sufficiently fast.

Claim: *There is exactly one solution on* $[0,\infty)\times\mathbb{R}^n$ *to the initial-value problem*

$$\begin{cases} \partial_t u - \triangle u = 0, \quad (t,x) \in (0,\infty)\times\mathbb{R}^n, \\ \quad u(0,x) = u_0(x), \quad x \in \mathbb{R}^n \end{cases}$$

Justification. We perform partial Fourier transform (replacing x by ξ, but keeping t), and this gives us

$$\begin{cases} \partial_t \hat{u} + |\xi|^2 \hat{u} = 0, \quad (t,\xi) \in (0,\infty)\times\mathbb{R}^n, \\ \qquad \hat{u}(\xi) = \hat{u_0}(\xi), \quad \xi \in \mathbb{R}^n, \end{cases}$$

which has the solution $\hat{u}(t,\xi) = \exp(-t|\xi|^2)\hat{u_0}(\xi)$. Now we invert the partial Fourier transform and get

$$u(t,x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\xi - t|\xi|^2} \hat{u_0}(\xi)\,đ\xi, \qquad đ\xi := \frac{\mathrm{d}\xi}{(2\pi)^n}.$$

Let us substitute $\hat{u}_0$ here:

$$u(t,x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\xi - t|\xi|^2} \left(\int_{\mathbb{R}^n_y} e^{-\mathrm{i}\xi\cdot y} u_0(y)\,\mathrm{d}y \right) đ\xi,$$

and for positive t we are lucky and can swap the integrals:

$$u(t,x) = \int_{\mathbb{R}^n_y} u_0(y) \left(\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}\xi\cdot(x-y) - t|\xi|^2}\, đ\xi \right) \mathrm{d}y.$$

The inner integral can be evaluated, and the result then is

$$u(t,x) = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n_y} \exp\left(-\frac{|x-y|^2}{4t} \right) u_0(y)\,\mathrm{d}y. \tag{1.1}$$

□

Let us now have a look at the problem

$$-\triangle u(x) = f(x), \qquad x \in \mathbb{R}^n.$$

We perform the Fourier transform as before and get $|\xi|^2\hat{u}(\xi) = \hat{f}(\xi)$, from which then the inverse Fourier transform gives us ("no one has the intention of erecting a pole")

$$u(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} \frac{1}{|\xi|^2} \hat{f}(\xi)\,đ\xi = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} \frac{1}{|\xi|^2} \left(\int_{\mathbb{R}^n_y} e^{-\mathrm{i}y\cdot\xi} f(y)\,\mathrm{d}y \right) đ\xi.$$

And now ("can you see the three magpies outside the window ?") we have

$$u(x) = \int_{\mathbb{R}^n_y} \left(\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi} \frac{1}{|\xi|^2}\,đ\xi \right) f(y)\,\mathrm{d}y.$$

Continuing in this fashion we then get the formula

$$u(x) = \begin{cases} -\frac{1}{2\pi}\int_{\mathbb{R}^2_y} \ln(|x-y|)f(y)\,\mathrm{d}y & : n = 2, \\ \frac{1}{n(n-2)\alpha(n)}\int_{\mathbb{R}^n_y} \frac{1}{|x-y|^{n-2}} f(y)\,\mathrm{d}y & : n > 2, \end{cases} \tag{1.2}$$

with $\alpha(n)$ being the volume of the n-dimensional ball of radius 1.

Claim: *There is exactly one solution u on $\mathbb{R}\times\mathbb{R}^n$ to the initial-value problem*

$$\begin{cases} (\partial_t^2 - \triangle)u = 0, & (t,x) \in \mathbb{R}\times\mathbb{R}^n, \\ u(0,x) = u_0(x), & x \in \mathbb{R}^n, \\ u_t(0,x) = u_1(x), & x \in \mathbb{R}^n. \end{cases}$$

Justification. We perform partial Fourier transform as in the parabolic case and get

$$\left(\partial_t^2 + |\xi|^2\right)\hat{u}(t,\xi) = 0,$$

which implies together with the initial conditions that

$$\hat{u}(t,\xi) = \cos(t|\xi|)\hat{u_0}(\xi) + \frac{\sin(t|\xi|)}{|\xi|}\hat{u_1}(\xi), \qquad (t,\xi) \in \mathbb{R}\times\mathbb{R}^n.$$

This formula can be re-arranged into

$$\begin{aligned} u(t,x) = {} & \frac{1}{2}\int_{\mathbb{R}^n} e^{\mathrm{i}(x\cdot\xi + t|\xi|)}\hat{u}_0(\xi)\,đ\xi + \frac{1}{2}\int_{\mathbb{R}^n} e^{\mathrm{i}(x\cdot\xi - t|\xi|)}\hat{u}_0(\xi)\,đ\xi \\ & + \frac{1}{2\mathrm{i}}\int_{\mathbb{R}^n} e^{\mathrm{i}(x\cdot\xi + t|\xi|)}\frac{1}{|\xi|}\hat{u}_1(\xi)\,đ\xi - \frac{1}{2\mathrm{i}}\int_{\mathbb{R}^n} e^{\mathrm{i}(x\cdot\xi - t|\xi|)}\frac{1}{|\xi|}\hat{u}_1(\xi)\,đ\xi. \end{aligned} \tag{1.3}$$

As before, we can now express $\hat{u}_0(\xi)$ and $\hat{u}_1(\xi)$ as $\int_{\mathbb{R}^n_y} \exp(-\mathrm{i}y\cdot\xi)u_{0,1}(y)\,\mathrm{d}y$, and (as dodgy as above) swap the two integrals over y and ξ, and then we get the following formulas after some calculation:

$$u(t,x) = \frac{u_0(x-t) + u_0(x+t)}{2} + \frac{1}{2}\int_{y=x-t}^{x+t} u_1(y)\,\mathrm{d}y, \qquad t \geq 0, \quad n = 1, \tag{1.4}$$

$$u(t,x) = \frac{1}{2\int_{B_t(x)} 1\,\mathrm{d}y}\int_{B_t(x)} \frac{tu_0(y) + t^2 u_1(y) + t\nabla u_0(y)\cdot(y-x)}{\sqrt{t^2 - |x-y|^2}}\,\mathrm{d}y, \qquad t > 0, \quad n = 2, \tag{1.5}$$

$$u(t,x) = \frac{1}{\int_{\partial B_t(x)} 1\,\mathrm{d}\sigma(y)}\int_{\partial B_t(x)} \Big(tu_1(y) + u_0(y) + \nabla u_0(y)\cdot(y-x)\Big)\,\mathrm{d}\sigma(y), \qquad t > 0, \quad n = 3. \tag{1.6}$$

□

We observe that we can can bring the solution formulas (1.1), (1.2), (1.4)–(1.6) into convolution form, which means to express the solution u as something like $u(x) = (K * u_0)(x) := \int_{\mathbb{R}^n} K(x-y)u_0(y)\,\mathrm{d}y$.

For the solution formula (1.1) of the heat equation we get

$$u(t,x) = (K_t * u_0)(x), \qquad K_t(z) = \frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{|z|^2}{4t}\right).$$

This integral kernel K_t has no pole (assuming $t > 0$).

For the solution formula (1.2), we have

$$u(x) = (K * f)(x), \qquad K(z) = \begin{cases} -\frac{1}{2\pi}\ln|z| & : n = 2, \\ \frac{1}{n(n-2)\alpha(n)}\cdot\frac{1}{|z|^{n-2}} & : n > 2. \end{cases}$$

Now the integral kernel K has singularities at $z = 0$, which correspond to singularities at $y = x$ in the convolution integral $(K * f)(x)$.

And for the wave equations, things are a bit more complicated because we have two initial functions. First we define a function $K_t(z)$ of $z \in \mathbb{R}^n$ with parameter $t > 0$ by its Fourier transform like this:

$$\widehat{K_t}(\xi) := \text{const.}\frac{\sin(t|\xi|)}{|\xi|},$$

and then the solution becomes

$$u(t,x) = \partial_t\Big(K_t * u_0\Big)(x) + \Big(K_t * u_1\Big)(x).$$

The kernel K_t now turns out to be the following distributions or functions:

$$K_t(z) = \begin{cases} \frac{1}{2}\chi_{[-t,t]}(z) & : n = 1, \\ \frac{1}{4\pi t}\delta(|z| - t) & : n = 3, \end{cases}$$

and for $n = 2$, we have

$$K_t(z) = \begin{cases} \frac{1}{2\pi\sqrt{t^2-|z|^2}} & |z| < t, \\ 0 & |z| \geq t. \end{cases}$$

We observe that these three kernel functions have singularities at z whenever $|z| = t$, which is to be expected, because singularities of the initial data u_0 and u_1 should be propagated with velocity $c = 1$.

1.1.3 Problems in Bounded Domains

Proposition 1.1 (**Representation Formula of Poisson**[1]). *Let $B = B_R(0)$ be the open ball of radius R about 0 in $\mathbb{R}^n$ with $n > 2$, and let u be the solution to*

$$\begin{cases} \triangle u(x) = 0, & x \in B, \\ \quad u(x) = g(x), & x \in \partial B, \end{cases}$$

where $g \in C(\partial B)$. Then u has the representation

$$u(x) = \frac{R^2 - |x|^2}{n\alpha(n)R} \int_{y\in\partial B} \frac{g(y)}{|x-y|^n} \,\mathrm{d}\sigma_y, \qquad x \in B,$$

with $\alpha(n)$ being the volume of the n-dimensional ball of radius 1.

Proof. See [16]. □

1.2 Some Key Ideas of Microlocal Analysis

In this course, we are concerned with boundary value problems (or initial value problems) of the form

$$\begin{cases} \mathcal{A}(x, D_x)u(x) = f(x), & x \in \Omega, \\ \mathcal{B}(x, D_x)u(x) = g(x), & x \in \partial\Omega, \end{cases}$$

with given f and g. Here $\mathcal{A}$ and $\mathcal{B}$ are certain partial differential operators (PDOs).

We also are interesting in finding (at least abstract) solution formulas that look approximately like this:

$$u(x) = \int_\Omega K_f(x,y)f(y)\,\mathrm{d}y + \int_{\partial\Omega} K_g(x,y)g(y)\,\mathrm{d}y,$$

for some (not yet determined, or perhaps never determined) kernel functions K_f and K_g.

It turns out that all these operators (the PDOs and the solution operators) can be subsumed as Pseudodifferential Operators (ΨDOs) or Fourier–Integral Operators (FIOs), and the theory that studies all these operators is called *Microlocal Analysis.*

What are Pseudodifferential Operators ? They are generalisations of PDOs. Let $\mathcal{P} = \mathcal{P}(x, D_x)$ be such an operator on $\mathbb{R}^n$, hence

$$\mathcal{P}(x, D_x) = \sum_{|\alpha|\leq m} a_\alpha(x) D_x^\alpha,$$

[1] Siméon Denis Poisson, 1781–1840

where $\mathcal{P}$ has order $m \in \mathbb{N}$, and the coefficients a_α are smooth functions on $\mathbb{R}^n$, which are bounded together with all their derivatives, hence $a_\alpha \in C_b^\infty(\mathbb{R}^n)$.

We recall $D_x = \frac{1}{\mathrm{i}}\nabla$ (the whole point of this extra factor $\frac{1}{\mathrm{i}}$ is to simplify various formulas in which Fourier transforms appear). Using $(D_x^\alpha u)\hat{}(\xi) = \xi^\alpha \hat{u}(\xi)$, we then can write

$$\begin{aligned}(\mathcal{P}(x, D_x)u)(x) &= \sum_{|\alpha|\le m} a_\alpha(x)(\mathcal{F}^{-1}_{\xi\to x}\xi^\alpha\hat{u}(\xi))(x)\\ &= \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\sum_{|\alpha|\le m} a_\alpha(x)\xi^\alpha\hat{u}(\xi)\frac{\mathrm{d}\xi}{(2\pi)^n}.\end{aligned}$$

Again we set $\text{đ}\xi := (2\pi)^{-n}\,\mathrm{d}\xi$ (in order to simplify the notation to follow in later chapters), as well as $p(x,\xi) := \sum_{|\alpha|\le m} a_\alpha(x)\xi^\alpha$, and we obtain the compact notation

$$(\mathcal{P}(x, D_x)u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}p(x,\xi)\hat{u}(\xi)\,\text{đ}\xi. \tag{1.7}$$

The function p is called *pseudodifferential symbol* of the operator $\mathcal{P}$, and sometimes we will write $p = \sigma(\mathcal{P})$. If $\mathcal{P}$ is a PDO, then p is a polynomial in ξ.

The following properties of p are quite obvious:

- for $|\xi| \to \infty$, the symbol $p = p(x,\xi)$ has polynomial growth, and the growth order equals the order of the operator,
- if we take derivatives of p with respect to ξ, then the growth order is being reduced,
- if we take derivatives of p with respect to x, then the growth order will typically not change.

The key idea of ΨDOs is to replace the polynomial symbol p by another function, which is not necessarily a polynomial in ξ, but it still satisfies the above three properties. The associated operator P is then called a ΨDO.

Example: *The solution operator on the full space to* $1-\triangle$ *is a ΨDO with the symbol* $\sigma((1-\triangle)^{-1}) = \frac{1}{1+|\xi|^2}$, *which has growth order* -2.

Now some ideas have been presented, and some questions arise:

- Which of the properties of a PDO continue to hold for a ΨDO ? (mapping properties between Sobolev spaces; adjoint operators and composed operators are again operators of the same type; spectral properties)
- What are the advantages ?
- How to define ΨDO reasonable in $\Omega \subset \mathbb{R}^n$ (a difficulty is how to define $\hat{u}$ for some function u that is given only in a part of $\mathbb{R}^n$, not everywhere)
- How to generalise ΨDO reasonably ? (this will bring us to Fourier–Integral operators (FIOs), which are solution operators to hyperbolic partial differential equations).

We begin to generalise our notation (1.7) of ΨDOs. First we note that we have the equivalent formula

$$(\mathcal{P}(x, D_x)u)(x) = \int_{\mathbb{R}^n_\xi}\left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi}p(x,\xi)u(y)\,\mathrm{d}y\right)\text{đ}\xi,$$

in the sense of an *iterated integral*. Next we remark that the formula $(\mathcal{P}u)(x) = \sum_{|\alpha|\le m} a_\alpha(x)D_x^\alpha u(x)$ means that we take first the derivative D_x^α of u, and afterwards we multiply by the coefficient $a_\alpha(x)$. This order is quite arbitrary. Indeed, supposing $x \in \mathbb{R}^1$ for simplicity, the operator $a(x)\partial_x^2$ can be re-written like this:

$$a(x)\partial_x^2 u(x) = \partial_x^2(a(x)u(x)) - 2a'(x)u'(x) - a''(x)u(x) = \partial_x^2\Big(a(x)u(x)\Big) - 2\partial_x\Big(a'(x)u(x)\Big) + a''(x)u(x),$$

and the same re-writing is applicable to any operator $\sum_{|\alpha|\le m} a_\alpha(x)D_x^\alpha$. Tedious, but doable.

But then an operator $P(x, D_x)$ with action as $(\mathcal{P}u)(x) = \sum_{|\alpha|\le m} D_x^\alpha(a_\alpha(x)u(x))$ will lead to the formula

$$(\mathcal{P}(x, D_x)u)(x) = \int_{\mathbb{R}^n_\xi} \left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} p(y,\xi)u(y)\,\mathrm{d}y \right) \text{đ}\xi, \qquad p(y,\xi) := \sum_{|\alpha|\le m} a_\alpha(y)\xi^\alpha,$$

and now the factor $p(x,\xi)$ has turned into $p(y,\xi)$. Observe that operators with this convention appear naturally when we build adjoint differential operators.

Therefore, we arrive at a first generalisation of (1.7):

$$(\mathcal{P}(x, D_x)u)(x) = \int_{\mathbb{R}^n_\xi} \left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)u(y)\,\mathrm{d}y \right) \text{đ}\xi, \tag{1.8}$$

and the function $a = a(x,y,\xi)$ has polynomial growth with respect to ξ. We call this function *amplitude function* of the operator $\mathcal{P}$, and we remark right away that each operator $\mathcal{P}$ will have many different amplitude functions.

Now we have the desire to swap the integrals in (1.8) and obtain

$$(\mathcal{P}(x, D_x)u)(x) \overset{?}{=} \int_{\mathbb{R}^n_y} K(x,y)u(y)\,\mathrm{d}y, \qquad K(x,y) \overset{?}{=} \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)\,\text{đ}\xi,$$

and this step needs a meaning and a justification. Note that we can write K also as an inverse Fourier transform of the amplitude function

$$K(x,y) \overset{?}{=} \left(\mathcal{F}^{-1}_{\xi\to z} a(x,y,\xi) \right)_{|z=x-y},$$

provided we can explain what this actually means: "(inverse) Fourier transform of a function that is a polynomial (or grows like a polynomial)". The justification that will be given in later chapters crucially rests on the oscillations of the factor $\exp(\mathrm{i}(x-y)\cdot\xi)$ which we observe if $|\xi|$ approaches ∞, provided that $x \neq y$.

The object $K(x,y)$ that will be produced by this inverse Fourier transform procedure then will not be a function, but a distribution (called the *Schwartz kernel of the operator* K), and it is advantageous to know where this distribution K is *singular*. The examples which you have seen so far indicate that singularities are to be expected for $x = y$.

On the other hand, these discussions seem to exclude the solution formula (1.3) of the wave equation. And indeed, (1.3) can not be brought into the form (1.7), because "the exponential term does not fit".

In order to handle (1.3), we need to generalise (1.8) to

$$(\mathcal{P}(x, D_x)u)(x) = \int_{\mathbb{R}^n_\xi} \left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}\Phi(x,y,\xi)} a(x,y,\xi)u(y)\,\mathrm{d}y \right) \text{đ}\xi, \tag{1.9}$$

and the function Φ is called *phase function.* For $\Phi(x,y,\xi) = (x-y)\cdot\xi$, we arrive again at (1.8). And $\Phi(x,y,\xi) = (x-y)\cdot\xi \pm t|\xi|$ brings us to (1.3). Operators of this type are called *Fourier Integral Operators (FIO).* Again, oscillations will help us in ensuring that an integral which gives us the Schwartz kernel actually converges, and therefore we should require that Φ is real-valued, and that certain derivatives of Φ do not vanish. If $\nabla_\xi\Phi(x,y,\xi) = 0$ for certain (x,y,ξ), then the oscillation is "not fast enough there", and some singularity in the Schwartz kernel appears at that point (x,y). If $\Phi(x,y,\xi) = (x-y)\cdot\xi + t|\xi|$, then we have $\nabla_\xi\Phi = 0$ exactly if $|x-y| = |t|$, which simply means that the wave equation lets singularities of the initial data propagate with speed 1, as it should be.

1.3 Typical Results. Challenges

We expect the following result: if $\Omega \subset \mathbb{R}^n$ is a domain (suppose its boundary $\partial\Omega$ is smooth enough, whatever that means), and $\mathcal{P}$ is a ΨDO of growth order m, then $\mathcal{P}$ should map the Sobolev space $H^s(\Omega)$ into $H^{s-m}(\Omega)$. This seems plausible, but in reality things are more complicated.

The reason is this: in order to calculate $\mathcal{P}u$ as in (1.7), we need the Fourier transform $\hat{u}$ of u, hence we need to know u outside of Ω. How to define u outside of Ω ? We could extend u with zero-values. But then the extended version u_{ext} of u will have a jump at $\partial\Omega$, and u_{ext} will not belong to $H^s(\mathbb{R}^n)$, and hence $\mathcal{P}u_{\text{ext}} \notin H^{s-m}(\mathbb{R}^n)$. And we do not know whether the operator $\mathcal{P}$ "drags the jump of u_{ext} into Ω".

A loophole for us could be the following: we demand that $u \in H^s(\Omega)$ has compact support inside Ω, which means that $u \equiv 0$ in an inside marginal strip along $\partial\Omega$, and then we can safely define

$$u_{\text{ext}}(x) = \begin{cases} u(x) & : x \in \Omega, \\ 0 & : x \notin \Omega, \end{cases}$$

without introducing artificial singularities. Then $\mathcal{P}u_{\text{ext}} \in H^{s-m}(\mathbb{R}^n)$ seems plausible (and is true, actually). If we restrict $\mathcal{P}u_{\text{ext}}$ back to Ω, we get a function of $H^{s-m}(\Omega)$. The result then is the mapping property $\mathcal{P}\colon H^s_{\text{comp}}(\Omega) \to H^{s-m}(\Omega)$, and the subscript "comp" is crucial here.

But the next challenge is waiting for us: ΨDOs do not preserve the support of a function. This means the following. If $\mathcal{P}(x, D_x) = \sum_{|\alpha|\le m} a_\alpha(x) D_x^\alpha$ is a PDO, then it does not increase the support, hence $\operatorname{supp} \mathcal{P}u \subseteq \operatorname{supp} u$.

But this is no longer true if $\mathcal{P}$ is a ΨDO. The reason is that the Schwartz kernel $K(x, y)$ of a PDO is zero for $x \ne y$, hence $\operatorname{supp} K \subseteq \operatorname{diag}(\Omega \times \Omega)$. But ΨDOs do not behave that way — their Schwartz kernels are typically everywhere non-zero, which means $\operatorname{supp} K = \overline{\Omega \times \Omega}$. The consequence then is that we can not define what the composition $\mathcal{P} \circ \mathcal{Q}$ of two ΨDOs $\mathcal{P}$ and $\mathcal{Q}$ is supposed to mean if both are to act on a domain $\Omega \ne \mathbb{R}^n$. A loophole could be to forget about domains Ω, and only ever consider the full space $\mathbb{R}^n$, which has no boundary. Or to only ever consider the torus $\mathbb{T}^n$ as domain for the x variable, which also has no boundary.

There is one more challenge waiting for us, which is a bit harder to see. Consider two PDOs $\mathcal{P}$ and $\mathcal{Q}$, having the form

$$\mathcal{P}(x, D_x) = \sum_{|\alpha|\le m} a_\alpha(x) D_x^\alpha, \qquad \mathcal{Q}(x, D_x) = \sum_{|\beta|\le m} b_\beta(x) D_x^\beta.$$

Then their composition $\mathcal{P} \circ \mathcal{Q}$ is again a PDO, which we compute as follows:

$$\begin{aligned} ((\mathcal{P} \circ \mathcal{Q})u)(x) &= \sum_{|\alpha|\le m, |\beta|\le m} a_\alpha(x) D_x^\alpha \Big(b_\beta(x) D_x^\beta u(x) \Big) \\ &= \sum_{|\alpha|\le m, |\beta|\le m} a_\alpha(x) \sum_{\gamma\le\alpha} \binom{\alpha}{\gamma} \Big(D_x^\gamma b_\beta(x) \Big) D_x^{\beta+\alpha-\gamma} u(x), \end{aligned}$$

and therefore the pseudodifferential symbol of $\mathcal{P} \circ \mathcal{Q}$ evaluates as

$$\begin{aligned} \sigma(\mathcal{P} \circ \mathcal{Q})(x, \xi) &= \sum_{|\alpha|\le m, |\beta|\le m} \sum_{\gamma\le\alpha} \binom{\alpha}{\gamma} a_\alpha(x) \Big(D_x^\gamma b_\beta(x) \Big) \xi^{\alpha+\beta-\gamma} \\ &= \sum_{|\gamma|=0}^{\infty} \left(\sum_{|\alpha|\le m, \alpha\ge\gamma} \binom{\alpha}{\gamma} a_\alpha(x) \xi^{\alpha-\gamma} \right) \cdot \left(\sum_{|\beta|\le m} D_x^\gamma b_\beta(x) \xi^\beta \right) \\ &= \sum_{|\gamma|=0}^{\infty} \left(\sum_{|\alpha|\le m} \frac{1}{\gamma!} a_\alpha(x) \partial_\xi^\gamma \xi^\alpha \right) \cdot \left(\sum_{|\beta|\le m} D_x^\gamma b_\beta(x) \xi^\beta \right) \\ &= \sum_{|\gamma|=0}^{\infty} \frac{1}{\gamma!} \Big(\partial_\xi^\gamma p(x, \xi) \Big) \cdot \Big(D_x^\gamma q(x, \xi) \Big), \qquad p(x, \xi) := \sigma(\mathcal{P}), \qquad q(x, \xi) := \sigma(\mathcal{Q}). \end{aligned} \tag{1.10}$$

Note the following: although we have let $|\gamma|$ run to ∞, in reality only a finite number of terms appear in this series, because p is a polynomial in ξ of degree at most m, and therefore $\partial_\xi^\gamma p(x, \xi)$ will be zero whenever $|\gamma| > m$.

Now let us replace $\mathcal{P}$ and $\mathcal{Q}$ by ΨDOs, and the situation changes. Suddenly we have to add an infinite number of terms. It gets worse — typically this series does not converge anymore. We have to replace the concept of a *converging series* by the concept of an *asymptotic series*.

This means the following. Consider a series $s(\xi) = \sum_{\ell=0}^{\infty} a_\ell(\xi)$. Define its partial sum

$$s_L(\xi) := \sum_{\ell=0}^{L} a_\ell(\xi).$$

We say that $\sum_{\ell=0}^{\infty} a_\ell(\xi)$ is a *converging series with value* $s(\xi)$ if the following holds:

$$\forall\, \xi \text{ (fixed)}: \quad |s(\xi) - s_L(\xi)| \to 0 \quad \text{for} \quad L \to \infty.$$

To define an asymptotic series, we suppose that each item $a_\ell(\xi)$ has growth order $m - \ell$ for $|\xi| \to \infty$, for some common $m \in \mathbb{R}$.
We say that $\sum_{\ell=0}^{\infty} a_\ell(\xi)$ is an *asymptotic series with value* $s(\xi)$ if the following holds:

$$\forall\, L \text{ (fixed)}: \quad |s(\xi) - s_L(\xi)| \to 0 \quad \text{with speed} \quad \mathcal{O}(|\xi|^{m-L-1}) \quad \text{for} \quad |\xi| \to \infty.$$

Only large L are interesting. This is the only sense of convergence we will have available. To make matters worse, the limit $s(\xi)$ is never unique, because you can add some item like $\exp(-73|\xi|^2)$ to $s(\xi)$.

After all these remarks, we aspire to achieve the following results:

- ΨDOs and FIOs of order $m \in \mathbb{R}$ generate maps from $H^t(\Omega)$ into $H^{t-m}(\Omega)$, for $t \in \mathbb{R}$, possibly under some extra condition to manage the supports.
- the composition of two ΨDOs is again a ΨDO (under some extra condition to manage the support). And we have an asymptotic series for the symbol of the composition.
- the adjoint operator of a ΨDO is again a ΨDO (with asymptotic series for its symbol).
- elliptic operators are "almost invertible". That means that for each elliptic operator $\mathcal{P}$ on the full space $\mathbb{R}^n$, we have another operator $\mathcal{Q}$ (which is again elliptic) such that $\mathcal{P} \circ \mathcal{Q} = \mathrm{id} - \mathcal{R}$, and $\mathcal{R}$ is a smoothing operator (its Schwartz kernel is a C^∞ function). We have an asymptotic formula for $\mathcal{Q}$.
- the same holds for elliptic boundary value problems on a bounded domain Ω, provided we can define what an elliptic BVP is. The proof will be much more work.

We will focus on L^2 based Sobolev spaces, and will give only a few remarks concerning L^p for $1 < p < \infty$. Things would get nasty for $p = 1$ and $p = \infty$.

About the literature: the course follows Chapter 1 of Shubin[2] [22]. The standard reference is Hörmander[3] [14], which is a monumental work full of advanced results.

[2]Mikhail Shubin, 1944–
[3]Lars Valter Hörmander, 1931–2012, Fields medallist 1962

Chapter 2

Tools from Other Parts of Analysis

2.1 The Fourier Transform

2.1.1 The Schwartz Function Space $\mathcal{S}(\mathbb{R}^n)$

Definition 2.1 (Schwartz function space $\mathcal{S}(\mathbb{R}^n)$). *The* SCHWARTZ[1] space $\mathcal{S}(\mathbb{R}^n)$ *consists of all those functions* $f \in C^\infty(\mathbb{R}^n)$ *with*

$$p_{k,\alpha}(f) := \sup_{x \in \mathbb{R}^n} \left(1 + |x|^k\right) |\partial_x^\alpha f(x)| < \infty,$$

for all $k \in \mathbb{N}_0$ *and all* $\alpha \in \mathbb{N}^n$.

These Schwartz functions are infinitely smooth, they decay at infinity faster than all powers of $|x|^{-1}$, and all their derivatives decay at infinity faster than all powers of $|x|^{-1}$, too.

We remark that $\mathcal{S}(\mathbb{R}^n)$ is a complete locally convex linear topological vector space, whose topology is being generated by a countable collection of semi-norms $p_{k,\alpha}$; and therefore it is also a Fréchet [2] space. This is useful because for each Fréchet space there is a metric on that space that generates the same topology. Such *metrisable* spaces are so much nicer to handle than general linear topological vector spaces. We do not explicitely write down the topology (understood as a collection of open sets) of $\mathcal{S}(\mathbb{R}^n)$, because this is quite easy.

Definition 2.2 (Convergence in $\mathcal{S}(\mathbb{R}^n)$). *We say that a sequence* $(\varphi_1, \varphi_2, \dots) \subset \mathcal{S}(\mathbb{R}^n)$ *converges to* $\varphi \in \mathcal{S}$ in the topology of $\mathcal{S}(\mathbb{R}^n)$ *if* $\lim_{j\to\infty} p_{k,\alpha}(\varphi_j - \varphi) = 0$ *for all* k, α. *We write*

$$\varphi_j \xrightarrow{\mathcal{S}} \varphi \qquad (j \to \infty)$$

for this convergence.

This means that the sequence $(\varphi_1, \varphi_2, \dots)$ converges to φ uniformly, and all the sequences of derivatives enjoy uniform convergence, too. The convergence in the topology of $\mathcal{S}$ is extremely strong and powerful.

2.1.2 The Fourier Transform on $L^1(\mathbb{R}^n)$, $\mathcal{S}(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$

Definition 2.3 (Fourier[3] transform on $L^1(\mathbb{R}^n)$). *For* $f \in L^1(\mathbb{R}^n)$, *we define its Fourier transform* $\mathcal{F}f = \hat{f}$ *by*

$$(\mathcal{F}f)(\xi) := \int_{\mathbb{R}^n} e^{-\mathrm{i}x\cdot\xi} f(x)\,\mathrm{d}x, \qquad \xi \in \mathbb{R}^n, \quad x \cdot \xi := x_1\xi_1 + \dots + x_n\xi_n.$$

[1] LAURENT SCHWARTZ, 1915–2002, inventor of the distributions (after Sobolev), Fields medallist 1950
[2] MAURICE RENÉ FRÉCHET, 1878–1973
[3] JEAN BAPTISTE JOSEPH FOURIER, 1768–1830

We introduce the notations

$$D := \frac{1}{\mathrm{i}}\nabla, \qquad đ\xi := \frac{\mathrm{d}\xi}{(2\pi)^n}.$$

Proposition 2.4. *The Fourier transform has the following properties:*

$$\begin{aligned}
f \in L^1(\mathbb{R}^n) &\implies |\hat f(\xi)| \le \|f\|_{L^1(\mathbb{R}^n)} \quad \forall\, \xi \in \mathbb{R}^n,\\
f \in \mathcal{S}(\mathbb{R}^n) &\implies \hat f \in \mathcal{S}(\mathbb{R}^n_\xi),\\
f_j \xrightarrow{\mathcal{S}} f &\implies \hat f_j \xrightarrow{\mathcal{S}(\mathbb{R}^n_\xi)} \hat f,\\
f \in \mathcal{S}(\mathbb{R}^n) &\implies (\mathcal{F}(D_x^\alpha f))(\xi) = \xi^\alpha (\mathcal{F}f)(\xi), \qquad \forall\, \alpha \in \mathbb{N}^n,\\
f, g \in \mathcal{S}(\mathbb{R}^n) &\implies \int_{\mathbb{R}^n_x} f(x)\overline{g(x)}\,\mathrm{d}x = \int_{\mathbb{R}^n_\xi} \hat f(\xi)\overline{\hat g(\xi)}\,đ\xi,\\
f, g \in \mathcal{S}(\mathbb{R}^n) &\implies (f\cdot g)\widehat{\ }(\xi) = (\hat f * \hat g)(\xi),
\end{aligned}$$

with $*$ *as the convolution operator:*

$$(\hat f * \hat g)(\xi) := \int_{\mathbb{R}^n_\eta} \hat f(\eta)\cdot \hat g(\xi-\eta)\,đ\eta. \tag{2.1}$$

The penultimate property yields the Parseval[4] *identity*:

$$\|f\|_{L^2(\mathbb{R}^n_x)} = \left\|\hat f\right\|_{L^2(\mathbb{R}^n_\xi)} := \left(\int_{\mathbb{R}^n_\xi} |\hat f(\xi)|^2\,đ\xi\right)^{1/2}, \qquad \forall\, f \in \mathcal{S}(\mathbb{R}^n).$$

Example 2.5. *Take* $f = f(x) = \exp(-x^2/2)$ *on* $\mathbb{R}^1$*. Then*

$$\begin{aligned}
\partial_\xi \hat f(\xi) &= \partial_\xi \int_{\mathbb{R}_x} e^{-\mathrm{i}x\cdot\xi} f(x)\,\mathrm{d}x = \int_{\mathbb{R}_x} (-\mathrm{i}x)e^{-\mathrm{i}x\cdot\xi} f(x)\,\mathrm{d}x = \mathrm{i}\int_{\mathbb{R}_x} e^{-\mathrm{i}x\cdot\xi}\partial_x f(x)\,\mathrm{d}x\\
&= -\int_{\mathbb{R}_x} e^{-\mathrm{i}x\cdot\xi}(D_x f)(x)\,\mathrm{d}x = -(D_x f)\widehat{\ }(\xi) = -\xi \hat f(\xi),
\end{aligned}$$

and this ODE has the solutions

$$\hat f(\xi) = c\exp(-\xi^2/2),$$

with an unknown constant c *which can be found by*

$$c = \hat f(0) = \int_{\mathbb{R}_x} f(x)\,\mathrm{d}x = \int_{\mathbb{R}_x} e^{-x^2/2}\,\mathrm{d}x = (2\pi)^{1/2}.$$

Example 2.6. *Take* $f = f(x) = \exp(-|x|^2/2)$ *on* $\mathbb{R}^n$*. Then*

$$\begin{aligned}
\hat f(\xi) &= \int_{\mathbb{R}^n_x} e^{-\mathrm{i}(x_1\xi_1+\dots+x_n\xi_n)} f(x)\,\mathrm{d}x = \prod_{k=1}^n \left(\int_{\mathbb{R}^1_t} e^{-\mathrm{i}t\xi_k}\exp(-t^2/2)\,\mathrm{d}t\right)\\
&= \prod_{k=1}^n (2\pi)^{1/2}\exp(-\xi_k^2/2) = (2\pi)^{n/2}\exp(-|\xi|^2/2).
\end{aligned}$$

By repeated substitution we find that $f_\varepsilon(x) = \exp(-|\varepsilon x|^2/2)$ *has the Fourier transform* $\hat f_\varepsilon(\xi) = \varepsilon^{-n}(2\pi)^{n/2}\exp(-|\xi/\varepsilon|^2/2)$*. This function* $\hat f_\varepsilon$ *has a peak at* $\xi = 0$*, and we know that* $\int_{\mathbb{R}^n_\xi} \hat f_\varepsilon(\xi)\,\mathrm{d}\xi = (2\pi)^n$*.*

Proposition 2.7 (Inverse Fourier transform on $\mathcal{S}(\mathbb{R}^n)$). *The Fourier transform is an isomorphism from* $\mathcal{S}(\mathbb{R}^n_x)$ *onto* $\mathcal{S}(\mathbb{R}^n_\xi)$*, and the inverse Fourier transform is given by*

$$\varphi(x) = \int_{\mathbb{R}^n_\xi} e^{+\mathrm{i}x\cdot\xi}\hat\varphi(\xi)\,đ\xi, \qquad x \in \mathbb{R}^n, \quad \hat\varphi \in \mathcal{S}(\mathbb{R}^n). \tag{2.2}$$

[4] Marc–Antoine Parseval des Chênes, 1755–1836

Proof. We wish to show that

$$(2\pi)^n\varphi(x) = \int_{\mathbb{R}^n_\xi}\left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi}\varphi(y)\,\mathrm{d}y\right)\mathrm{d}\xi,$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and all $x \in \mathbb{R}^n$. However, the integral on the RHS does not converge absolutely, so we cannot swap the integrals.

Pick some $\psi \in \mathcal{S}(\mathbb{R}^n)$. Then we get

$$\begin{aligned}\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\hat\varphi(\xi)\psi(\xi)\,\mathrm{d}\xi &= \int_{\mathbb{R}^n_\xi}\left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi}\varphi(y)\,\mathrm{d}y\right)\psi(\xi)\,\mathrm{d}\xi = \int_{\mathbb{R}^n_y}\varphi(y)\left(\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi}\psi(\xi)\,\mathrm{d}\xi\right)\mathrm{d}y\\ &= \int_{\mathbb{R}^n_y}\varphi(y)\hat\psi(y-x)\,\mathrm{d}y = \int_{\mathbb{R}^n_y}\varphi(x+y)\hat\psi(y)\,\mathrm{d}y.\end{aligned}$$

Now let us replace $\psi(\xi)$ by $\psi(\varepsilon\xi)$, for some positive ε. Then we have to replace the Fourier transform $\hat\psi(y)$ by $\varepsilon^{-n}\hat\psi(y/\varepsilon)$, and it follows that

$$\begin{aligned}\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\hat\varphi(\xi)\psi(\varepsilon\xi)\,\mathrm{d}\xi &= \int_{\mathbb{R}^n_y}\varphi(x+y)\varepsilon^{-n}\hat\psi\left(\frac{y}{\varepsilon}\right)\mathrm{d}y\\ &= \int_{\mathbb{R}^n_y}\varphi(x+\varepsilon y)\hat\psi(y)\,\mathrm{d}y.\end{aligned} \tag{2.3}$$

We send ε to $+0$ and apply the Convergence theorem of Lebesgue, which yields

$$\psi(0)\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\hat\varphi(\xi)\,\mathrm{d}\xi = \varphi(x)\int_{\mathbb{R}^n_y}\hat\psi(y)\,\mathrm{d}y.$$

And now we can choose ψ like this: $\psi(\xi) = \exp(-\frac{1}{2}|\xi|^2)$. Then we have $\psi(0) = 1$ and $\hat\psi(y) = (2\pi)^{n/2}\exp(-\frac{1}{2}|y|^2)$, as well as $\int_{\mathbb{R}^n_y}\hat\psi(y)\,\mathrm{d}y = (2\pi)^n$. □

Corollary 2.8. *From (2.3) we immediately obtain*

$$\int_{\mathbb{R}^n}\hat\varphi\psi\,\mathrm{d}x = \int_{\mathbb{R}^n}\varphi\hat\psi\,\mathrm{d}x, \tag{2.4}$$

by setting $x = 0$ and $\varepsilon = 1$.

We have only shown that the Fourier transform is an isomorphism between algebraic vector spaces. However, Proposition 2.4 enables us to quickly prove that the Fourier transform is also a topological isomorphism (exercise).

Example 2.9. *Take $f = f(x) \in L^1(\mathbb{R}^1)$ with $f(x) = 1$ for $0 \le x \le 1$, and $f(x) = 0$ for all other x. Then*

$$\hat f(\xi) = \int_{x=-\infty}^{\infty} e^{-\mathrm{i}x\cdot\xi}f(x)\,\mathrm{d}x = \int_{x=0}^{1} e^{-\mathrm{i}x\cdot\xi}\,\mathrm{d}x = \frac{1-e^{-\mathrm{i}\xi}}{\mathrm{i}\xi} \qquad (\xi\neq 0), \qquad \hat f(0) = 1,$$

and this is a continuous function on $\mathbb{R}^1$, but $\hat f$ does not belong to $L^1(\mathbb{R}^1_\xi)$ because it does not decay fast enough for $\xi \to \infty$. Unfortunately, the inversion formula (2.2) has no meaning as an integral in $L^1(\mathbb{R}^1)$. However, the next lemma will give a positive result.

Lemma 2.10 (**Inverse Fourier transform for some non-smooth functions**)**.** *Let $f \in L^1(\mathbb{R}^1)$ be continuous, except a finite number of jumps. Then*

$$\frac{1}{2\pi}\lim_{R\to\infty}\int_{\xi=-R}^{R} e^{\mathrm{i}x\cdot\xi}\hat f(\xi)\,\mathrm{d}\xi = \begin{cases} f(x) & : \ f \text{ is continuous at } x,\\ \frac{1}{2}(f(x+0)+f(x-0)) & : \ f \text{ jumps at } x.\end{cases}$$

Proof. This is quoted from [7], Satz 8.2. □

The space $L^2(\mathbb{R}^n)$ is of great physical importance because it has a scalar product, in contrast to $L^1(\mathbb{R}^n)$. The Fourier transform is not yet defined on $L^2(\mathbb{R}^n)$ because there are functions in $L^2(\mathbb{R}^n)$ which are not in $L^1(\mathbb{R}^n)$. The definition of $\mathcal{F}$ on $L^2(\mathbb{R}^n)$ will be made possible by $\mathcal{S}(\mathbb{R}^n)$ being *sequentially dense* in $L^2(\mathbb{R}^n)$: for each $f \in L^2(\mathbb{R}^n)$, there is a sequence $(f_1, f_2, \dots) \subset \mathcal{S}(\mathbb{R}^n)$ with $\lim_{j\to\infty} \|f_j - f\|_{L^2(\mathbb{R}^n)} = 0$. And because $\mathcal{S}(\mathbb{R}^n)$ is a Fréchet space (and $L^2(\mathbb{R}^n)$ is a metric space), we are happy because in this setting *sequentially dense* means the same as *topologically dense* (the topological closure of the smaller set is the bigger set).

Definition 2.11 (**Fourier transform on** $L^2(\mathbb{R}^n)$)**.** *For $f \in L^2(\mathbb{R}^n)$, let $(f_1, f_2, \dots) \subset \mathcal{S}(\mathbb{R}^n)$ be a sequence approximating f. Then we define*

$$\hat{f}(\xi) := \lim_{j\to\infty} \hat{f}_j(\xi).$$

This limit is independent of the choice of the sequence $(f_1, f_2, \dots)$. The convergence of the sequence $(\hat{f}_1, \hat{f}_2, \dots)$ in the norm of $L^2(\mathbb{R}^n_\xi)$ follows from the Parseval formula.

We draw an intermediate summary: the difference between the Fourier transform $\mathcal{F}$ and the inverse transform $\mathcal{F}^{-1}$ are the exchange of $\exp(+\mathrm{i}x\cdot\xi)$ against $\exp(-\mathrm{i}x\cdot\xi)$, and an additional factor $(2\pi)^{-n}$. Then it is no surprise that similar rules as given in Proposition 2.4 hold also for the inverse transform, and in particular we mention

$$\mathcal{F}^{-1}_{\xi\to x}(\hat{f}\cdot\hat{g})(x) = (f * g)(x),$$

with the convolution in the x–world defined as

$$(f * g)(x) := \int_{\mathbb{R}^n_y} f(y)\cdot g(x-y)\,\mathrm{d}y.$$

Note that the differential is now $\mathrm{d}y$, instead of $\text{đ}\eta$ as in (2.1).

2.2 Distribution Theory

This chapter is following [14]. Another nice presentation can be found in [28], and a further approach (which is quite nice from the functional analytic aspect) is in [18]. Compare also [26] and [27].

2.2.1 Purpose

In this chapter, we assume $\Omega \subset \mathbb{R}^n$ to be a domain (hence open and connected), not necessarily bounded. We make no assumption on the regularity of the boundary $\partial\Omega$.

We wish to define generalised functions (called *distributions*) with the following properties:

- every "reasonable function" over Ω can be understood as a distribution, where "reasonable" could mean $L^1_{\mathrm{loc}}(\Omega)$;
- every distribution can be differentiated as often as we want, and the result is again a distribution;
- we have a meaning of a distribution being the limit of a sequence of distributions;
- we have various operations that we can let act upon distributions (adding them, multiplying them by numbers and by smooth functions, computing their Fourier transforms);
- these operations are continuous in a sense of appropriately chosen topological vector spaces.

Unfortunately, we have a price to pay: it will not be possible to define products of distributions. In that sense, the distribution theory will always be a linear theory.

2.2.2 Test Functions and Distributions

Definition 2.12. *Let A and B be subsets of $\mathbb{R}^n$. We write $A \Subset B$ if the closure of A is a compact subset of $\mathbb{R}^n$, and the closure of A is contained in the interior of B.*

After this preparation, we can start:

Definition 2.13 (Test function space $\mathcal{D}(\Omega)$). *We set*

$$C_0^\infty(\Omega) = \{u \colon \Omega \to \mathbb{C} \colon u \text{ is infinitely often differentiable, } \operatorname{supp} u \Subset \Omega\},$$

and we write $\mathcal{D}(\Omega) := C_0^\infty(\Omega)$. We define a convergence in this space as follows: we say that a sequence $(\varphi_j)_{j\to\infty}$ converges to φ in $\mathcal{D}(\Omega)$ if there is a compact set $K \Subset \Omega$ with $\operatorname{supp} \varphi_j \subset K$ for all j, and

$$\lim_{j\to\infty} \sup_{x\in K} |\partial_x^\alpha(\varphi_j(x) - \varphi(x))| = 0$$

for all multi-indices $\alpha \in \mathbb{N}^n$. This convergence is also written as

$$\varphi_j \xrightarrow{\mathcal{D}} \varphi \quad (j \to \infty).$$

Obviously, $\mathcal{D}(\Omega)$ is a vector space over the field $\mathbb{C}$. It is also a topological vector space, but we make no attempt at explaining what its collection of open sets is, because we do not really need it, and we only mention that the gory details can be found in Section 4.6 of [15], or in Appendix B of [13]. Unfortunately, $\mathcal{D}(\Omega)$ is no Fréchet space.

Now we are in a position to define distributions:

Definition 2.14 (Set $\mathcal{D}'(\Omega)$ of distributions). *The set*

$$\mathcal{D}'(\Omega) = \{T \colon \mathcal{D}(\Omega) \to \mathbb{C} \colon T \text{ is linear and sequentially continuous}\}$$

is called set of distributions over Ω.

Clearly, linearity means $T(\alpha\varphi + \beta\psi) = \alpha T(\varphi) + \beta T(\psi)$ for all $\alpha, \beta \in \mathbb{C}$ and all $\varphi, \psi \in \mathcal{D}(\Omega)$; and the sequential continuity means that $\lim_{j\to\infty} T(\varphi_j) = T(\varphi)$ for all sequences $(\varphi_j)_{j\to\infty}$ with

$$\varphi_j \xrightarrow{\mathcal{D}} \varphi \quad (j \to \infty).$$

We quickly check that $\mathcal{D}'(\Omega)$ is again $\mathbb{C}$ vector space (even a topological one).

An equivalent definition is this one (without proof of equivalency):

Definition 2.15 (Distribution space $\mathcal{D}'(\Omega)$). *A linear set $T \colon \mathcal{D}(\Omega) \to \mathbb{C}$ is called distribution if: for each $K \Subset \Omega$ there are constants $C \in \mathbb{R}$, $k \in \mathbb{N}$ with*

$$|T(\varphi)| \le C \|\varphi\|_{C_b^k(K)} \qquad \forall \varphi \in C_0^\infty(\Omega), \qquad \operatorname{supp} \varphi \subset K.$$

If we can choose the same number k for all compact sets K, then the smallest such k is called the order *of T.*

Example: *Let $f \in C(\Omega)$ and $T(\varphi) = \int_\Omega f(x)\varphi(x)\,\mathrm{d}x$ for $\varphi \in C_0^\infty(\Omega)$. Note that f can have arbitrarily strong singularities at the boundary of Ω. Obviously, we could relax the assumption on f to $f \in L^1_{\mathrm{loc}}(\Omega)$, which means $f \in L^1(K)$ for each $K \Subset \Omega$.*

Example: *Pick $x_0 \in \Omega$ and $\alpha \in \mathbb{N}^n$. Then we set $T(\varphi) = (\partial_x^\alpha \varphi)(x_0)$ for $\varphi \in C_0^\infty(\Omega)$.*

Example: *Let $(x_1, x_2, \dots)$ be a sequence in Ω with limit on the boundary $\partial\Omega$. Let $(\alpha^{(1)}, \alpha^{(2)}, \dots)$ be a sequence of multi-indices, and choose a sequence $(a_1, a_2, \dots)$ in $\mathbb{C}$. Then we define*

$$T(\varphi) = \sum_{j=1}^\infty a_j \cdot (\partial_x^{\alpha^{(j)}} \varphi)(x_j), \qquad \varphi \in C_0^\infty(\Omega).$$

For fixed φ, this series contains only a finite number of non-vanishing items. Because the sequence of integers $|\alpha^{(j)}|$ may be unbounded, this distribution T perhaps has infinite order.

Every function $f \in L^1_{\mathrm{loc}}(\Omega)$ generates a distribution T_f according to the formula $T_f(\varphi) := \int_\Omega f(x)\varphi(x)\,\mathrm{d}x$. We quickly check that this map $f \mapsto T_f$ is injective. And this concept is so important that it deserves a name: T_f is called a *regular distribution*.

It is in this sense that we consider distributions as generalised functions.

Proposition 2.16 (Dirac–**distribution**[5]). *Assume* $0 \in \Omega$. *Pick* $f \in C_0^\infty(\Omega)$ *with* $\int_\Omega f(x)\,\mathrm{d}x = 1$ *and* $f(x) \geq 0$ *for all* $x \in \Omega$. *We then define*

$$f_\varepsilon(x) := \frac{1}{\varepsilon^n} f\left(\frac{x}{\varepsilon}\right), \qquad \varepsilon > 0.$$

Then we have $\lim_{\varepsilon \to +0} \int_\Omega f_\varepsilon(x)\varphi(x)\,\mathrm{d}x = \varphi(0)$ *for all* $\varphi \in C_0^\infty(\Omega)$.

Proof. Exercise. □

That distribution which maps a test function φ to its value at $x = 0$ is called *Delta Distribution of Dirac*. We observe that this distribution can be approximated by a sequence of regular distributions T_{f_ε}, in the sense of convergence in $\mathcal{D}'$, which we define now:

Definition 2.17 (**Convergence in** $\mathcal{D}'$). *A sequence* $(T_1, T_2, \dots)$ *of distributions converges to a distribution* $T \in \mathcal{D}'(\Omega)$ *if* $\lim_{j\to\infty} T_j(\varphi) = T(\varphi)$ *for each* $\varphi \in \mathcal{D}(\Omega)$.

In other words, the vector space $\mathcal{D}'(\Omega)$ is being equipped with the weak-$*$ topology.

Every distribution can be approximated by regular distributions:

Proposition 2.18. *Let* $T \in \mathcal{D}'(\Omega)$. *Then there is a sequence* $(f_1, f_2, \dots) \subset C_0^\infty(\Omega)$ *with* $\lim_{j\to\infty} T_{f_j} = T$ *in* $\mathcal{D}'(\Omega)$, *in other words*

$$\lim_{j\to\infty} \int_\Omega f_j(x)\varphi(x)\,\mathrm{d}x = T(\varphi), \qquad \forall \varphi \in C_0^\infty(\Omega).$$

This can be expressed as $\mathcal{D}(\Omega)$ being *sequentially dense* in $\mathcal{D}'(\Omega)$.

Proof. See [14], volume 1, Theorem 4.1.5. □

2.2.3 Operations for Distributions

Up to now, only a few operations have been defined in $\mathcal{D}'(\Omega)$, namely addition, and multiplication by numbers.

Now we will define:

- multiplication by a smooth function of $C^\infty(\Omega)$,
- differentiation.

And we choose the definition in such a way that

- we get the old meaning back if the distribution under consideration is a regular distribution generated by a smooth function;
- the operation is a continuous linear operation from $\mathcal{D}'(\Omega)$ into $\mathcal{D}'(\Omega)$.

Let $T \in \mathcal{D}'(\Omega)$ be given, and take a sequence $(f_1, f_2, \dots) \subset C_0^\infty(\Omega)$ with

$$T(\varphi) = \lim_{j\to\infty} \int_\Omega f_j(x)\varphi(x)\,\mathrm{d}x, \qquad \forall \varphi \in C_0^\infty(\Omega),$$

which we can abbreviate as $T_{f_j} \xrightarrow{\mathcal{D}'} T$ for $j \to \infty$.

For some function $a = a(x) \in C^\infty(\Omega)$, we wish to define $aT \in \mathcal{D}'(\Omega)$, under the condition $(aT_{f_j}) \xrightarrow{\mathcal{D}'} aT$ for $j \to \infty$, which is equivalent to

$$(aT_{f_j})(\varphi) \xrightarrow{\mathbb{C}} (aT)(\varphi), \quad (j \to \infty), \qquad \forall \varphi \in C_0^\infty(\Omega).$$

[5] Paul Adrien Maurice Dirac, 1902–1984

Now $af_j \in C_0^\infty(\Omega)$ allows to rewrite the LHS into

$$(aT_{f_j})(\varphi) = \int_\Omega a(x)f_j(x)\cdot\varphi(x)\,\mathrm{d}x = \int_\Omega f_j(x)\cdot a(x)\varphi(x)\,\mathrm{d}x = T_{f_j}(a\varphi) \xrightarrow{\mathbb{C}} T(a\varphi) \quad (j\to\infty).$$

And this therefore enforces the following definition:

Definition 2.19 (**Multiplication of a function and a distribution**). *Let $T \in \mathcal{D}'(\Omega)$ and $a \in C^\infty(\Omega)$. Then we define a distribution $aT \in \mathcal{D}'(\Omega)$ by means of*

$$(aT)(\varphi) := T(a\varphi), \qquad \varphi \in C_0^\infty(\Omega).$$

Remark 2.20. *The product aT can also be meaningfully defined for $a, T \in \mathcal{D}'(\Omega)$ under the assumption* sing-supp $a \cap$ sing-supp $T = \emptyset$*, with the singular support defined below in Section 2.2.5. Otherwise it gets tricky: it can be shown that it is not possible to define a multiplication of distributions in such a way that the law of associativity holds, and that we get the classical multiplication back in case of functions. A partial remedy are the* Colombeau*–algebras (which we don't have time for, unfortunately).*

Next we intend to define $\partial_x^\alpha T$, under the above two conditions. In particular $\partial_x^\alpha T_{f_j} \xrightarrow{\mathcal{D}'} \partial_x^\alpha T$, which means

$$(\partial_x^\alpha T_{f_j})(\varphi) \xrightarrow{\mathbb{C}} (\partial_x^\alpha T)(\varphi), \qquad \forall \varphi \in C_0^\infty(\Omega).$$

And again, we can rewrite the LHS into

$$(\partial_x^\alpha T_{f_j})(\varphi) = \int_\Omega (\partial_x^\alpha f_j(x))\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|}\int_\Omega f_j(x)(\partial_x^\alpha\varphi(x))\,\mathrm{d}x = (-1)^{|\alpha|}T_{f_j}(\partial_x^\alpha\varphi) \xrightarrow{\mathbb{C}} (-1)^{|\alpha|}T(\partial_x^\alpha\varphi).$$

Hence we must define:

Definition 2.21 (**Derivatives of distributions**). *Let $T \in \mathcal{D}'(\Omega)$ and $\alpha \in \mathbb{N}^n$. Then we define a distribution $\partial_x^\alpha T \in \mathcal{D}'(\Omega)$ as*

$$(\partial_x^\alpha T)(\varphi) := (-1)^{|\alpha|}T(\partial_x^\alpha\varphi), \qquad \varphi \in C_0^\infty(\Omega).$$

Remark 2.22. *It holds $\partial_x^\alpha\partial_x^\beta T = \partial_x^\beta\partial_x^\alpha T$.*

Proposition 2.23. *The mappings $T \mapsto aT$ and $T \mapsto \partial_x^\alpha T$ are linear and continuous as mappings from $\mathcal{D}'(\Omega)$ into $\mathcal{D}'(\Omega)$.*

Proof. This follows directly from Proposition A.20. □

Example 2.24. *Let $H\colon \mathbb{R}^1 \to \mathbb{R}^1$ be the* Heaviside[6] *function, defined as*

$$H(x) = \begin{cases} 1\colon x > 0, \\ 0\colon x \le 0. \end{cases}$$

Then $H' = \delta$. You can also choose $H(0) = 1$, and again you get $H' = \delta$.

We introduce a notation.

Definition 2.25. *For a distribution $T \in \mathcal{D}'(\Omega)$ and a test function $\varphi \in \mathcal{D}(\Omega)$, we let*

$$\langle T, \varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} \in \mathbb{C}$$

stand for the $T(\varphi)$, which is the result of applying the distribution T to the test function φ.

In case T is a regular distribution generated by a function f, we get

$$\langle T_f, \varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \int_\Omega f(x)\cdot\varphi(x)\,\mathrm{d}x.$$

Then we have the identities

$$\langle aT, \varphi\rangle_{\mathcal{D}'(?\Omega)\times\mathcal{D}(\Omega)} = \langle T, a\varphi\rangle_{\mathcal{D}'(?\Omega)\times\mathcal{D}(\Omega)}, \qquad T \in \mathcal{D}'(\Omega), \quad \varphi \in \mathcal{D}(\Omega), \quad a \in C^\infty(\Omega),$$
$$\langle \partial_x^\alpha T, \varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = (-1)^{|\alpha|}\langle T, \partial_x^\alpha\varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)}, \qquad T \in \mathcal{D}'(\Omega), \quad \varphi \in \mathcal{D}(\Omega), \quad \alpha \in \mathbb{N}_0^n.$$

[6] Oliver Heaviside, 1850–1925

2.2.4 Fourier Transformation of Distributions

We wish to explain what the Fourier transform of $\sin(x)$ or x^2 is. Unfortunately, these functions belong neither to $L^1(\mathbb{R}^n)$, nor to $L^2(\mathbb{R}^n)$ or $\mathcal{S}(\mathbb{R}^n)$. At least, they have at most polynomial growth for $|x| \to \infty$.

We now follow the (at the beginning confusing) convention of writing a slowly growing function f and its associated Schwartz distribution $T_f \in \mathcal{D}'(\mathbb{R}^n)$ by the same symbol f. Then the expression $f = f(\cdot)$ becomes ambiguous (the dot could be an x or a test function φ), and we resolve this equivocation by writing $\langle f, \varphi\rangle_{\mathcal{D}'\times\mathcal{D}}$ when we apply the distribution $T_f \in \mathcal{D}'$ to the test function $\varphi \in \mathcal{D}$.

In the same way we have defined a distribution space $\mathcal{D}'(\Omega)$ as the topological dual vector space of $\mathcal{D}(\Omega)$, we now define one more distribution space $\mathcal{S}'(\mathbb{R}^n)$ as topological dual vector space of $\mathcal{S}(\mathbb{R}^n)$:

Definition 2.26 (**Schwartz distributions**). *A map $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is called a* Schwartz distribution *if it is linear and sequentially continuous. Here sequential continuity means that*

$$\text{if} \quad \varphi_j \xrightarrow{\mathcal{S}} \varphi \quad \text{then} \quad \lim_{j\to\infty} T(\varphi_j) = T(\varphi).$$

We introduce the notation

$$\langle T, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} := T(\varphi), \qquad T \in \mathcal{S}'(\mathbb{R}^n), \quad \varphi \in \mathcal{S}(\mathbb{R}^n)$$

as we did for $\mathcal{D}'$. The set (it is even a vector space) of all Schwartz distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$.

Example 2.27. *The Delta distribution located at a point $x_0 \in \mathbb{R}^n$ is a Schwartz distribution:*

$$\delta_{x_0}(\varphi) := \langle \delta_{x_0}, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} := \varphi(x_0), \qquad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Example 2.28. *Each function f which is piecewise continuous and grows at infinity at most polynomially generates a Schwartz distribution T_f:*

$$T_f(\varphi) := \int_{\mathbb{R}^n} f(x)\varphi(x)\,\mathrm{d}x, \qquad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Exponentially growing functions f will in general not generate a Schwartz distribution, but the devil is in the detail *(more on that below). Anyway, since typically only a moderate growth of f is admissible, these distributions are often called* temperate distributions[7].

Why do we need yet another space of distributions ? We clearly see that $\mathcal{D}(\mathbb{R}^n)$ is "smaller" than $\mathcal{S}(\mathbb{R}^n)$, so we may conjecture that $\mathcal{D}'(\mathbb{R}^n)$ is "bigger" than $\mathcal{S}'(\mathbb{R}^n)$ (more on that below), and the hope is that $\mathcal{S}'(\mathbb{R}^n)$ has less "monsters" than $\mathcal{D}'(\mathbb{R}^n)$, which makes $\mathcal{S}'(\mathbb{R}^n)$ easier to handle than $\mathcal{D}'(\mathbb{R}^n)$. Well, why did we introduce $\mathcal{D}$ and $\mathcal{D}'$ then ? The answer is that $\mathcal{S}$ and $\mathcal{S}'$ can not be defined on domains Ω, only on $\mathbb{R}^n$. If we are only interested in pseudodifferential operators on the full space $\mathbb{R}^n$, then we can indeed skip the discussion of $\mathcal{D}$ and $\mathcal{D}'$. A second motivation for defining $\mathcal{D}'$ is the Schwartz kernel theorem coming later.

Now we figure out the relations between $\mathcal{D}(\mathbb{R}^n)$ and $\mathcal{S}$, and between $\mathcal{D}'(\mathbb{R}^n)$ and $\mathcal{S}'$.

Lemma 2.29. *The test function space $\mathcal{D}(\mathbb{R}^n)$ is sequentially dense in $\mathcal{S}(\mathbb{R}^n)$, and the embedding operator is sequentially continuous.*

Proof. Exercise. You have to show:

$$\forall\, \varphi \in \mathcal{S}(\mathbb{R}^n) \quad \exists\,(\psi_1, \psi_2, \dots) \subset \mathcal{D}(\mathbb{R}^n)\colon \quad \psi_j \xrightarrow{\mathcal{S}} \varphi \quad (j \to \infty);$$
$$\forall\,(\psi_1, \psi_2, \dots) \subset \mathcal{D}(\mathbb{R}^n), \quad \forall \psi \in \mathcal{D}(\mathbb{R}^n)\colon \quad \psi_j \xrightarrow{\mathcal{D}(\mathbb{R}^n)} \psi \quad (j \to \infty) \Longrightarrow \psi_j \xrightarrow{\mathcal{S}(\mathbb{R}^n)} \psi \quad (j \to \infty).$$

□

Therefore, every linear and sequentially continuous map from $\mathcal{S}(\mathbb{R}^n)$ into $\mathbb{C}$ is also a linear and sequentially continuous map from $\mathcal{D}(\mathbb{R}^n)$ into $\mathbb{C}$. This gives us $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$ as a subset relation.

[7] Perhaps you will read sometimes the expression *tempered* distribution, even written by a native speaker of English. Such authors have no taste.

We make the agreement

$$\boxed{\text{If } T \in \mathcal{S}'(\mathbb{R}^n) \text{ and } \varphi \in \mathcal{D}(\mathbb{R}^n) \text{ then } \langle T, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle T, \varphi\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)}.}$$

We have even more: that map which reads some $T \in \mathcal{S}'(\mathbb{R}^n)$ as an element of $\mathcal{D}'(\mathbb{R}^n)$, is *injective*. Because: assume $T_1, T_2 \in \mathcal{S}'(\mathbb{R}^n)$ are equal as elements of $\mathcal{D}'(\mathbb{R}^n)$. This means (by definition) $\langle T_1, \psi\rangle_{\mathcal{D}'\times\mathcal{D}} = \langle T_2, \psi\rangle_{\mathcal{D}'\times\mathcal{D}}$ for each $\psi \in \mathcal{D}(\mathbb{R}^n)$. Choose some $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then there is a sequence $(\psi_1, \psi_2, \dots) \subset \mathcal{D}(\mathbb{R}^n)$ with $\psi_j \xrightarrow{\mathcal{S}} \varphi$. Then we get

$$\langle T_1, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} \xleftarrow{\mathcal{S}} \langle T_1, \psi_j\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle T_1, \psi_j\rangle_{\mathcal{D}'\times\mathcal{D}} = \langle T_2, \psi_j\rangle_{\mathcal{D}'\times\mathcal{D}} = \langle T_2, \psi_j\rangle_{\mathcal{S}'\times\mathcal{S}} \xrightarrow{\mathcal{S}} \langle T_2, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}},$$

which implies $\langle T_1, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle T_2, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}}$, hence $T_1 = T_2$ as elements of $\mathcal{S}'$.

Corollary 2.30. *We have the continuous embedding* $\mathcal{S}'(\mathbb{R}^n) \hookrightarrow \mathcal{D}'(\mathbb{R}^n)$.

We can also lift these considerations onto a more formal level. The natural inclusion map $\iota\colon \mathcal{D}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is injective. And it is sequentially continuous (every sequence in $\mathcal{D}(\mathbb{R}^n)$ that converges in $\mathcal{D}(\mathbb{R}^n)$, also converges in $\mathcal{S}(\mathbb{R}^n)$). And $\mathcal{S}$ is a Fréchet space, and the topology in $\mathcal{D}$ has been designed in such a way that then ι is continuous in the topological sense. We have shown above that ι is also a sequentially dense embedding, which makes the transposed map $\iota^t\colon \mathcal{S}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ injective. And the transposed map ι^t is always continuous (when we equip $\mathcal{S}'$ and $\mathcal{D}'$ with their weak-$*$ topologies). By definition of ι^t, we have

$$\langle T, \iota\varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = \left\langle \iota^t T, \varphi\right\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)}, \qquad \forall\, T \in \mathcal{S}'(\mathbb{R}^n), \quad \forall\, \varphi \in \mathcal{D}(\mathbb{R}^n),$$

which is the boxed agreement from above in a different notation. Hence ι^t is indeed the map that restricts the domain of a $\mathcal{S}'$ distribution from $\mathcal{S}$ to $\mathcal{D}(\mathbb{R}^n)$.

Now comes the crazy example of a function with exponential growth that is still a member of $\mathcal{S}'$:

Lemma 2.31. *The function* $f = f(x) = \exp(x)\cdot\cos(\exp(x)) = \partial_x(\sin(\exp(x)))$ *generates a distribution* $T_f \in \mathcal{S}'(\mathbb{R}^1)$ *in the sense of*

$$\langle T_f, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} := \lim_{R\to\infty} \int_{|x|\le R} f(x)\varphi(x)\,\mathrm{d}x.$$

Proof. This follows from

$$\int_{|x|\le R} f(x)\varphi(x)\,\mathrm{d}x = \sin(\exp(x))\cdot\varphi(x)\Big|_{x=-R}^{x=R} - \int_{|x|\le R} \sin(\exp(x))\cdot\varphi'(x)\,\mathrm{d}x,$$

hence

$$\langle T_f, \varphi\rangle_{\mathcal{S}'\times\mathcal{S}} = -\int_{\mathbb{R}^1} \sin(\exp(x))\cdot\varphi'(x)\,\mathrm{d}x.$$

□

The reason is that there are two types of integrals in $\mathbb{R}^n$ that have to be distinguished. One is the usual Lebesgue integral as in the definition of $L^1(\mathbb{R}^n)$, the other is $\lim_{R\to\infty}\int_{|x|<R}\dots\,\mathrm{d}x$. These are not the same, as can be seen from the function $g(x) = \frac{\sin(x)}{x}$ which is integrable in the second sense, but not in the first (recall that if g were a member of $L^1(\mathbb{R}^n)$, then so would $|g|$).

Now we define the Fourier transform in $\mathcal{S}'$. To this end, we recall (2.4), which can be recast into

$$\left\langle \hat{f}, g\right\rangle_{\mathcal{S}'\times\mathcal{S}} = \langle f, \hat{g}\rangle_{\mathcal{S}'\times\mathcal{S}}, \qquad f, g \in \mathcal{S}(\mathbb{R}^n).$$

Keep in mind that a distribution from $\mathcal{S}'(\mathbb{R}^n)$ need not be a function; it could be also a Delta distribution.

Definition 2.32 (**Fourier transform on** $\mathcal{S}'(\mathbb{R}^n)$). *For a distribution $T \in \mathcal{S}'(\mathbb{R}^n)$, we define its Fourier transform $\hat{T} \in \mathcal{S}'(\mathbb{R}^n)$ via*

$$\left\langle \hat{T}, \varphi \right\rangle_{\mathcal{S}' \times \mathcal{S}} := \langle f, \hat{\varphi} \rangle_{\mathcal{S}' \times \mathcal{S}}, \qquad \forall\, \varphi \in \mathcal{S}(\mathbb{R}^n).$$

We also write $\mathcal{F}T$ in place of $\hat{T}$.

Example 2.33. *What is $\hat{\delta}$?*

$$\left\langle \hat{\delta}, \varphi \right\rangle_{\mathcal{S}' \times \mathcal{S}} := \langle \delta, \hat{\varphi} \rangle_{\mathcal{S}' \times \mathcal{S}} = \hat{\varphi}(0) = \int_{\mathbb{R}^n} e^{-\mathrm{i}0 \cdot x} \cdot \varphi(x)\, \mathrm{d}x = \langle 1, \varphi \rangle_{\mathcal{S}' \times \mathcal{S}}.$$

Answer: the Fourier transform of Dirac's delta is that function which is one everywhere.

Lemma 2.34. *The Fourier transform is a topological isomorphism from $\mathcal{S}'(\mathbb{R}^n)$ onto $\mathcal{S}'(\mathbb{R}^n)$. Taking a derivative ∂_x^α maps $\mathcal{S}'(\mathbb{R}^n)$ continuously into itself. Multiplying by $a \in C^\infty(\mathbb{R}^n)$ maps continuously from $\mathcal{S}'(\mathbb{R}^n)$ into $\mathcal{D}'(\mathbb{R}^n)$ (even into $\mathcal{S}'(\mathbb{R}^n)$ in case of $a \in C_b^\infty(\mathbb{R}^n)$).*

Proof. Repeated application of Proposition A.20. □

Lemma 2.35. *The formula $(D_x^\alpha f)\hat{\ }(\xi) = \xi^\alpha \hat{f}(\xi)$ holds also for $f \in \mathcal{S}'$.*

Proof. From $f \in \mathcal{S}'(\mathbb{R}^n)$ we get $D_x^\alpha f \in \mathcal{S}'(\mathbb{R}^n)$, and then also $\mathcal{F}(D_x^\alpha f) \in \mathcal{S}'(\mathbb{R}^n)$. Take some arbitrary $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then we have

$$\begin{aligned}
\langle \mathcal{F}(D_x^\alpha f), \varphi \rangle_{\mathcal{S}' \times \mathcal{S}} &= \langle D_x^\alpha f, \mathcal{F}(\varphi) \rangle_{\mathcal{S}' \times \mathcal{S}} && \text{because of Definition 2.32} \\
&= (-1)^{|\alpha|} \langle f, D_x^\alpha \mathcal{F}(\varphi) \rangle_{\mathcal{S}' \times \mathcal{S}} && \text{because of Definition 2.21 and Lemma 2.34} \\
&= \langle f, \mathcal{F}(x^\alpha \varphi(x)) \rangle_{\mathcal{S}' \times \mathcal{S}} && \text{from a variant of Proposition 2.4} \\
&= \langle \mathcal{F}(f), x^\alpha \varphi(x) \rangle_{\mathcal{S}' \times \mathcal{S}} && \text{from Definition 2.32} \\
&= \left\langle x^\alpha \hat{f}(x), \varphi(x) \right\rangle_{\mathcal{S}'(\mathbb{R}^n_x) \times \mathcal{S}(\mathbb{R}^n_x)}.
\end{aligned}$$

Now we only have to rename x to ξ. □

2.2.5 Support of Distributions

For functions f it is nice to know,

- where they have non-trivial values (support of f, $\operatorname{supp} f$),
- where they have singularities (singular support of f, $\operatorname{sing-supp} f$).

We say that a function has a singularity at a certain point if it is there *not* infinitely differentiable.

We wish to extend these notions to distributions, in such a way that the old meaning and the new meaning are compatible for regular distributions.

Definition 2.36 (**Support** $\operatorname{supp} T$). *Let $T \in \mathcal{D}'(\Omega)$. We say that a point $x_0 \in \Omega$ belongs to the support of T ($x_0 \in \operatorname{supp} T$) if for each open neighbourhood $U \subset \Omega$ of x_0, there is a function $\varphi \in \mathcal{D}(U)$ with $\langle T, \varphi \rangle_{\mathcal{D}'(\Omega) \times \mathcal{D}(\Omega)} \neq 0$.*

Another wording: $x_0 \notin \operatorname{supp} T$ if and only if there is an open neighbourhood U of x_0 such that $U \subset \Omega$ and $\langle T, \varphi \rangle_{\mathcal{D}'(\Omega) \times \mathcal{D}(\Omega)} = 0$ for all $\varphi \in \mathcal{D}(U)$.

Definition 2.37 (**Singular support** $\operatorname{sing-supp} T$). *Let $T \in \mathcal{D}'(\Omega)$. We say that a point $x_0 \in \Omega$* does not *belong to the singular support of T ($x_0 \notin \operatorname{sing-supp} T$) if an open neighbourhood $U \subset \Omega$ of x_0 and a function $f \in C^\infty(U)$ exist with $\langle T, \varphi \rangle_{\mathcal{D}'(\Omega) \times \mathcal{D}(\Omega)} = \int_U f(x)\varphi(x)\, \mathrm{d}x$ for all $\varphi \in \mathcal{D}(U)$.*

This means: in a neighbourhood of x_0, T is identical to a regular distribution that is being generated by a C^∞ function.

Another wording: $x_1 \in \text{sing-supp}\, T$ if and only if there is no open neighbourhood U of x_1, for which the restriction of T to U equals a C^∞ function.

Here we have defined the restriction T_U of T to U by means of $\langle T_U, \varphi\rangle_{\mathcal{D}'(U)\times\mathcal{D}(U)} := \langle T, \varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)}$ for all $\varphi \in \mathcal{D}(U) \subset \mathcal{D}(\Omega)$.

Now we take a PDO $\mathcal{P} = \mathcal{P}(x, D_x) = \sum_{|\alpha|\le m} a_\alpha(x) D_x^\alpha$ with coefficients $a_\alpha \in C^\infty(\Omega)$. Then we clearly have, for each function $u \in C^m(\Omega)$, that

$$\text{sing-supp}(\mathcal{P}u) \subseteq \text{sing-supp}\, u.$$

Later we will learn that equality holds in case $\mathcal{P}$ being an *elliptic* PDO.

Sometimes the following is useful:

Proposition 2.38. *Let $T \in \mathcal{D}'(\Omega)$ be a distribution of order k with support* $\operatorname{supp} T = \{x_0\} \in \Omega$. *Then there are numbers $a_\alpha \in \mathbb{C}$ for $|\alpha| \le k$ such that*

$$\langle T, \varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \sum_{|\alpha|\le k} a_\alpha \cdot (\partial_x^\alpha \varphi)(x_0), \qquad \forall \varphi \in \mathcal{D}(\Omega).$$

Proof. See [14], Volume 1, Theorem 2.3.4. □

2.2.6 The Schwartz Kernel Theorem

We start with some preparations. Consider the vector spaces $\mathbb{R}^n$ and $\mathbb{R}^m$. How do linear maps $\mathcal{K}$ from $\mathbb{R}^n$ into $\mathbb{R}^m$ look like ? We know that each such linear map $\mathcal{K}$ is being generated by a matrix K, in the sense of

$$(\mathcal{K}u)_j = \sum_{i=1}^n K_{ji} u_i, \qquad \forall u \in \mathbb{R}^n, \quad \forall j \in \{1, \dots, m\}.$$

Now we ask for all linear maps $\mathcal{K}$ from $L^2(\Omega_1)$ into $L^2(\Omega_2)$. Reading (j, i) as (x, y) and $\sum$ as $\int$, we guess that the answer might be

$$(\mathcal{K}u)(x) \stackrel{?}{=} \int_{\Omega_1} K(x, y) u(y)\, \mathrm{d}y, \qquad u \in L^2(\Omega_1), \quad x \in \Omega_2.$$

This guess is true provided that we require $\mathcal{K}$ to be a continuous map between the two normed spaces, and provided that we allow K to be a distribution from $\mathcal{D}'(\Omega_1 \times \Omega_2)$. Then we will have to define what this formula actually means.

However, for practical applications we want more. The requirement that $\mathcal{K}$ be a continuous map from $L^2(\Omega_1)$ into $L^2(\Omega_2)$ will for instance exclude differential operators, which would be embarrassing because this whole course is about differential operators. The good news however is that the *Schwartz Kernel Theorem* applies even to linear maps $\mathcal{K}$ from $\mathcal{D}(\Omega_1)$ into $\mathcal{D}'(\Omega_2)$ that are continuous in the respective topologies. Every operator we will ever be interested in does satisfy this condition.

Now let us get more rigorous.

Definition 2.39 (Tensor product of functions). *Let $\Omega_1 \subset \mathbb{R}^{n_1}$ and $\Omega_2 \subset \mathbb{R}^{n_2}$ be domains. For functions $u_j \in C(\Omega_j)$, $j = 1, 2$, we define their tensor product $u_1 \otimes u_2 \in C(\Omega_1 \times \Omega_2)$ as*

$$(u_1 \otimes u_2)(x_1, x_2) := u_1(x_1) \cdot u_2(x_2), \quad (x_1, x_2) \in \Omega_1 \times \Omega_2.$$

Now let us consider a function $K = K(x_1, x_2) \in C(\Omega_1 \times \Omega_2)$, and we define

$$(\mathcal{K}\varphi)(x_1) := \int_{\Omega_2} K(x_1, x_2)\varphi(x_2)\,\mathrm{d}x_2$$

for any function $\varphi \in C_0(\Omega_2)$, which is defined as the space of continuous functions with compact support in Ω_2. Then we have $\mathcal{K}\varphi \in C(\Omega_1)$. Observe that K may have arbitrarily strong poles for $x_1 \to \partial\Omega_1$ or $x_2 \to \partial\Omega_2$.

If we pair this equation with some $\psi \in C_0(\Omega_1)$, then we get

$$\int_{\Omega_1} (\mathcal{K}\varphi)(x_1) \cdot \psi(x_1)\,\mathrm{d}x_1 = \iint_{\Omega_1\times\Omega_2} K(x_1, x_2) \cdot (\psi \otimes \varphi)(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2.$$

Now we may read the left factor in the integrand of the LHS as a distribution. This shall help us in understanding the meaning of the Schwartz Kernel Theorem.

Theorem 2.40 (Schwartz **kernel theorem).**

1. *Let $K \in \mathcal{D}'(\Omega_1 \times \Omega_2)$. For each $\varphi \in \mathcal{D}(\Omega_2)$, we define a distribution $\mathcal{K}\varphi \in \mathcal{D}'(\Omega_1)$ via*

   $$\langle \mathcal{K}\varphi, \psi\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)} := \langle K, \psi \otimes \varphi\rangle_{\mathcal{D}'(\Omega_1\times\Omega_2)\times\mathcal{D}(\Omega_1\times\Omega_2)}, \qquad \psi \in \mathcal{D}(\Omega_1).$$

   Then the following holds: this map $\varphi \mapsto \mathcal{K}\varphi$ is linear and sequentially continuous from $\mathcal{D}(\Omega_2)$ into $\mathcal{D}'(\Omega_1)$ in the following sense:

   If $\varphi_j \overset{\mathcal{D}(\Omega_2)}{\longrightarrow} 0$, then $\mathcal{K}\varphi_j \overset{\mathcal{D}'(\Omega_1)}{\longrightarrow} 0$ (both for $j \to \infty$).

2. *Let $\mathcal{K}$ be a map from $\mathcal{D}(\Omega_2)$ into $\mathcal{D}'(\Omega_1)$, which is linear and continuous in the above sense. Then there is exactly one distribution $K \in \mathcal{D}'(\Omega_1 \times \Omega_2)$ with*

   $$\langle \mathcal{K}\varphi, \psi\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)} = \langle K, \psi \otimes \varphi\rangle_{\mathcal{D}'(\Omega_1\times\Omega_2)\times\mathcal{D}(\Omega_1\times\Omega_2)},$$

   for all $\psi \in \mathcal{D}(\Omega_1)$ and all $\varphi \in \mathcal{D}(\Omega_2)$.

This distribution K is called kernel *of the operator $\mathcal{K}$.*

Proof. See [14], Volume 1, Theorem 5.2.1. □

Example 2.41. *Let $\Omega_1 = \Omega_2$ and $\mathcal{K}\varphi = \varphi$, which makes $\mathcal{K}$ be the identity map. Then we have $\langle \mathcal{K}\varphi, \psi\rangle = \int_{\Omega_1} \varphi(x_1)\psi(x_1)\,\mathrm{d}x_1$, and the distribution K turns out to act like this,*

$$\langle K, \Phi\rangle_{\mathcal{D}'(\Omega_1\times\Omega_1)\times\mathcal{D}(\Omega_1\times\Omega_1)} = \int_{\Omega_1} \Phi(x_1, x_1)\,\mathrm{d}x_1,$$

on test functions $\Phi \in \mathcal{D}(\Omega_1 \times \Omega_1)$. The support of the kernel K is the diagonal:

$$\operatorname{supp} K = \operatorname{diag}(\Omega_1 \times \Omega_1) := \{(x_1, x_1) \in \Omega_1 \times \Omega_1 :\ x_1 \in \Omega_1\}.$$

Observe that K is no function or regular distribution. In physicists' jargon, we have $K(x_1, x_2) = \delta_{x_1=x_2}$.

2.2.7 Smoothing Operators

What happens if the Schwartz kernel K is a C^∞ function ?

Unfortunately, we need more test functions and more distributions.

Definition 2.42 (The test function space $\mathcal{E}$ and its topology). *We define $\mathcal{E}(\Omega) := C^\infty(\Omega)$, equipped with the semi-norms*

$$\varphi \mapsto \sum_{|\alpha|\le m} \sup_{x\in K} |\partial_x^\alpha \varphi(x)|,$$

where $m \in \mathbb{N}$, and $K \Subset \Omega$ runs through all compact sets. We endow that space with the locally convex topology that is being generated by these semi-norms.

Since it suffices to take a countable selection of compact sets K, $\mathcal{E}$ turns out to be a Fréchet space.

Definition 2.43 (**Distribution space** $\mathcal{E}'(\Omega)$)**.** *We define a set*

$$\mathcal{E}'(\Omega) := \{T \in \mathcal{D}'(\Omega) \colon \operatorname{supp} T \Subset \Omega\}.$$

The elements of $\mathcal{E}'(\Omega)$ *are called* distributions with compact support*.*

This is just a set, for now. Next we define how to apply $T \in \mathcal{E}'(\Omega)$ to some $\varphi \in C^\infty(\Omega)$. The trouble is that typically $\varphi \notin C_0^\infty(\Omega)$.

To this end, we choose compact sets K_1 and K_2 with

$$\operatorname{supp} T \Subset K_1 \Subset K_2 \Subset \Omega,$$

and we choose a cut-off function $\psi = \psi(x)$ with $\psi \in C_0^\infty(\Omega)$ and $\operatorname{supp}\psi \subset K_2$ as well as $\psi \equiv 1$ on K_1. Then we define

$$T(\varphi) := \langle T, \psi\varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)}, \qquad \varphi \in C^\infty(\Omega).$$

This construction gives us a complex number that does not depend on ψ or the choices of K_1, K_2. Because let $\tilde{K}_1$, $\tilde{K}_2$ be other compact sets, and let $\tilde{\psi}$ be another cut-off function (with the same properties as above). Then we have

$$\left\langle T, \tilde{\psi}\varphi\right\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \langle T, \psi\varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)}$$

owing to $\operatorname{supp} T \cap \operatorname{supp}(\tilde{\psi}\varphi - \psi\varphi) = \emptyset$.

We see that every $T \in \mathcal{E}'(\Omega)$ generates a linear map from $\mathcal{E}(\Omega)$ into $\mathbb{C}$ (nothing said about continuity).

Proposition 2.44. *The vector space* $\mathcal{E}'(\Omega)$ *coincides with the topologically defined dual space of the Fréchet space* $\mathcal{E}(\Omega)$*.*

Then we may introduce the notation

$$\langle T, \varphi\rangle_{\mathcal{E}'(\Omega)\times\mathcal{E}(\Omega)}$$

with the meaning $T(\varphi)$ defined above, for $T \in \mathcal{E}'(\Omega)$ and $\varphi \in \mathcal{E}(\Omega)$.

We remark that, by this very construction, we have the compatibility

$$\langle T, \varphi\rangle_{\mathcal{E}'(\Omega)\times\mathcal{E}(\Omega)} = \langle T, \varphi\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)}, \qquad T \in \mathcal{E}'(\Omega), \quad \varphi \in \mathcal{D}(\Omega).$$

Now we come to the key result that can be paraphrased as "smooth kernels generate smoothing operators, and vice versa".

Theorem 2.45 (**Smoothing operators**)**.**

1. *Let* $K \in C^\infty(\Omega_1 \times \Omega_2)$*, and define* $\mathcal{K} \colon \mathcal{D}(\Omega_2) \to \mathcal{D}'(\Omega_1)$ *by*

   $$\langle \mathcal{K}\varphi, \psi\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)} := \langle K, \psi \otimes \varphi\rangle, \qquad \forall \varphi \in \mathcal{D}(\Omega_2), \quad \psi \in \mathcal{D}(\Omega_1).$$

   Then the operator $\mathcal{K}$ *can be continuously extended to a linear map*

   $$\mathcal{K} \colon \mathcal{E}'(\Omega_2) \to \mathcal{E}(\Omega_1)$$

   by means of

   $$(\mathcal{K}\varphi)(x_1) := \langle \varphi, K(x_1, \cdot)\rangle_{\mathcal{E}'(\Omega_2)\times\mathcal{E}(\Omega_2)}, \qquad \varphi \in \mathcal{E}'(\Omega_2), \quad x_1 \in \Omega_1. \tag{2.5}$$

2. *If* $\mathcal{K} \colon \mathcal{E}'(\Omega_2) \to \mathcal{E}(\Omega_1)$ *is linear and continuous, then there is exactly one* $K \in C^\infty(\Omega_1 \times \Omega_2)$ *with (2.5).*

Example 2.46. *Choose* $\Omega = \mathbb{R}^n$*. Let* $(\mathcal{K}\varphi)(x) = (\varphi * f_\varepsilon)(x)$*, for* $\varphi \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$*, and* $f_\varepsilon(x) = \varepsilon^{-n} f(x/\varepsilon)$*, where* $f \in C_0^\infty(\mathbb{R}^n)$ *with* $f(x) \ge 0$ *for all* $x \in \mathbb{R}^n$*, and* $\int_{\mathbb{R}^n} f(x)\,\mathrm{d}x = 1$*. In other words*

$$(\mathcal{K}\varphi)(x) = \int_{\mathbb{R}^n_y} f_\varepsilon(x-y)\varphi(y)\,\mathrm{d}y,$$

and this integral operator has the kernel function $K(x,y) = f_\varepsilon(x-y)$*, which is obviously smooth, and it has its support in a tubular neighbourhood of width* $\mathcal{O}(\varepsilon)$ *about the diagonal of* $\mathbb{R}^n \times \mathbb{R}^n$*.*

2.3 Integral Operators

2.3.1 Convolutions

For integrable functions φ_1 on φ_2 on $\mathbb{R}^n$, we define their convolution $\varphi_1 * \varphi_2$ on $\mathbb{R}^n$ as

$$(\varphi_1 * \varphi_2)(x) := \int_{\mathbb{R}^n_y} \varphi_1(x-y)\varphi_2(y)\,\mathrm{d}y,$$

provided this integral exists in the Lebesgue sense.

It is easy to show using substitution that the convolution is commutative if one factor has compact support. And the convolution is associative if two factors have compact support, which can be shown by Fubini[8].

It is common belief that the "the convolution is associative whenever all integrals exist". Well, *the devil is in the detail.* The following counter-example has been provided by Rudin[9] [19].

Example 2.47. *Take a function Φ on $\mathbb{R}$,*

$$\Phi(x) := \begin{cases} 30x^2(1-x)^2 & : 0 \le x \le 1, \\ 0 & : \textit{else}, \end{cases}$$

and define $f(x) := 1$ for all $x \in \mathbb{R}$, $g(x) := \Phi'(x)$, and $h(x) := \int_{-\infty}^x \Phi(t)\,\mathrm{d}t$.
Then we calculate

$$(f*g)(x) = \int_{y=-\infty}^{\infty} f(x-y)g(y)\,\mathrm{d}y = \int_{y=-\infty}^{\infty} g(y)\,\mathrm{d}y = \int_{y=0}^{1} g(y)\,\mathrm{d}y = \Phi(1)-\Phi(0) = 0,$$

and therefore $(f*g)*h \equiv 0$ on $\mathbb{R}$.
On other hand, a partial integration reveals

$$\begin{aligned}(g*h)(x) &= \int_{y=-\infty}^{\infty} g(x-y)h(y)\,\mathrm{d}y = \int_{y=x-1}^{x} \Phi'(x-y)h(y)\,\mathrm{d}y \\ &= -\Phi(x-y)h(y)\Big|_{y=x-1}^{y=x} + \int_{y=x-1}^{x} \Phi(x-y)h'(y)\,\mathrm{d}y \\ &= \int_{y=x-1}^{x} \Phi(x-y)\Phi(y)\,\mathrm{d}y,\end{aligned}$$

which is positive whenever $0 < x < 2$. The conclusion then is $(f*(g*h))(x) > 0$ for all $x \in \mathbb{R}$.
Find what went wrong !

The following presentation is guided by [14, Volume 1].

Lemma 2.48. *If $p_1, \ldots, p_k \in [1,\infty]$ are Lebesgue exponents, with $\frac{1}{p_1} + \ldots + \frac{1}{p_k} = k-1+\frac{1}{q}$, for some $q \in [1,\infty]$, and $\varphi_1, \ldots, \varphi_k$ are integrable functions with compact support, then*

$$\|\varphi_1 * \varphi_2 * \cdots * \varphi_k\|_{L^q(\mathbb{R}^n)} \le \|\varphi_1\|_{L^{p_1}(\mathbb{R}^n)} \cdot \ldots \cdot \|\varphi_k\|_{L^{p_k}(\mathbb{R}^n)}.$$

Proof. This is Corollary 4.5.2 in [14, Volume 1]. □

Lemma 2.49. *Let $\Omega = \Omega_1 \times \Omega_2$. Take $u_1 \in \mathcal{D}'(\Omega_1)$ and $\varphi \in \mathcal{D}(\Omega)$. Then the function*

$$x_2 \mapsto \psi(x_2) := \langle u_1, \varphi(\cdot, x_2)\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)}$$

is a member of $\mathcal{D}(\Omega_2)$, and we have, for each α,

$$\partial^\alpha_{x_2}\psi(x_2) = \left\langle u_1, \partial^\alpha_{x_2}\varphi(\cdot, x_2)\right\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)}.$$

The same holds with $u_1 \in \mathcal{E}'(\Omega_1)$, $\varphi \in \mathcal{E}(\Omega)$ and $\psi \in \mathcal{E}(\Omega_2)$.

[8] Guido Fubini, 1879–1943
[9] Walter Rudin, 1921–2010

For a proof, it is recommended to go back to the definition of derivatives as limits of quotients of differences.

Definition 2.50. *For a distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ and a function $\varphi \in \mathcal{D}(\mathbb{R}^n)$, we define their convolution*

$$(u * \varphi)(x) := \langle u, \varphi(x - \cdot)\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)} .$$

By the previous lemma, this is a smooth function, depending on $x \in \mathbb{R}^n$.

Lemma 2.51. *If $u \in \mathcal{D}'(\mathbb{R}^n)$ and $\varphi \in \mathcal{D}(\mathbb{R})$, then $u * \varphi \in C^\infty(\mathbb{R}^n)$, and we can differentiate like this:*

$$D^\alpha(u * \varphi) = (D^\alpha u) * \varphi = u * (D^\alpha \varphi).$$

Proof. This is Theorem 4.1.1 in [14, Volume 1]. □

Lemma 2.52. *If φ, $\psi \in \mathcal{D}(\mathbb{R}^n)$ and $u \in \mathcal{D}'(\mathbb{R}^n)$, then*

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

Proof. This is Theorem 4.1.2 in [14, Volume 1]. □

Corollary 2.53. *Let $u \in \mathcal{D}'(\mathbb{R}^n)$, and φ, $\psi \in \mathcal{D}(\mathbb{R}^n)$. Put $\omega(y) := \varphi(-y)$. Then*

$$\langle u * \varphi, \psi\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)} = \langle u, \omega * \psi\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)} .$$

Proof. Put $\zeta(y) := \psi(-y)$. Then we calculate

$$\begin{aligned}
\langle u * \varphi, \psi\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)} &= \int_{\mathbb{R}^n_y} (u * \varphi)(y) \cdot \psi(y)\,\mathrm{d}y = \int_{\mathbb{R}^n_y} (u * \varphi)(y) \cdot \zeta(0 - y)\,\mathrm{d}y \\
&= \Big((u * \varphi) * \zeta\Big)(0) = \Big(u * (\varphi * \zeta)\Big)(0) = \langle u, (\varphi * \zeta)(0 - \cdot)\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)} , \\
(\varphi * \zeta)(-y) &= \int_{\mathbb{R}^n_z} \varphi(z)\zeta(-y - z)\,\mathrm{d}z = \int_{\mathbb{R}^n_z} \varphi(z)\psi(y + z)\,\mathrm{d}z = \int_{\mathbb{R}^n_z} \varphi(-z)\psi(y - z)\,\mathrm{d}z \\
&= \int_{\mathbb{R}^n_z} \omega(z)\psi(y - z)\,\mathrm{d}z = (\omega * \psi)(y),
\end{aligned}$$

which completes the proof. □

Definition 2.54. *Let u_1, $u_2 \in \mathcal{D}'(\mathbb{R}^n)$, and one of them has compact support. Then $u_1 * u_2$ is defined as that uniquely determined distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ for which*

$$u_1 * (u_2 * \varphi) = u * \varphi, \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^n).$$

The discussion in Section 4.2 of [14] shows that such a distribution u indeed exists uniquely.

The convolution of two distributions is commutative provided at least one of them has compact support. The convolution of three distributions is associative provided at least two of them have compact support.

Proposition 2.55. *Let $u_1 \in \mathcal{E}'(\mathbb{R}^n)$ and $u_2 \in \mathcal{S}'(\mathbb{R}^n)$. Then $u_1 * u_2$ belongs to $\mathcal{S}'(\mathbb{R}^n)$, and its Fourier transform is*

$$(u_1 * u_2)\widehat{\ } = \hat{u}_1 \hat{u}_2.$$

Proof. This is Theorem 7.1.5 in [14, Volume 1]. Note that $\hat{u}_1 \in \mathcal{E}(\mathbb{R}^n)$ (it is even analytic) because u_1 has compact support. □

It should be clear that for each function φ, we have $\varphi * \delta = \varphi$. Now let u be a distribution. Then

$$(D^\alpha u) * \varphi = u * (D^\alpha \varphi) = u * ((D^\alpha \varphi) * \delta) = u * ((D^\alpha \delta) * \varphi) = (u * D^\alpha \delta) * \varphi,$$

which holds for all $\varphi \in \mathcal{D}(\mathbb{R}^n)$. The result is then $D^\alpha u = u * D^\alpha \delta$.

Now we take two distributions, u_1 and u_2, and one of them has compact support. Then

$$D^\alpha(u_1 * u_2) = (u_1 * u_2) * D^\alpha \delta = u_1 * (u_2 * D^\alpha \delta) = u_1 * D^\alpha u_2.$$

Similarly, we show $D^\alpha(u_1 * u_2) = (D^\alpha u_1) * u_2$.

Definition 2.56. *Consider a PDO $\mathcal{P}(D) = \sum_{|\alpha|\le m} a_\alpha D^\alpha$ with* ***constant*** *coefficients $a_\alpha \in \mathbb{C}$. A distribution $E \in \mathcal{D}'(\mathbb{R}^n)$ is called* fundamental solution for $\mathcal{P}$ *if*

$$\mathcal{P}(D)E = \delta.$$

Theorem 7.3.10 in [14, Volume 1] guarantees that such a fundamental solution always exists, and it is a distribution of finite order.

The key purpose of the fundamental solution E is an easy formula for the solution u to $\mathcal{P}u = f$:

$$\mathcal{P}(D)\Big(E * f\Big) = \Big(\mathcal{P}(D)E\Big) * f = \delta * f = f, \qquad f \in \mathcal{E}'(\mathbb{R}^n). \tag{2.6}$$

Therefore, $u := E * f$ solves $\mathcal{P}u = f$ provided $f \in \mathcal{E}'(\mathbb{R}^n)$. In the many examples in the motivation Chapter, the distribution E was typically a function (and called K), and we have seen there many convolution integrals.

And also, for a distribution $u \in \mathcal{E}'(\mathbb{R}^n)$, we have

$$E * \Big(\mathcal{P}(D)u\Big) = \mathcal{P}(D)\Big(E * u\Big) = \Big(\mathcal{P}(D)E\Big) * u = \delta * u = u,$$

which means that $E*$ is at the same time the right-inverse operator and the left-inverse operator for $\mathcal{P}(D)$, but on $\mathcal{E}'(\mathbb{R}^n)$. Nothing has been claimed about $\mathcal{D}'(\mathbb{R}^n)$ or $\mathcal{S}'(\mathbb{R}^n)$.

We apply the theory of convolutions to tensor products of distributions, which we will define soon:

Lemma 2.57. *Put $\Omega := \Omega_1 \times \Omega_2$, where $\Omega_1 \subset \mathbb{R}^n$ and $\Omega_2 \subset \mathbb{R}^n$. Given $u_1 \in \mathcal{D}'(\Omega_1)$ and $u_2 \in \mathcal{D}'(\Omega_2)$, there is* ***at most*** *one $u \in \mathcal{D}'(\Omega)$ with*

$$\langle u, \varphi_1 \otimes \varphi_2\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \langle u_1, \varphi_1\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)} \cdot \langle u_2, \varphi_2\rangle_{\mathcal{D}'(\Omega_2)\times\mathcal{D}(\Omega_2)}, \qquad \forall \varphi_j \in \mathcal{D}(\Omega_j).$$

Proof. Suppose $u \in \mathcal{D}'(\Omega)$ and $\langle u, \varphi_1 \otimes \varphi_2\rangle = 0$ for all φ_1, φ_2. Choose $K \Subset \Omega$. Take $\chi_1 \in \mathcal{D}(\Omega_1)$ and $\chi_2 \in \mathcal{D}(\Omega_2)$ with $\chi := \chi_1 \otimes \chi_2 \equiv 1$ on K. Then $\chi u \in \mathcal{D}'(\Omega)$ with compact support in Ω, hence $\chi u \in \mathcal{E}'(\mathbb{R}^n)$. Take an even function $\psi \in \mathcal{D}(\mathbb{R}^n)$, $\psi \ge 0$, $\int_{\mathbb{R}^n} \psi(x)\,\mathrm{d}x = 1$, and consider

$$\Psi_\varepsilon(x_1, x_2) := \varepsilon^{-2n}\psi(x_1/\varepsilon)\cdot\psi(x_2/\varepsilon), \qquad 0 < \varepsilon < 1.$$

For small ε and $x \in K$ we then can conclude

$$\Big((\chi u) * \Psi_\varepsilon\Big)(x) = \langle \chi u, \Psi_\varepsilon(x-\cdot)\rangle_{\mathcal{D}'(\mathbb{R}^{2n})\times\mathcal{D}(\mathbb{R}^{2n})} = \langle u, \chi(\cdot)\Psi_\varepsilon(x-\cdot)\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = 0, \tag{2.7}$$

because $\chi(\cdot)\Psi_\varepsilon(x-\cdot)$ has tensor product form.

On the other hand, take any $\varphi \in \mathcal{D}(\Omega)$ with $\operatorname{supp}\varphi \subset K$. By Corollary 2.53, and Ψ_ε being even, we have

$$\langle (\chi u) * \Psi_\varepsilon, \varphi\rangle_{\mathcal{D}'(K)\times\mathcal{D}(K)} = \langle (\chi u) * \Psi_\varepsilon, \varphi\rangle_{\mathcal{D}'(\mathbb{R}^{2n})\times\mathcal{D}(\mathbb{R}^{2n})} = \langle \chi u, \Psi_\varepsilon * \varphi\rangle_{\mathcal{D}'(\mathbb{R}^{2n})\times\mathcal{D}(\mathbb{R}^{2n})},$$

and $\Psi_\varepsilon * \varphi$ converges, for $\varepsilon \to +0$, to φ, in the topology of $\mathcal{D}(\mathbb{R}^{2n})$. The result then is

$$\lim_{\varepsilon\to+0} \langle (\chi u) * \Psi_\varepsilon, \varphi\rangle_{\mathcal{D}'(K)\times\mathcal{D}(K)} = \langle \chi u, \varphi\rangle_{\mathcal{D}'(\mathbb{R}^{2n})\times\mathcal{D}(\mathbb{R}^{2n})}.$$

Reconciling this with (2.7) then requires $\chi u \equiv 0$ on K. But $K \Subset \Omega$ was arbitrary. □

Lemma 2.58. *Put $\Omega := \Omega_1 \times \Omega_2$, where $\Omega_1 \subset \mathbb{R}^n$ and $\Omega_2 \subset \mathbb{R}^n$. Given $u_1 \in \mathcal{D}'(\Omega_1)$ and $u_2 \in \mathcal{D}'(\Omega_2)$, there is* ***exactly*** *one $u \in \mathcal{D}'(\Omega)$ with*

$$\langle u, \varphi_1 \otimes \varphi_2\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \langle u_1, \varphi_1\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)} \cdot \langle u_2, \varphi_2\rangle_{\mathcal{D}'(\Omega_2)\times\mathcal{D}(\Omega_2)}, \qquad \forall \varphi_j \in \mathcal{D}(\Omega_j).$$

Proof. Define some linear map $u\colon \mathcal{D}(\Omega) \to \mathbb{C}$, applied to $\varphi \in \mathcal{D}(\Omega)$, by

$$u(\varphi) := \Big\langle u_1(:\,), \langle u_2(\cdot), \varphi(:\,,\cdot)\rangle_{\mathcal{D}'(\Omega_2)\times\mathcal{D}(\Omega_2)}\Big\rangle_{\mathcal{D}'(\Omega_1)\times\mathcal{D}(\Omega_1)}. \tag{2.8}$$

From Lemma 2.49 we obtain the continuity of u as a map from $\mathcal{D}(\Omega)$ to $\mathbb{C}$. □

Definition 2.59 (Tensor product of distributions). *Under the notations of Lemma 2.58, we define the tensor product* $u := u_1 \otimes u_2$ *by (2.8).*

There is no reason why u_1 has to be the "outer" distribution and u_2 the "inner". Swapping them gives the same u, and this observation is written as $u_1 \otimes u_2 = u_2 \otimes u_1$. Some people call it FUBINI.

Example: *Take* $u_1 = f \in \mathcal{D}'(\mathbb{R}^n)$ *and* $u_2 \equiv 1 \in \mathcal{D}'(\mathbb{R}^n)$. *Then we have, for all* $\varphi \in \mathcal{D}(\mathbb{R}^n \times \mathbb{R}^n)$,

$$\left\langle f(\cdot), \int_{\mathbb{R}^n_y} \varphi(\cdot, y)\,\mathrm{d}y \right\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)} = \int_{\mathbb{R}^n_y} \langle f(\cdot), \varphi(\cdot, y)\rangle_{\mathcal{D}'(\mathbb{R}^n)\times\mathcal{D}(\mathbb{R}^n)}\,\mathrm{d}y.$$

By density, the seven $\mathcal{D}$ *may be replaced by seven* $\mathcal{S}$.

2.3.2 Integral Operators on Domains

We generalise the convolution integrals.

Proposition 2.60. *Let* $\Omega \subset \mathbb{R}^n$ *be a domain, and let* $K = K(x, y)$ *be a measurable function on* $\Omega \times \Omega$. *Suppose that*

$$\forall y \in \Omega\colon \int_{\Omega_x} |K(x, y)|\,\mathrm{d}x \le C_1, \qquad \forall x \in \Omega\colon \int_{\Omega_y} |K(x, y)|\,\mathrm{d}y \le C_2.$$

Then the integral operator $\mathcal{K}$, *given by*

$$(\mathcal{K}u)(x) := \int_{\Omega_y} K(x, y)u(y)\,\mathrm{d}y \tag{2.9}$$

has the bound

$$\|\mathcal{K}u\|_{L^p(\Omega)} \le C_1^{1/p} C_2^{1/q} \|u\|_{L^p(\Omega)}, \qquad 1 \le p \le \infty, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{2.10}$$

Proof. The cases $p = 1$ and $p = \infty$ are left to the readers.

Choose some function $v \in L^q(\Omega)$, and put

$$T_v := \iint_{\Omega^2} K(x, y)u(y)v(x)\,\mathrm{d}x\,\mathrm{d}y.$$

For some $t \in (0, \infty)$, estimate the product $u(y)v(x)$ in the integrand according to the principle

$$|uv| = \left| tu \cdot \frac{1}{t} v \right| \le \frac{1}{p}|tu|^p + \frac{1}{q}\left|\frac{1}{t}v\right|^q,$$

resulting in

$$|T_v| \le \frac{C_1}{p}\|tu\|^p_{L^p(\Omega)} + \frac{C_2}{q}\left\|\frac{1}{t}v\right\|^q_{L^q(\Omega)}.$$

Now minimise the RHS over t, resulting in

$$|T_v| \le C_1^{1/p} C_2^{1/q} \|u\|_{L^p(\Omega)} \|v\|_{L^q(\Omega)}.$$

The mapping $v \mapsto T_v$, from $L^q(\Omega)$ into $\mathbb{C}$, is linear and bounded, hence an element of the dual space $(L^q(\Omega))'$, and its operator norm $\|T\|_{\mathrm{op}}$ is bounded by

$$\|T\|_{\mathrm{op}} \le C_1^{1/p} C_2^{1/q} \|u\|_{L^p(\Omega)}.$$

But each element of the dual space $(L^q(\Omega))'$ can be represented (in a unique way) via an integral pairing with an element from $L^p(\Omega)$, whose L^p norm equals the operator norm of T; and, by the very definition of T_v, this element of $L^p(\Omega)$ is $\mathcal{K}u$. Thus

$$\|\mathcal{K}u\|_{L^p(\Omega)} = \|T\|_{\mathrm{op}},$$

which concludes the proof. □

This result resembles an interpolation between the *row sum norm* and the *column sum norm* of a matrix. However, the following estimate resembles the FROBENIUS[10] norm of a matrix:

Lemma 2.61. *Let $\Omega \subset \mathbb{R}^n$ be a domain, and let $K = K(x, y)$ be a measurable function on $\Omega \times \Omega$. Suppose*

$$C_0 := \left(\iint_{\Omega^2} |K(x, y)|^2 \, \mathrm{d}x \, \mathrm{d}y \right)^{1/2} < \infty.$$

Then the integral operator $\mathcal{K}$, given by (2.9), maps from $L^2(\Omega)$ into $L^2(\Omega)$, with norm $\|\mathcal{K}\|_{\mathrm{op}} \le C_0$.

Proof. This is a special case of Statement 3.13 in [4]. □

[10] FERDINAND GEORG FROBENIUS, 1849–1917

Chapter 3

Oscillating Integrals

This chapter follows [22].

3.1 Definition

Consider a domain $\Omega \subset \mathbb{R}^n$ and a function $u \in C_0^\infty(\Omega)$, which we may extend to all of $\mathbb{R}^n$ by zero if necessary.

Our intention is to make sense of terms

$$I_\Phi(au) := \iint_{\mathbb{R}^N_\theta \times \Omega_z} e^{\mathrm{i}\Phi(z,\theta)} a(z,\theta) u(z) \,\mathrm{d}z \,\mathrm{d}\theta \tag{3.1}$$

where $\theta \in \mathbb{R}^N$, and a may grow polynomially for $|\theta| \to \infty$. We remark that $N \neq n$ is allowed.

Definition 3.1 (**Symbol classes**). *Let $m \in \mathbb{R}$ and $0 \leq \delta, \varrho \leq 1$. We say that a function $a = a(z,\theta) \in C^\infty(\Omega \times \mathbb{R}^N)$ belongs to the* symbol class $S^m_{\varrho,\delta}(\Omega \times \mathbb{R}^N)$ *if the following holds: for each pair of multi-indices α, β and for each compact set $K \Subset \Omega$, there is a constant $C_{\alpha\beta K}$*

$$\left|\partial_\theta^\alpha \partial_z^\beta a(z,\theta)\right| \leq C_{\alpha\beta K} \langle\theta\rangle^{m-\varrho|\alpha|+\delta|\beta|}, \quad \forall (z,\theta) \in K \times \mathbb{R}^N.$$

Here we have set $\langle\theta\rangle := \sqrt{1+|\theta|^2}$. Moreover, we define $S^{-\infty} = \bigcap_{m\in\mathbb{R}} S^m_{\varrho,\delta}$. (This intersection is independent of ϱ and δ.)

These inequalities are called *symbol estimates*.

Since the constants $C_{\alpha\beta K}$ are allowed to depend on the compact set K, the function a may have arbitrarily strong poles at the boundary $\partial\Omega$ (called *local* symbol estimates). The most common case however is $\varrho = 1$, $\delta = 0$, and $C_{\alpha\beta K}$ can be chosen independently of K (*global* symbol estimates).

Definition 3.2 (**Phase function**). *A function $\Phi = \Phi(z,\theta)$ is called* phase function *if the following holds:*

- *$\Phi \in C^\infty(\Omega \times (\mathbb{R}^N \setminus \{0\}))$, with values in $\mathbb{R}$,*
- *$\Phi(z, t\theta) = t\Phi(z,\theta)$, for all $(z,\theta) \in \Omega \times (\mathbb{R}^N \setminus \{0\})$ and all $t > 0$,*
- *Φ has no critical points if $\theta \neq 0$, which means $\nabla_{z,\theta}\Phi(z,\theta) \neq 0$ for all $(z,\theta) \in \Omega \times (\mathbb{R}^N \setminus \{0\})$.*

Definition 3.3. *If $a \in S^m_{\varrho,\delta}$, $u \in C^\infty(\Omega)$, and Φ is a phase function, then the expression $I_\Phi(au)$ is called* oscillating integral *(we claim nothing about its existence as a converging integral).*

Proposition 3.4. *Let Φ be a phase function. Then there is a PDO L on $\Omega \times \mathbb{R}^N$ of the form*

$$\begin{cases} L = \sum_{j=1}^N a_j(z,\theta)\partial_{\theta_j} + \sum_{k=1}^n b_k(z,\theta)\partial_{z_k} + c(z,\theta), \\ a_j \in S^0_{1,0}(\Omega \times \mathbb{R}^N), \quad b_k \in S^{-1}_{1,0}(\Omega \times \mathbb{R}^N), \quad c \in S^{-1}_{1,0}(\Omega \times \mathbb{R}^N), \\ a_j(z,\theta) = b_k(z,\theta) = 0, \quad z \in \Omega, \quad |\theta| \leq 1, \end{cases} \tag{3.2}$$

such that

$$L^t(z,\theta,\partial_z,\partial_\theta)e^{\mathrm{i}\Phi(z,\theta)} = e^{\mathrm{i}\Phi(z,\theta)}, \qquad (z,\theta)\in\Omega\times(\mathbb{R}^N\setminus 0), \tag{3.3}$$

where L^t is the formally transposed operator to L:

$$(L^t u)(z,\theta) = -\sum_{j=1}^N \partial_{\theta_j}(a_j u) - \sum_{k=1}^n \partial_{z_k}(b_k u) + cu.$$

We have introduced the shorter notation $\mathbb{R}^N\setminus 0 := \mathbb{R}^N\setminus\{0\}$.

The proof will show that there infinitely many such operators L.

Remark 3.5. *Recall that the formally transposed operator satisfies*

$$\iint_{\Omega\times\mathbb{R}^N}(Lu)v\,\mathrm{d}z\,\mathrm{d}\theta = \iint_{\Omega\times\mathbb{R}^N} u(L^t v)\,\mathrm{d}z\,\mathrm{d}\theta, \qquad u,v\in C_0^\infty(\Omega\times\mathbb{R}^N),$$

and also $(L^t)^t = L$.

Proof. We quickly check that we have, for all $(z,\theta)\in\Omega\times(\mathbb{R}^N\setminus 0)$:

$$\partial_{\theta_j}e^{\mathrm{i}\Phi} = \mathrm{i}\frac{\partial\Phi}{\partial\theta_j}e^{\mathrm{i}\Phi}, \qquad \partial_{z_k}e^{\mathrm{i}\Phi} = \mathrm{i}\frac{\partial\Phi}{\partial z_k}e^{\mathrm{i}\Phi},$$

$$\left(-\mathrm{i}\sum_{j=1}^N \frac{\partial\Phi}{\partial\theta_j}|\theta|^2\partial_{\theta_j} - \mathrm{i}\sum_{k=1}^n \frac{\partial\Phi}{\partial z_k}\partial_{z_k}\right)e^{\mathrm{i}\Phi} = \left(\sum_{j=1}^N |\theta|^2\left(\frac{\partial\Phi}{\partial\theta_j}\right)^2 + \sum_{k=1}^n\left(\frac{\partial\Phi}{\partial z_k}\right)^2\right)e^{\mathrm{i}\Phi}$$
$$=: \frac{1}{\psi(z,\theta)}e^{\mathrm{i}\Phi(z,\theta)}.$$

We observe $\psi\in C^\infty(\Omega\times(\mathbb{R}^N\setminus 0))$ with positive homogeneity in θ of order -2, which means $\psi(z,t\theta) = t^{-2}\psi(z,\theta)$ for all $(z,\theta)\in\Omega\times(\mathbb{R}^N\setminus 0)$ and all $t>0$.

From this we obtain

$$-\mathrm{i}\psi\cdot\left(\sum_{j=1}^N|\theta|^2\frac{\partial\Phi}{\partial\theta_j}\partial_{\theta_j} + \sum_{k=1}^n\frac{\partial\Phi}{\partial z_k}\partial_{z_k}\right)e^{\mathrm{i}\Phi} = e^{\mathrm{i}\Phi}.$$

We are now almost done. The only problem is that the coefficients have a singularity for $\theta=0$. To remove them, we introduce a cut-off function $\chi=\chi(\theta)\in C_0^\infty(\mathbb{R}^N)$ with $\chi(\theta)=1$ for $|\theta|\le 1$ and $\chi(\theta)=0$ for $|\theta|\ge 2$. Then we define

$$M(z,\theta,\partial_z,\partial_\theta) := -\sum_{j=1}^N(1-\chi(\theta))\mathrm{i}\psi|\theta|^2\frac{\partial\Phi}{\partial\theta_j}\partial_{\theta_j} - \sum_{k=1}^n(1-\chi(\theta))\mathrm{i}\psi\frac{\partial\Phi}{\partial z_k}\partial_{z_k} + \chi(\theta)$$

and we check that $Me^{\mathrm{i}\Phi} = e^{\mathrm{i}\Phi}$ for all $(z,\theta)\in\Omega\times\mathbb{R}^N$. The coefficients of M are smooth, and they are in the desired symbol classes $S^0_{1,0}$ and $S^{-1}_{1,0}$. It suffices to put $L := M^t$. □

Now we come to a first high-light of this chapter.

Theorem 3.6. *Let Φ be a phase function on $\Omega\times\mathbb{R}^N$, let $a\in S^m_{\varrho,\delta}(\Omega\times\mathbb{R}^n)$ with $0<\varrho\le 1$ and $0\le\delta<1$. Assume $u\in C_0^\infty(\Omega)$, and for $\Omega=\mathbb{R}^n$, this may be relaxed to $u\in\mathcal{S}(\mathbb{R}^n)$. Let $\chi\in C_0^\infty(\mathbb{R}^N)$ be a function that is identically equal to 1 in a neighbourhood of $\theta=0$. Then the limit*

$$\lim_{\varepsilon\to+0}\iint_{\Omega\times\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)}\chi(\varepsilon\theta)a(z,\theta)u(z)\,\mathrm{d}z\,\mathrm{d}\theta$$

exists.

Moreover, let L be a differential operator of the form (3.2). We assume that for each $(z_0,\theta_0)\in\Omega\times(\mathbb{R}^N\setminus 0)$, at least one of the following two conditions holds:

- $((L^t)^k e^{i\Phi})(z_0,\theta_0) = e^{i\Phi(z_0,\theta_0)}$, *for all* $k \in \mathbb{N}$,
- *the function* $(z,\theta) \mapsto a(z,\theta)u(z)$ *vanishes at* (z_0,θ_0).

Then we have, for all $k \in \mathbb{N}$ *with* $m - k\min(\varrho, 1-\delta) < -N$, *the identity*

$$\lim_{\varepsilon\to+0} \iint_{\Omega\times\mathbb{R}^N} e^{i\Phi(z,\theta)}\chi(\varepsilon\theta)a(z,\theta)u(z)\,dz\,d\theta = \iint_{\Omega\times\mathbb{R}^N} e^{i\Phi(z,\theta)} L^k(a(z,\theta)u(z))\,dz\,d\theta,$$

with the right-hand side being a converging integral in the usual sense.

The last line directly implies that the LHS does not depend on the choice of χ, and the RHS does neither depend on L nor on k.

Proof. For all $(z,\theta) \in \Omega \times (\mathbb{R}^N \setminus 0)$, we have

$$e^{i\Phi(z,\theta)}a(z,\theta)u(z) = \left((L^t)^k e^{i\Phi(z,\theta)}\right) a(z,\theta)u(z).$$

Then we obtain

$$\lim_{\varepsilon\to+0} \iint_{\Omega\times\mathbb{R}^N} e^{i\Phi(z,\theta)}\chi(\varepsilon\theta)a(z,\theta)u(z)\,dz\,d\theta = \lim_{\varepsilon\to+0} \iint_{\Omega\times\mathbb{R}^N} \left((L^t)^k e^{i\Phi(z,\theta)}\right)\chi(\varepsilon\theta)a(z,\theta)u(z)\,dz\,d\theta$$
$$= \lim_{\varepsilon\to+0} \iint_{\Omega\times\mathbb{R}^N} e^{i\Phi(z,\theta)} L^k\left(\chi(\varepsilon\theta)a(z,\theta)u(z)\right)\,dz\,d\theta. \tag{3.4}$$

Now we have $\chi \in S^0_{1,0}$, with symbol estimates that do not depend on ε. Moreover, we have $u \in S^0_{1,0}$ and $a \in S^m_{\varrho,\delta}$. Hence the function $(z,\theta) \mapsto \chi(\varepsilon\theta)a(z,\theta)u(z)$ belongs to $S^m_{\varrho,\delta}$, with symbol estimates that do not depend on ε. Then we have $\partial_{\theta_j}(\chi a u) \in S^{m-\varrho}_{\varrho,\delta}$ and $\partial_{z_\ell}(\chi a u) \in S^{m+\delta}_{\varrho,\delta}$, and it follows that $La \in S^{m-\min(\varrho,1-\delta)}_{\varrho,\delta}$. Similarly, we have $L^k a \in S^{m-k\min(\varrho,1-\delta)}_{\varrho,\delta}$. By the choice of k, the last integral in (3.4) then is absolutely integrable also without the factor $\chi(\varepsilon\theta)$, and we may apply LEBESGUE's[1] convergence theorem to conclude the proof. □

Then the final definition of an oscillating integral becomes:

Definition 3.7 (Oscillating integral). *Let the functions* Φ, a, u, χ *be given as in Theorem 3.6. Then we define the two notations*

$$\mathrm{Os}\!\!\iint_{\Omega\times\mathbb{R}^N} e^{i\Phi(z,\theta)}a(z,\theta)u(z)\,dz\,d\theta := I_\Phi(au) := \lim_{\varepsilon\to+0} \iint_{\Omega\times\mathbb{R}^N} e^{i\Phi(z,\theta)}\chi(\varepsilon\theta)a(z,\theta)u(z)\,dz\,d\theta,$$

and this concept is called oscillating integral.

Now we give examples that motivate why this concept is useful.

Example 3.8. *Take a function* $u = u(x)$ *on* $\mathbb{R}^N$ *that is smooth and grows at most polynomially. We assume that it belongs to some symbol class* $S^m_{\varrho,0}$ *for some* $m > 0$ *and some* $\varrho \in (0,1]$. *Here* θ *means* x *and* z *does not feature. Then the Fourier transform* $\hat{u}(\xi)$ *will typically be a distribution, and the notation*

$$\hat{u}(\xi) = \int_{\mathbb{R}^n_x} e^{-ix\cdot\xi}u(x)\,dx$$

shall be understood in the sense that after pairing with a test function $\varphi = \varphi(\xi)$ *from* $\mathcal{S}(\mathbb{R}^N)$, *we get an oscillating integral*

$$\mathrm{Os}\!\!\iint_{\mathbb{R}^{2n}_{x,\xi}} e^{-ix\cdot\xi}u(x)\varphi(\xi)\,d\xi\,dx. \tag{3.5}$$

Now z *has become* ξ. *The condition* $\nabla_{z,\theta}\Phi \neq 0$ *turns into* $(x,\xi) \neq 0$, *required for all* (ξ, x) *with* $x \neq 0$. *Later we will see that those* $(z,\theta) = (\xi, x)$ *become interesting for which* $\nabla_\theta\Phi = 0$, *hence* $\xi = 0$. *Recall that Fourier transforms of polynomials explode for* $\xi = 0$.

[1] HENRI LEBESGUE, 1875–1941

Example 3.9. *Let $\mathcal{A} = \mathcal{A}(x, D_x)$ be a ΨDO with pseudodifferential symbol $a(x, \xi)$, on a domain Ω. For $u \in \mathcal{D}(\Omega)$ (which we may extend outside Ω by zero-values), we then can write*

$$(\mathcal{A}u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} a(x,\xi)\hat{u}(\xi)\,\text{đ}\xi = \int_{\mathbb{R}^n_\xi} \left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} a(x,\xi) u(y)\,\mathrm{d}y \right) \text{đ}\xi$$
$$= \overset{\mathrm{Os}}{\iint_{\Omega\times\mathbb{R}^n}} e^{\mathrm{i}(x-y)\cdot\xi} a(x,\xi)u(y)\,\mathrm{d}y\,\text{đ}\xi.$$

Now z is y, and θ is ξ, and $N = n$. We have $\Phi(z,\theta) = (x-y)\cdot\xi$, and the condition $\nabla_{z,\theta} \neq 0$ for $\theta \neq 0$ becomes

$$0 \neq \nabla_{x,\xi}\Phi(x,\xi) = (\xi, x-y),$$

which obviously holds. Later we will see that those $(z,\theta) = (y,\xi)$ become interesting for which $\nabla_\theta\Phi = 0$, hence $y = x$. These are exactly the places where the Schwartz kernel A of the operator $\mathcal{A}$ becomes singular.

Example 3.10. *From (1.3) we recall integrals of the form*

$$\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x\cdot\xi - t|\xi|)} \cdot \frac{1}{2}\hat{u}_0(\xi)\,\text{đ}\xi = \int_{\mathbb{R}^n_\xi} \left(\int_{\mathbb{R}^n_y} e^{\mathrm{i}((x-y)\cdot\xi - t|\xi|)} \cdot \frac{1}{2} \cdot u_0(y)\,\mathrm{d}y \right) \text{đ}\xi$$
$$= \overset{\mathrm{Os}}{\iint_{\mathbb{R}^n_y\times\mathbb{R}^n_\xi}} e^{\mathrm{i}((x-y)\cdot\xi - t|\xi|)} \cdot \frac{1}{2} \cdot u_0(y)\,\mathrm{d}y\,\text{đ}\xi.$$

Now z is y, x is only a parameter, θ is ξ, and $N = n$. We have $\Phi(z,\theta) = (x-y)\cdot\xi - t|\xi|$, and the condition $\nabla_{z,\theta} \neq 0$ for $\theta \neq 0$ becomes

$$0 \neq \nabla_{y,\xi}\Phi(y,\xi) = \left(-\xi, (x-y) - \frac{\xi}{|\xi|}t \right), \qquad \text{for } \xi \neq 0,$$

which is always true. Later we will see that those $(z,\theta) = (y,\xi)$ become interesting for which $\nabla_\theta\Phi = 0$, which means $|x-y| = |t|$ and $\xi \parallel (x-y)$. Compare the Propositions 3.15 and 3.22.

3.2 Regularity Properties

We keep Φ and a fixed, and we let u run through $\mathcal{D}(\Omega)$. Then we obtain a map

$$u \mapsto I_\Phi(au), \qquad \mathcal{D}(\Omega) \to \mathbb{C}$$

which is obviously linear. Looking at Theorem 3.6 we see that there is a representation formula for $I_\Phi(au)$ with at most k derivatives acting upon u, for some k. Hence this map is also continuous in the topologies of $\mathcal{D}(\Omega)$ and of $\mathbb{C}$. Therefore, a distribution $A \in \mathcal{D}'(\Omega)$ exists with

$$I_\Phi(au) = \langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)}, \qquad \forall u \in \mathcal{D}(\Omega). \tag{3.6}$$

We wish to know more about sing-supp A.

Definition 3.11. *For a phase function Φ on $\Omega \times \mathbb{R}^N$, we define*

$$C_\Phi := \left\{ (z,\theta) \in \Omega \times (\mathbb{R}^N \setminus 0) \colon \nabla_\theta\Phi(z,\theta) = 0 \right\}.$$

Let π be the canonical projection[2]

$$\pi\colon \Omega \times (\mathbb{R}^N \setminus 0) \to \Omega,$$
$$\pi\colon (z,\theta) \mapsto z.$$

Then we define

$$S_\Phi := \pi C_\Phi, \qquad R_\Phi := \Omega \setminus C_\Phi.$$

[2] which “throws away the frequency variable”

We draw some conclusions: first $\nabla_\theta \Phi$ is a continuous function from $\Omega \times (\mathbb{R}^N \setminus 0) \subset \mathbb{R}^{n+N}$ into $\mathbb{R}^N$. This means that for any closed set in $\mathbb{R}^N$, its pre-image under $\nabla_\theta \Phi$ is a closed set in the induced topology on $\Omega \times (\mathbb{R}^N \setminus 0)$. But $\{0\}$ is a closed set in $\mathbb{R}^N$, and therefore C_Φ is closed in the induced topology. Second we remark that π is also continuous, and therefore S_Φ is closed in the induced topology of Ω. Hence R_Φ is open in the induced topology. Since Ω is itself open in $\mathbb{R}^n$, also R_Φ is open in $\mathbb{R}^n$.

Remark 3.12. *Next we recall the* Euler[3] differential equation*: If $p = p(y)\colon \mathbb{R}^N \to \mathbb{R}$ is positively homogeneous of degree $\alpha \in \mathbb{R}$, which means $p(ty) = t^\alpha p(y)$ for all $(t, y) \in \mathbb{R}_+ \times (\mathbb{R}^N \setminus 0)$, then*

$$y \cdot \nabla p(y) = \alpha p(y), \qquad y \in \mathbb{R}^N \setminus 0,$$

with the scalar product in $\mathbb{R}^N$ on the LHS.

By assumption, Φ is positively homogeneous in θ of order 1, which then implies $\Phi = 0$ on C_Φ.

Moreover, C_Φ is a *conic set*.

Definition 3.13 (**Conic set**)**.** *A set $C \subset \Omega \times (\mathbb{R}^N \setminus 0)$ is called* conic *if the following holds: whenever $(x, \theta) \in C$ and $t > 0$, then also $(x, t\theta) \in C$.*

Definition 3.14 (**Conic neighbourhood**)**.** *Let $C \subset \Omega \times (\mathbb{R}^N \setminus 0)$ be a conic set. A set $D \subset \Omega \times (\mathbb{R}^N \setminus 0)$ is called* conic neighbourhood of C *if the following holds:*

- *D is an open set in the induced topology of $\Omega \times (\mathbb{R}^N \setminus 0)$,*
- *$\overline{C} \subset D$, where $\overline{C}$ is the closure of C in the induced topology of $\Omega \times (\mathbb{R}^N \setminus 0)$,*
- *D is a conic set.*

Proposition 3.15. *It holds* $\operatorname{sing-supp} A \subset S_\Phi$.

For the proof it is recommended to familiarise with the proofs of Proposition 3.4 and of Theorem 3.6, because now we will have to retrace their steps, with modifications.

Proof. By the definition of the singular support, we have to show: there is a function $\tilde{A} \in C^\infty(R_\Phi)$ such that $\langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \int_\Omega \tilde{A}(z)u(z)\,\mathrm{d}z$ for all $u \in C_0^\infty(R_\Phi)$. In other words, the distribution $A \in \mathcal{D}'(\Omega)$ has a restriction $A_{|R_\Phi}$ that is a regular distribution in $\mathcal{D}'(R_\Phi)$ which is being generated by a smooth function $\tilde{A}$.

Take $u \in \mathcal{D}(\Omega)$. By definition we have

$$\langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = I_\Phi(au) = \iint_{\Omega\times\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)} L^k\Big(a(z,\theta)u(z)\Big)\,\mathrm{d}z\,\mathrm{d}\theta.$$

Now let $u \in \mathcal{D}(R_\Phi)$. Then $K := \operatorname{supp} u$ is a compact set, and $K \Subset R_\Phi$.

Next we construct L in a clever way. Whenever $a(z,\theta)u(z) \neq 0$, then $z \in K$, and therefore $\nabla_\theta \Phi(z,\theta) \neq 0$. Hence we can construct L in such a way that it only differentiates with respect to θ, and satisfies $L^t \exp(\mathrm{i}\Phi) = \exp(\mathrm{i}\Phi)$. And for z being outside K, we have $a(z,\theta)u(z) = 0$ identically (including all its derivatives), and we can choose the coefficients of L there whatever we want.

The result is then: if $u \in \mathcal{D}(R_\Phi)$ with $K := \operatorname{supp} u \Subset R_\Phi$, then the operator L can be chosen in such a way (depending on K) that

$$\langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \iint_{\Omega\times\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)} u(z) L^k\Big(a(z,\theta)\Big)\,\mathrm{d}z\,\mathrm{d}\theta.$$

The improvement is that no derivative ever arrives at u. But now we can apply Fubini's theorem,

$$\langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \int_\Omega \left(\int_{\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)} L^k a(z,\theta)\,\mathrm{d}\theta\right) u(z)\,\mathrm{d}z,$$

and it suffices to set

$$\tilde{A}(z) := \int_{\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)} L^k a(z,\theta)\,\mathrm{d}\theta, \qquad z \in K.$$

This is a smooth function, because of $a \in S^m_{\varrho,\delta}$, and the coefficients of L are C^∞. □

[3] Leonhard Euler, 1707–1783

We apply this idea once more:

Proposition 3.16. *Let $a \in S^m_{\varrho,\delta}(\Omega \times \mathbb{R}^N)$, and suppose that $a \equiv 0$ in a conic neighbourhood D_1 of C_Φ. Then $A \in \mathcal{D}'(\Omega)$ is a regular distribution that is being generated by a smooth function $\tilde{A} \in C^\infty(\Omega)$.*

This function $\tilde{A}$ is obviously unique, and it has already been determined outside C_Φ in the previous result.

Proof. We choose another conic neighbourhood D_0 in-between:

$$C_\Phi \subset D_0, \qquad \overline{D_0} \subset D_1,$$

with D_0 and D_1 being open.

Outside of D_0, we have $\nabla_\theta \Phi(x,\theta) \neq 0$. Hence we can choose there the operator L in such a way that it only differentiates with respect to θ, but not z. And inside D_1, we know that the function $(z,\theta) \mapsto a(z,\theta)u(z)$ vanishes identically, giving us the freedom to choose the coefficients of L there however we want. Again we obtain

$$\langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)} = \int_\Omega \left(\int_{\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)} L^k a(z,\theta)\,\mathrm{d}\theta \right) u(z)\,\mathrm{d}z, \qquad u \in \mathcal{D}(\Omega),$$

because no derivative ever arrives at $u(z)$. And again we put

$$\tilde{A}(z) = \int_{\mathbb{R}^N} e^{\mathrm{i}\Phi(z,\theta)} L^k a(z,\theta)\,\mathrm{d}\theta, \qquad z \in \Omega,$$

and we check the smoothness of $\tilde{A}$ quickly. □

Example: *Consider the operator* id*, which maps $u \in C_0^\infty(\Omega)$ to u. Let us tacitly extend u to the full space $\mathbb{R}^n$ by zero values. Then*

$$(\mathrm{id}\, u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} \cdot 1 \cdot \hat{u}(\xi)\,\text{đ}\xi = \int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} \cdot 1 \cdot u(y)\,\mathrm{d}y\,\text{đ}\xi.$$

Now x is just a parameter, $a = a(y,\xi) = 1 \in S^0_{1,0}(\Omega \times \mathbb{R}^n)$, $\Phi(y,\xi) = (x-y)\cdot\xi$,

$$C_\Phi = \{(y,\xi) \in \Omega \times (\mathbb{R}^n \setminus 0) \colon \nabla_\xi \Phi(y,\xi) = 0\} = \{(x,\xi) \colon \xi \in \mathbb{R}^n \setminus 0\},$$

$S_\Phi = \{x\}$, and indeed we have

$$u(x) = \langle \delta_x(y), u(y)\rangle_{\mathcal{D}'(\Omega_y)\times\mathcal{D}(\Omega)} = \langle A, u\rangle_{\mathcal{D}'(\Omega)\times\mathcal{D}(\Omega)},$$

where δ_x is the Delta distribution, shifted to the location x. Therefore, the operator id *indeed satisfies* sing-supp $A = \{x\} = S_\Phi$.

Example: *Choose $\Omega = \mathbb{R}^n$, and let f be a mollifying kernel, which means $f \in C_0^\infty(\mathbb{R}^n)$, $f(x) \geq 0$ everywhere, $f(0) > 0$, and $\int_{\mathbb{R}^n} f(x)\,\mathrm{d}x = 1$. Consider the operator*

$$P \colon u \mapsto Pu := f * u.$$

Now we have $f(z) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}z\cdot\xi} \hat{f}(\xi)\,\text{đ}\xi$, hence

$$(Pu)(x) = \int_{\mathbb{R}^n_y} f(x-y)u(y)\,\mathrm{d}y = \int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} \hat{f}(\xi)u(y)\,\mathrm{d}y\,\text{đ}\xi.$$

In this setting, x is just a parameter, $a = a(y,\xi) = \hat{f}(\xi) \in S^{-\infty}(\mathbb{R}^n \times \mathbb{R}^n)$, $\Phi(y,\xi) = (x-y)\cdot\xi$ as in the previous example, hence also $S_\Phi = \{x\}$. However, we have

$$(Pu)(x) = \langle f(x-y), u(y)\rangle_{\mathcal{D}'(\mathbb{R}^n_y)\times\mathcal{D}(\mathbb{R}^n_y)},$$

which means $\tilde{A}(y) = f(x-y)$. The result therefore is sing-supp $A = \emptyset \subset S_\Phi$*, with a strict inclusion.*

3.3 Fourier Integral Operators

We now intend to study more general operators

$$(\mathcal{A}u)(x) = \overset{\text{Os}}{\iint\limits_{\Omega_y \times \mathbb{R}^N_\theta}} e^{\mathrm{i}\Phi(x,y,\theta)} a(x,y,\theta) u(y) \, \mathrm{d}y \, \mathrm{d}\theta, \tag{3.7}$$

which are oscillating integrals with a parameter x, where

$$\begin{aligned} &x \in \Omega_x \subset \mathbb{R}^{n_x}, \qquad y \in \Omega_y \subset \mathbb{R}^{n_y}, \\ &u \in C_0^\infty(\Omega_y), \\ &a \in S^m_{\varrho,\delta}((\Omega_x \times \Omega_y) \times (\mathbb{R}^N \setminus 0)), \qquad 0 < \varrho \le 1, \quad 0 \le \delta < 1, \end{aligned}$$

and Φ is a phase function on $(\Omega_x \times \Omega_y) \times (\mathbb{R}^N \setminus 0)$.

It is not yet fully clear in which sense $\mathcal{A}u$ exists.

Definition 3.17 (**Operator phase function**). *A phase function Φ on $(\Omega_x \times \Omega_y) \times (\mathbb{R}^N \setminus 0)$ is called* Operator phase function *if the following condition holds additionally:*

$$\left. \begin{aligned} \nabla_{y,\theta}\Phi(x,y,\theta) \neq 0, \\ \nabla_{x,\theta}\Phi(x,y,\theta) \neq 0, \end{aligned} \right\} \quad \forall (x,y,\theta) \in \Omega_x \times \Omega_y \times (\mathbb{R}^N \setminus 0).$$

Proposition 3.18. *Under these assumptions, $\mathcal{A}$ is a continuous map from $\mathcal{D}(\Omega_y)$ into $\mathcal{E}(\Omega_x)$. And $\mathcal{A}$ has a continuous extension*[4] *from $\mathcal{E}'(\Omega_y)$ into $\mathcal{D}'(\Omega_x)$.*

Proof. We start from the assumption $\nabla_{y,\theta}\Phi(x,y,\theta) \neq 0$. Then the oscillating integral

$$(\mathcal{A}u)(x) = \overset{\text{Os}}{\iint\limits_{\Omega_y \times \mathbb{R}^N_\theta}} e^{\mathrm{i}\Phi(x,y,\theta)} a(x,y,\theta) u(y) \, \mathrm{d}y \, \mathrm{d}\theta, \qquad u \in \mathcal{D}(\Omega_y),$$

exists for each $x \in \Omega_x$. The derivatives $\partial_x^\alpha (\mathcal{A}u)(x)$ also exist, and they are represented by oscillating integrals of the same shape. Consequently $\mathcal{A}u \in C^\infty(\Omega_x)$, and the continuity of $\mathcal{A}$, in the topologies of $\mathcal{D}(\Omega_y)$ and $\mathcal{E}(\Omega_x)$, is easy to show. We know that there is a differential operator $L_a(\not{D}_x)$ with

$$(\mathcal{A}u)(x) = \int_{\Omega_y} \int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)} (L_a(\not{D}_x))^k \Big(a(x,y,\theta) u(y) \Big) \, \mathrm{d}y \, \mathrm{d}\theta,$$

as a converging integral, where $\not{D}_x$ means that this operator L_a differentiates with respect to perhaps y and perhaps θ, but never with respect to x.

We define another operator $\mathcal{B}$ as

$$(\mathcal{B}v)(y) := \overset{\text{Os}}{\iint\limits_{\Omega_x \times \mathbb{R}^N_\theta}} e^{\mathrm{i}\Phi(x,y,\theta)} a(x,y,\theta) v(x) \, \mathrm{d}x \, \mathrm{d}\theta, \qquad v \in \mathcal{D}(\Omega_x),$$

which exists for each $y \in \Omega_y$. Now we bring the assumption $\nabla_{x,\theta}\Phi(x,y,\theta) \neq 0$ into play, and we conclude (as above) the continuity of the map $\mathcal{B} \colon \mathcal{D}(\Omega_y) \to \mathcal{E}(\Omega_x)$, and we have the representation formula

$$(\mathcal{B}v)(y) = \int_{\Omega_x} \int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)} (L_b(\not{D}_y))^k \Big(a(x,y,\theta) v(x) \Big) \, \mathrm{d}x \, \mathrm{d}\theta,$$

with some operator L_b that never differentiates with respect to y.

[4] which we again denote by $\mathcal{A}$

For $u \in \mathcal{D}(\Omega_y)$ and $v \in \mathcal{D}(\Omega_x)$, let us pair $\mathcal{A}u \in \mathcal{E}(\Omega_x)$ with v:

$$\begin{aligned}\langle \mathcal{A}u, v\rangle_{\mathcal{D}'(\Omega_x)\times\mathcal{D}(\Omega_x)} &= \int_{\Omega_x} (\mathcal{A}u)(x)\cdot v(x)\,\mathrm{d}x \\ &= \int_{\Omega_x}\left(\int_{\Omega_y}\int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)}(L_a(\not{D}_x))^k\Big(a(x,y,\theta)u(y)\Big)\,\mathrm{d}y\,\mathrm{d}\theta\right)v(x)\,\mathrm{d}x \\ &= \int_{\Omega_x}\int_{\Omega_y}\int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)}(L_a(\not{D}_x))^k\Big(a(x,y,\theta)u(y)v(x)\Big)\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}x \\ &= \lim_{\varepsilon\to+0}\int_{\Omega_x}\int_{\Omega_y}\int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)}(L_a(\not{D}_x))^k\Big(\chi(\varepsilon\theta)a(x,y,\theta)u(y)v(x)\Big)\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}x \\ &= \lim_{\varepsilon\to+0}\int_{\Omega_x}\int_{\Omega_y}\int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)}\chi(\varepsilon\theta)a(x,y,\theta)u(y)v(x)\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}x.\end{aligned}$$

And if we pair $\mathcal{B}v \in \mathcal{E}(\Omega_y)$ with $u \in \mathcal{D}(\Omega_y)$, we get (following a highly similar calculation)

$$\langle \mathcal{B}v, u\rangle_{\mathcal{D}'(\Omega_y)\times\mathcal{D}(\Omega_y)} = \lim_{\varepsilon\to+0}\int_{\Omega_y}\int_{\Omega_x}\int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)}\chi(\varepsilon\theta)a(x,y,\theta)v(x)u(y)\,\mathrm{d}x\,\mathrm{d}\theta\,\mathrm{d}y.$$

The result therefore is

$$\langle \mathcal{A}u, v\rangle_{\mathcal{D}'(\Omega_x)\times\mathcal{D}(\Omega_x)} = \langle u, \mathcal{B}v\rangle_{\mathcal{E}'(\Omega_y)\times\mathcal{E}(\Omega_y)}, \qquad u \in \mathcal{D}(\Omega_y), \quad v \in \mathcal{D}(\Omega_x).$$

Now we extend $\mathcal{A}$ from $\mathcal{D}(\Omega_y)$ to the distribution space $\mathcal{E}'(\Omega_y)$. If $u \in \mathcal{E}'(\Omega_y)$, then we define $\mathcal{A}u \in \mathcal{D}'(\Omega_x)$ by means of

$$\langle \mathcal{A}u, v\rangle_{\mathcal{D}'(\Omega_x)\times\mathcal{D}(\Omega_x)} := \langle u, \mathcal{B}v\rangle_{\mathcal{E}'(\Omega_y)\times\mathcal{E}(\Omega_y)}, \qquad \forall v \in \mathcal{D}(\Omega_x).$$

We have to check that this is a continuous extension. First of all, it is an extension (because we regain the previous identity in case of $u \in \mathcal{D}(\Omega_y)$).

And this extension is indeed *continuous*, because: take some $u \in \mathcal{E}'(\Omega_y)$. Proposition 2.18 gives us a sequence $(u_1, u_2, \dots) \subset \mathcal{D}(\Omega_y)$ with

$$u_j \overset{\mathcal{E}'(\Omega_y)}{\longrightarrow} u, \qquad (j\to\infty),$$

and then we have (for each $v \in \mathcal{D}(\Omega_x)$)

$$\langle \mathcal{A}u_j, v\rangle_{\mathcal{D}'(\Omega_x)\times\mathcal{D}(\Omega_x)} = \langle u_j, \mathcal{B}v\rangle_{\mathcal{E}'(\Omega_y)\times\mathcal{E}(\Omega_y)} \overset{\mathbb{C}}{\longrightarrow} \langle u, \mathcal{B}v\rangle_{\mathcal{E}'(\Omega_y)\times\mathcal{E}(\Omega_y)}, \qquad (j\to\infty),$$

and the last convergence holds because of the very definition of the topology in $\mathcal{E}'(\Omega_y)$. □

Remark 3.19. *The operator $\mathcal{A}$ is a continuous map between the topological vector spaces $\mathcal{D}(\Omega_y)$ and $\mathcal{E}(\Omega_x)$. Then the transposed operator (defined in the appendix) $\mathcal{A}^t$ is a continuous map between the dual vector spaces $\mathcal{E}'(\Omega_x)$ and $\mathcal{D}'(\Omega_y)$. It turns out that $\mathcal{B}$ is simply the restriction of $\mathcal{A}^t$ to $\mathcal{D}(\Omega_x)$.*

Definition 3.20 (**Fourier Integral Operator (FIO)**). *Let Φ be a phase function on $(\Omega_x\times\Omega_y)\times(\mathbb{R}^N\setminus 0)$, $a \in S^m_{\varrho,\delta}((\Omega_x\times\Omega_y)\times\mathbb{R}^N)$ with $0<\varrho\le 1$ and $0\le\delta<1$. Then the operator $\mathcal{A}$ (and its above defined extension denoted by the same letter) is called a* Fourier integral operator (FIO).

Remark 3.21. *We mention that different amplitude functions $a = a(x,y,\theta)$ can generate the same FIO $\mathcal{A}$. Let us take $\Omega_x = \Omega_y = \mathbb{R}^n$, $\theta = \xi \in \mathbb{R}^n$, and $\Phi(x,y,\xi) = (x-y)\cdot\xi$. Choose $a(x,y,\xi)\equiv 0$, and $\mathcal{A}$ becomes the zero operator (which maps every function u to the zero function).*

On the other hand, choose $a(x,y,\xi) = \varphi(x)\psi(y)$ with $\varphi,\psi \in C_0^\infty(\mathbb{R}^n)$, but $\operatorname{supp}\varphi\cap\operatorname{supp}\psi = \emptyset$. Then

$$(\mathcal{A}u)(x) = \iint_{\mathbb{R}^n_y\times\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi}\varphi(x)\psi(y)u(y)\,\mathrm{d}y\,\text{đ}\xi = \varphi(x)\psi(x)u(x),$$

which is also equal to zero for all x.

Since $\mathcal{A}$ maps from $\mathcal{D}(\Omega_y)$ to $\mathcal{D}(\Omega_x)$, it possesses a Schwartz kernel $A \in \mathcal{D}'(\Omega_x \times \Omega_y)$ which we now (partly) calculate. By definition of the Schwartz kernel, we have, for all $u \in \mathcal{D}(\Omega_y)$ and all $v \in \mathcal{D}(\Omega_x)$,

$$\begin{aligned}\langle A, v \otimes u\rangle_{\mathcal{D}'(\Omega_x\times\Omega_y)\times\mathcal{D}(\Omega_x\times\Omega_y)} &= \langle \mathcal{A}u, v\rangle_{\mathcal{D}'(\Omega_x)\times\mathcal{D}(\Omega_x)} \\ &= \lim_{\varepsilon\to+0}\int_{\Omega_x}\int_{\Omega_y}\int_{\mathbb{R}^N_\theta} e^{\mathrm{i}\Phi(x,y,\theta)}\chi(\varepsilon\theta)a(x,y,\theta)u(y)v(x)\,\mathrm{d}y\,\mathrm{d}\theta\,\mathrm{d}x.\end{aligned}$$

Now we join the variables:

$$z := (x, y) \in \Omega_z := \Omega_x \times \Omega_y.$$

The result then is

$$\langle A, v \otimes u\rangle_{\mathcal{D}'(\Omega_x\times\Omega_y)\times\mathcal{D}(\Omega_x\times\Omega_y)} = \overset{\mathrm{Os}}{\iint\limits_{\Omega_z\times\mathbb{R}^N_\theta}} e^{\mathrm{i}\Phi(z,\theta)}a(z,\theta)(v\otimes u)(z)\,\mathrm{d}z\,\mathrm{d}\theta.$$

Proposition 3.22. *Let $\pi\colon (\Omega_x \times \Omega_y) \times (\mathbb{R}^N \setminus 0) \to (\Omega_x \times \Omega_y)$ be the canonical projection, defined as $\pi(x, y, \theta) := (x, y)$, and put*

$$\begin{aligned}C_\Phi &:= \left\{(x,y,\theta) \in \Omega_x \times \Omega_y \times (\mathbb{R}^N \setminus 0)\colon \nabla_\theta\Phi(x,y,\theta) = 0\right\}, \\ S_\Phi &:= \pi C_\Phi.\end{aligned}$$

Then the Schwartz kernel $A \in \mathcal{D}'(\Omega_x \times \Omega_y)$ coincides outside S_Φ with a smooth function, which means sing-supp $A \subset S_\Phi$. *And if $a(x, y, \theta)$ vanishes identically in a conic neighbourhood of S_Φ, then A is a smooth function everywhere in $\Omega_x \times \Omega_y$.*

Proof. This is Proposition 3.15 and Proposition 3.16. □

What is now the relation between sing-supp$(\mathcal{A}u)$ and sing-supp u ?

Definition 3.23. *Let X and Y be sets, and suppose $S \subset X \times Y$ and $K \subset Y$. Then we define*

$$S \circ K := \{x \in X\colon \exists y \in K \text{ with } (x, y) \in S\} = \bigcup_{y\in K}\{x \in X\colon (x, y) \in S\}.$$

We need some short result with an embarrassingly long proof.

Lemma 3.24. *Let $X \subset \mathbb{R}^{n_x}$, $Y \subset \mathbb{R}^{n_y}$ be open, $K \Subset Y$, and S closed in $X \times Y$. Then $S \circ K$ is closed in X.*

Proof. We show that $X \setminus (S \circ K)$ is open in X. Now take some $x_0 \in X \setminus (S \circ K)$. We intend to prove that all $x \in X$ in a certain neighbourhood of x_0 also belong to $X \setminus (S \circ K)$. Note that (since X is open and $x_0 \in X$) there is some $\varepsilon > 0$ with $B_\varepsilon(x_0) \subset X$. And since $K \Subset Y$, we can also assume $B_\varepsilon(y_0) \subset Y$ for each $y_0 \in K$. Next we observe that

$$X \setminus (S \circ K) = \bigcap_{y\in K}\Big\{x \in X\colon (x, y) \in (X \times Y) \setminus S\Big\}.$$

From $x_0 \in X \setminus (S \circ K)$ we then deduce that

$$\forall y_0 \in K\colon (x_0, y_0) \in (X \times Y) \setminus S$$

However $(X \times Y) \setminus S$ is open, since S is assumed to be closed in $X \times Y$. This yields

$$\forall y_0 \in K\colon\ \ \exists\delta \in (0, \varepsilon)\colon\ \ B_\delta(x_0, y_0) \cap (X \times Y) \subset (X \times Y) \setminus S.$$

But $\delta < \varepsilon$, and the choice of ε then gives us $B_\delta(x_0, y_0) \subset X \times Y$, and we can simplify to

$$\forall y_0 \in K\colon\ \ \exists\delta > 0\colon\ \ B_\delta(x_0, y_0) \subset (X \times Y) \setminus S.$$

We can always put a smaller square into a disk, consequently

$$\forall y_0 \in K\colon \quad \exists \delta > 0\colon \quad B_\delta(x_0) \times B_\delta(y_0) \subset (X \times Y) \setminus S.$$

Now let y_0 run through K (and keep in mind that δ will depend on y_0). The $B_\delta(y_0)$ covers y_0 and is open in Y. This gives us a covering of K by open balls. But K is compact, which tells us that a finite number of such open sets is enough to cover K. We call them $B_{\delta_1}(y_1)$, …, $B_{\delta_M}(y_M)$. Then

$$\forall y \in K\colon \quad \exists \ell \in \{1, 2, \dots, M\}\colon \quad y \in B_{\delta_\ell}(y_\ell), \quad B_{\delta_\ell}(x_0) \times B_{\delta_\ell}(y_\ell) \subset (X \times Y) \setminus S.$$

Put $\delta_0 := \min\{\delta_1, \dots, \delta_M\}$ and note that $B_{\delta_0}(x_0) \subset B_{\delta_\ell}(x_0)$ and $\{y\} \subset B_{\delta_\ell}(y_\ell)$. Then

$$\forall y \in K\colon \quad \exists \ell \in \{1, 2, \dots, M\}\colon \quad y \in B_{\delta_\ell}(y_\ell), \quad B_{\delta_0}(x_0) \times \{y\} \subset (X \times Y) \setminus S.$$

This is an awkward way of saying

$$\forall y \in K\colon \quad B_{\delta_0}(x_0) \times \{y\} \subset (X \times Y) \setminus S,$$

which can be simplified to

$$B_{\delta_0}(x_0) \times K \subset (X \times Y) \setminus S,$$

hence $B_{\delta_0}(x_0) \subset X \setminus (S \circ K)$. Consequently, $X \setminus (S \circ K)$ is indeed open. □

Proposition 3.25. *If* $u \in \mathcal{E}'(\Omega_y)$ *then* $\text{sing-supp}(\mathcal{A}u) \subset \text{sing-supp}(A) \circ \text{sing-supp}\, u$.

Proof. By its very definition, $\text{sing-supp}\, u$ is closed in Ω_y, and $\text{sing-supp}\, u \subset \text{supp}\, u \Subset \Omega_y$. This yields that $\text{sing-supp}\, u$ is compact.

We consider a sequence of compacta $(K_1, K_2, \dots)$ with

$$\Omega \Supset K_1 \Supset K_2 \Supset K_3 \Supset \dots$$

and $\text{sing-supp}\, u = \cap_{\ell=1}^\infty K_\ell$. Take functions $\psi_\ell = \psi_\ell(y) \in C_0^\infty \Omega_y$ with

$$\text{supp}\, \psi_\ell \Subset K_{l-1}, \qquad \psi_\ell \equiv 1 \text{ on } K_\ell.$$

In the sense of multiplying a distribution by a smooth function, we have

$$\mathcal{A}u = \mathcal{A}(\psi_\ell u) + \mathcal{A}((1 - \psi_\ell)u).$$

We remark $(1 - \psi_\ell)u \in C_0^\infty(\Omega_y)$, hence $\mathcal{A}((1 - \psi_\ell)u) \in C^\infty(\Omega_x)$, which implies

$$\text{sing-supp}(\mathcal{A}u) = \text{sing-supp}(\mathcal{A}(\psi_\ell u)), \qquad \ell \in \mathbb{N}.$$

Now it suffices to show that

$$\text{sing-supp}(\mathcal{A}(\psi_\ell u) \subset \text{sing-supp}(K) \circ K_{\ell-1}, \qquad \forall \ell.$$

To this end, we pick some $x_0 \in \Omega_x \setminus (\text{sing-supp}(A) \circ K_{\ell-1})$, and our intention is to prove $x_0 \not\in \text{sing-supp}(\mathcal{A}(\psi_\ell u))$. By the previous lemma, the set $\Omega_x \setminus (\text{sing-supp}(A) \circ K_{\ell-1})$ is open, hence there is some positive ε such that

$$B_\varepsilon(x_0) \subset \Omega_x \setminus (\text{sing-supp}(A) \circ K_{\ell-1}).$$

Take some $\varphi \in C_0^\infty(B_\varepsilon(x_0))$ with $\varphi \equiv 1$ on $B_{\varepsilon/2}(x_0)$. We wish to show that $\varphi \cdot \mathcal{A}(\psi_\ell u) \in C^\infty(\Omega_x)$.

But this is easy, since the operator that maps a function $v \in \mathcal{D}(\Omega_y)$ to $\varphi \cdot \mathcal{A}(\psi_\ell v)$ is a FIO with operator phase function Φ and amplitude given by

$$(x, y, \theta) \mapsto \varphi(x) a(x, y, \theta) \psi_\ell(y).$$

The last claim of Proposition 3.22 applies now. It suffices to use the first claim of Theorem 2.45. □

Example 3.26 (ΨDO). *In case of a ΨDO $\mathcal{A}$ on a domain $\Omega = \Omega_x = \Omega_y$ we have*

$$\Phi(x,y,\xi) = (x-y)\cdot\xi,$$
$$C_\Phi = \{(x,y,\xi)\colon (x-y)=0\} = \operatorname{diag}(\Omega\times\Omega)\times(\mathbb{R}^n\setminus 0),$$
$$S_\Phi = \operatorname{diag}(\Omega\times\Omega) = \{(x,x)\colon x\in\Omega\},$$

and the proposition says $\operatorname{sing-supp}(\mathcal{A}u)\subset\operatorname{sing-supp} u$.

Example 3.27 (Solution operator to the wave equation). *Take $u_1\in\mathcal{S}(\mathbb{R}^n)$, and let $u=u(t,x)$, with $(t,x)\in\mathbb{R}^{1+n}$, be the solution to*

$$u_{tt} - c^2 \triangle u = 0, \qquad u(0,x)=0, \qquad u_t(0,x) = u_1(x).$$

Then we have

$$\begin{aligned} u(t,x) &= \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\frac{\sin(ct|\xi|)}{c|\xi|}\hat{u}_1(\xi)\,đ\xi \\ &= \frac{1}{2\mathrm{i}c}\int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}((x-y)\cdot\xi+ct|\xi|)}\frac{\chi(\xi)}{|\xi|}u_1(y)\,\mathrm{d}y\,đ\xi \\ &\quad - \frac{1}{2\mathrm{i}c}\int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}((x-y)\cdot\xi-ct|\xi|)}\frac{\chi(\xi)}{|\xi|}u_1(y)\,\mathrm{d}y\,đ\xi \\ &\quad + \int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi}\frac{\sin(ct|\xi|)}{c|\xi|}(1-\chi(\xi))u_1(y)\,\mathrm{d}y\,đ\xi, \end{aligned}$$

where $\chi\in C^\infty(\mathbb{R}^n_\xi)$ is an excision function *with $\chi(\xi)=1$ for $|\xi|\geq 2$ and $\chi(\xi)=0$ for $|\xi|\leq 1$. The only purpose of this function is to avoid a division by zero in the first two integrals.*

The third integral is harmless, because we can obviously swap the integrals, and it acts as a smoothing operator. Focusing on the first two integrals we then have

$$\Phi(x,y,\xi) = (x-y)\cdot\xi \pm ct|\xi|,$$
$$a(x,y,\xi) = \frac{\chi(\xi)}{|\xi|}\in S^{-1}_{1,0}(\mathbb{R}^n\times\mathbb{R}^n),$$
$$\nabla_\xi\Phi(x,y,\xi) = (x-y)\pm ct\frac{\xi}{|\xi|},$$
$$S_\Phi = \{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n\colon |x-y| = c|t|\}.$$

Now $|x-y|$ is the distance between x and y, and $c|t|$ is the length of the path which you travel in a time interval of length $|t|$ at speed c.

The relation $\operatorname{sing-supp}(\mathcal{A}u_1)\subset S_\Phi\circ\operatorname{sing-supp}(u_1)$ *then tells us that singularities of u_1 are being propagated with speed c.*

Chapter 4

Symbol Calculus

Let us be given a ΨDO $\mathcal{A}$ with amplitude function $a \in S^m$,

$$(\mathcal{A}u)(x) = \overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^{2n}_{y\times\xi}}} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi) u(y)\,\mathrm{d}y\,\text{đ}\xi, \qquad u \in C_0^\infty(\mathbb{R}^n). \tag{4.1}$$

According to Proposition 3.18, $\mathcal{A}$ is then a continuous operator from $\mathcal{D}(\mathbb{R}^n)$ into $\mathcal{E}(\mathbb{R}^n)$, with a continuous extension as an operator from $\mathcal{E}'(\mathbb{R}^n)$ into $\mathcal{D}'(\mathbb{R}^n)$.

Our questions are:

- Is it possible to "get rid of" the argument y in the amplitude function ? Then we could perform the integration over y and obtain the easier formula

  $$(\mathcal{A}u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} p(x,\xi)\hat{u}(\xi)\,\text{đ}\xi, \qquad u \in C_0^\infty(\mathbb{R}^n),$$

  assuming that such a function p existed. This integral makes sense classically, because of the fast decay of $\hat{u}(\xi)$. We wish to have a formula how to build p from a. Such a function p then is called *symbol of* $\mathcal{A}$, and we write $p = \sigma_{\mathcal{A}}$, and also $p = \sigma(\mathcal{A})$.

- What can be said about the transposed operator $\mathcal{A}^t$ and its symbol $\sigma_{\mathcal{A}^t}$?

- Let $\mathcal{B}$ be one more ΨDO. What can be said about $\mathcal{A} \circ \mathcal{B}$ (even its existence is unclear, considering the above mapping properties) ?

- Under which assumptions does there exist an inverse operator $\mathcal{A}^{-1}$, and how to find it ? If we don't find it, can someone give us at least a "consolation prize" ?

In this chapter, we will determine the symbols of the mentioned operators, sometimes only up to a remainder from $S^{-\infty}(\Omega \times \Omega \times \mathbb{R}^n)$.

To simplify notation, we introduce one more class:

Definition 4.1. *Let $a = a(x,y,\xi) \in S^m_{\varrho,\delta}(\Omega\times\Omega\times\mathbb{R}^n)$ be an amplitude function. We say that its associated operator $\mathcal{A}$ belongs to the class $\Psi^m_{\varrho,\delta}(\Omega)$. In case of $(\varrho,\delta) = (1,0)$ we simply write Ψ^m instead of $\Psi^m_{1,0}$. We set $\Psi^{-\infty} = \bigcap_{m\in\mathbb{R}} \Psi^m_{\varrho,\delta}$, and this intersection is independent of $\varrho,\delta \in [0,1]$. Sometimes we write $\Psi^\infty = \bigcup_{m\in\mathbb{R}} \Psi^m_{\varrho,\delta}$.*

4.1 Constructing a Symbol from an Amplitude

Consider $a \in S^m_{1,0}(\mathbb{R}^n_x \times \mathbb{R}^n_y \times \mathbb{R}^n_\xi)$ with $m \in \mathbb{R}$, and let

$$(\mathcal{A}u)(x) = \overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^{2n}_{y,\xi}}} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi) u(y)\,\mathrm{d}y\,\text{đ}\xi, \qquad u \in \mathcal{S}(\mathbb{R}^n).$$

We could also take $S^m_{\varrho,\delta}$ with $0 \le \delta < \varrho \le 1$, and the following calculations would basically not change (except their length). Similarly, we could substitute $\mathbb{R}^n_x$ against $\Omega_x \subset \mathbb{R}^n$.

We are looking for a $p(x,\xi) = \sigma_{\mathcal{A}}(x,\xi)$ such that (in the sense of an iterated integral), we have

$$(\mathcal{A}u)(x) = \int_{\mathbb{R}^n_\xi} \int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} p(x,\xi) u(y) \,\mathrm{d}y \,\text{đ}\xi.$$

Several ideas now have to come together. First of all, we can ignore those y that are far away from x, in the following sense. Take $\chi \in C_0^\infty(\mathbb{R}^n)$ with $\chi(s) \equiv 1$ for $|s| \le \delta$ (choose some positive δ). Then we can write

$$\begin{aligned}(\mathcal{A}u)(x) = &\overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^{2n}_{y,\xi}}} e^{\mathrm{i}(x-y)\cdot\xi} \chi(x-y) a(x,y,\xi) u(y) \,\mathrm{d}y \,\text{đ}\xi \\ &+ \overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^{2n}_{y,\xi}}} e^{\mathrm{i}(x-y)\cdot\xi} (1-\chi(x-y)) a(x,y,\xi) u(y) \,\mathrm{d}y \,\text{đ}\xi.\end{aligned}$$

According to Proposition 3.16, the second integral represents an operator with a smooth Schwartz kernel, hence it is a smoothing operator, *i.e.* an operator which maps from $\mathcal{E}'(\mathbb{R}^n)$ into $\mathcal{E}(\mathbb{R}^n)$. Such operators will be ignored in the following.

Second, we have a look at the Schwartz kernel A of $\mathcal{A}$. As suggested in the motivation chapter, we expect the formula

$$A(x,y) = \left(\mathcal{F}^{-1}_{\xi\to z} a(x,y,\xi)\right)\Big|_{z=x-y}$$

to hold, in distributional sense. Now A has its singular support on the diagonal of $\Omega \times \Omega$, hence for $x = y$. If $\mathcal{A}$ is a zeroth order PDO, then $A(x,y) \sim \delta_x(y)$, the Delta distribution acting with respect to y, shifted at location x. And if $\mathcal{A}$ is a first order PDO, then A will carry derivatives of the Delta distributions on the diagonal. The higher the order of $\mathcal{A}$ gets, the stronger the singularities of A on the diagonal of $\Omega \times \Omega$ become.

Third, we have the representation for $\mathcal{A}$, and we may assume $|y - x| < \delta$. Hence $y \approx x$, and we wish to get rid of y, hence a Taylor expansion seems reasonable:

$$a(x,y,\xi) = \sum_{|\alpha|\le M} \frac{1}{\alpha!} (\partial_z^\alpha a(x,z,\xi))\Big|_{z=x} (y-x)^\alpha + r_M(x,y,\xi),$$

and each power of $(y - x)$ will "soften" the singularity of A on the diagonal $\{(x,y) \in \Omega \times \Omega \colon y = x\}$. Then we can hope for the expansion

$$p(x,\xi) = a(x,x,\xi) + \mathcal{O}(S^{m-1}) + \mathcal{O}(S^{m-2}) + \dots + \mathcal{O}(S^{m-M}) + ???,$$

where each $\mathcal{O}(S^{m-j})$ stands for some symbol which seems to belong to S^{m-j}, and it is easily computable. The last item (???) comes from r_M, and it will be a bit more laborious.

Now the actual work begins. We assume that the amplitude function $a = a(x,y,\xi)$ satisfies

$$a(x,y,\xi) = 0 \qquad \text{if } |x-y| > \delta,$$

for some positive δ of our choice. In Proposition 3.15, the Schwartz kernel A has been calculated outside of S_Φ (which is the diagonal in our case), and the result then is $A(x,y) = 0$ for $|x-y| > \delta$. Such operators $\mathcal{A}$ are called *properly supported*, and they are interesting because the value $(\mathcal{A}u)(x)$ only depends on those values $u(y)$ for which $|x - y| \le \delta$. In particular, functions u with compact support will be mapped to functions $\mathcal{A}u$ with compact support. We then even have

$$\mathcal{A} \colon \mathcal{D}(\mathbb{R}^n) \to \mathcal{D}(\mathbb{R}^n), \qquad \mathcal{A} \colon \mathcal{E}(\mathbb{R}^n) \to \mathcal{E}(\mathbb{R}^n),$$

with continuity in the respective topologies.

Now let $u \in \mathcal{S}(\mathbb{R}^n)$ with compact support of its Fourier transform $\hat{u}$. Such functions are a dense set in $\mathcal{S}(\mathbb{R}^n)$. Then u is even analytic, and we have

$$u(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\hat{u}(\xi)\,đ\xi.$$

Let us write $e_\xi(x) := e^{\mathrm{i}x\cdot\xi}$. For each ξ, we have $e_\xi \in \mathcal{E}(\mathbb{R}^n_x)$, and this collection of functions depends continuously (in the sense of the topology of $\mathcal{E}(\mathbb{R}^n_x)$) on the parameter ξ. And $\mathcal{A}$ is now continuous as a map of $\mathcal{E}(\mathbb{R}^n_x)$ into itself, hence we have[1]

$$(\mathcal{A}u)(x) = \left(\mathcal{A}\int_{\mathbb{R}^n_\xi} e_\xi\hat{u}(\xi)\,đ\xi\right)(x) = \int_{\mathbb{R}^n_\xi} (\mathcal{A}e_\xi)(x)\cdot\hat{u}(\xi)\,đ\xi.$$

Then we obtain

$$(\mathcal{A}u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi}\cdot(e_{-\xi}(x)\cdot(\mathcal{A}e_\xi)(x))\cdot\hat{u}(\xi)\,đ\xi, \qquad u\in\mathcal{S}(\mathbb{R}^n), \quad \operatorname{supp}\hat{u}\Subset\mathbb{R}^n_\xi,$$

and hence we may set

$$p(x,\xi) := e^{-\mathrm{i}x\cdot\xi}\cdot(\mathcal{A}e_\xi)(x), \qquad (x,\xi)\in\mathbb{R}^{2n}.$$

Due to the properties of $\mathcal{A}$ this is, for each fixed ξ, a function from $\mathcal{E}(\mathbb{R}^n_x)$.

By the definition of $\mathcal{A}$ (and since $\mathcal{A}$ is properly supported), we then have

$$\begin{aligned} p(x,\xi) &= e^{-\mathrm{i}x\cdot\xi}\ \overset{\mathrm{Os}}{\iint_{\mathbb{R}^{2n}_{y,\theta}}} e^{\mathrm{i}(x-y)\cdot\theta}a(x,y,\theta)e^{\mathrm{i}y\xi}\,\mathrm{d}y\,đ\theta \\ &= \overset{\mathrm{Os}}{\iint_{\mathbb{R}^{2n}_{y,\theta}}} e^{\mathrm{i}(x-y)\cdot(\theta-\xi)}a(x,y,\theta)\,\mathrm{d}y\,đ\theta \\ &= \lim_{\varepsilon\to+0}\iint_{\mathbb{R}^{2n}_{y,\theta}} e^{\mathrm{i}(x-y)\cdot(\theta-\xi)}\chi(\varepsilon\theta)a(x,y,\theta)\,\mathrm{d}y\,đ\theta. \end{aligned}$$

Keep in mind that the integration over y is being performed over a bounded domain anyway. We introduce $z := y-x$ and get

$$p(x,\xi) = \lim_{\varepsilon\to+0}\iint_{\mathbb{R}^{2n}_{z,\theta}} e^{-\mathrm{i}z\cdot(\theta-\xi)}\chi(\varepsilon\theta)a(x,x+z,\theta)\,\mathrm{d}z\,đ\theta.$$

---

[1] The justification why $\mathcal{A}$ can be dragged into the integral is like this:
we have $u\in\mathcal{S}(\mathbb{R}^n)$ and $\operatorname{supp}\hat{u}\subset[-R,R]^n$. Put $U_x(\xi) := \exp(\mathrm{i}x\cdot\xi)\hat{u}(\xi)$. Then $u(x) = \int_{[-R,R]^n} U_x(\xi)\,đ\xi$. For a continuous function f on $[-R,R]^n$, let $S_J\{f\}$ be a numerical quadrature formula with J nodes $\{\xi_1,\dots,\xi_J\}$ as an approximation for $\int_{[-R,R]^n} f(\xi)\,đ\xi$. This means $S_Jf = \sum_{j=1}^J \omega_j f(\xi_j)$. We then have positive constants C, ε and p such that the inequality

$$\left|\int_{[-R,R]^n} f(\xi)\,đ\xi - S_J\{f\}\right| \le CJ^{-\varepsilon}\,\|f\|_{C^p}$$

holds for all C^p functions f. Now we have, for all α, that $D^\alpha_x u(x) = \lim_{J\to\infty} S_J\{D^\alpha_x U_x\}$, with a convergence that is locally uniform with respect to x. Here we have used $\xi\in[-R,R]^n$. This then implies $u(x) = \mathcal{E}\text{–}\lim_{J\to\infty} S_J\{U_x\}$. Now $\mathcal{A}\colon\mathcal{E}\to\mathcal{E}$ is continuous, which means (by definition)

$$(\mathcal{A}u)(x) = \mathcal{A}(\lim_{J\to\infty} S_J\{U_x\}) = \lim_{J\to\infty}(\mathcal{A}S_J\{U_x\}).$$

Precall now $p(x,\xi) := (\mathcal{A}e_\xi)e^{-\mathrm{i}x\cdot\xi}$. To calculate $\mathcal{A}S_J\{U_x\}$, we remark that $\mathcal{A}U_x(\xi) = \mathcal{A}(e_\xi\hat{u}(\xi)) = e^{\mathrm{i}x\cdot\xi}p(x,\xi)\hat{u}(\xi) = p(x,\xi)U_x(\xi)$, and therefore

$$\mathcal{A}S_J\{U_x\} = \mathcal{A}\left(\sum_{j=1}^J \omega_j U_x(\xi_j)\right) = \sum_{j=1}^J \omega_j\mathcal{A}U_x(\xi_j) = \sum_{j=1}^J \omega_j p(x,\xi_j)U_x(\xi_j) = S_J\{p(x,\cdot)U_x\},$$

which then allows to resume the above calculation,

$$\lim_{J\to\infty}(\mathcal{A}S_J\{U_x\}) = \lim_{J\to\infty} S_J\{p(x,\cdot)U_x\} = \int_{[-R,R]^n} p(x,\xi)U_x(\xi)\,đ\xi = \int_{\mathbb{R}^n} e^{\mathrm{i}x\cdot\xi}p(x,\xi)\hat{u}(\xi)\,đ\xi.$$

Now we regularise once more:

$$p(x,\xi) = \lim_{\varepsilon\to+0}\left(\lim_{\varepsilon_0\to+0}\iint_{\mathbb{R}^{2n}_{z,\theta}} e^{-\mathrm{i}z\cdot(\theta-\xi)}\chi(\varepsilon\theta)\chi(\varepsilon_0 z)a(x,x+z,\theta)\,\mathrm{d}z\,\text{đ}\theta\right).$$

By Taylor expansion, we have, for some M of our choice,

$$a(x,x+z,\theta) = \sum_{|\alpha|\le M}\frac{1}{\alpha!}\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\cdot z^\alpha + \sum_{|\alpha|=M+1} r_\alpha(x,z,\theta)\cdot z^\alpha,$$

with $r_\alpha\in S^m$, in particular a smooth function of x and z.

We plug this in and make use of $\exp(-\mathrm{i}z\cdot(\theta-\xi))z^\alpha = (-D_\theta)^\alpha\exp(-\mathrm{i}z\cdot(\theta-\xi))$:

$$\begin{aligned}p(x,\xi) = \sum_{|\alpha|\le M}\frac{1}{\alpha!}\lim_{\varepsilon\to+0}&\left(\lim_{\varepsilon_0\to+0}\iint_{\mathbb{R}^{2n}_{z,\theta}}\Big((-D_\theta)^\alpha e^{-\mathrm{i}z\cdot(\theta-\xi)}\Big)\cdot\chi(\varepsilon\theta)\chi(\varepsilon_0 z)\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\mathrm{d}z\,\text{đ}\theta\right)\\ &+\sum_{|\alpha|=M+1}\lim_{\varepsilon\to+0}\left(\lim_{\varepsilon_0\to+0}\iint_{\mathbb{R}^{2n}_{z,\theta}}\Big((-D_\theta)^\alpha e^{-\mathrm{i}z\cdot(\theta-\xi)}\Big)\cdot\chi(\varepsilon\theta)\chi(\varepsilon_0 z)R_\alpha(x,z,\theta)\,\mathrm{d}z\,\text{đ}\theta\right).\end{aligned}$$

Observe that the functions $\theta\mapsto\chi(\varepsilon\theta)$ and $z\mapsto\chi(\varepsilon_0 z)$ have Fourier transforms

$$t\mapsto\varepsilon^{-n}\hat\chi(t/\varepsilon)\quad\text{and}\quad\zeta\mapsto\varepsilon_0^{-n}\hat\chi(\zeta/\varepsilon_0),$$

which converge (in $\mathcal{D}'(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$) to the Delta distribution, times $(2\pi)^n$.

By partial integration we then have

$$\begin{aligned}&\iint_{\mathbb{R}^{2n}_{z,\theta}}\Big((-D_\theta)^\alpha e^{-\mathrm{i}z\cdot(\theta-\xi)}\Big)\cdot\chi(\varepsilon\theta)\chi(\varepsilon_0 z)\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\mathrm{d}z\,\text{đ}\theta\\ &=\iint_{\mathbb{R}^{2n}_{z,\theta}} e^{-\mathrm{i}z\cdot(\theta-\xi)}\cdot D_\theta^\alpha\left(\chi(\varepsilon\theta)\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\right)\cdot\chi(\varepsilon_0 z)\,\mathrm{d}z\,\text{đ}\theta\\ &=\int_{\mathbb{R}^n_\theta}\left(\int_{\mathbb{R}^n_z} e^{-\mathrm{i}z\cdot(\theta-\xi)}\cdot\chi(\varepsilon_0 z)\,\mathrm{d}z\right)D_\theta^\alpha\left(\chi(\varepsilon\theta)\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\right)\text{đ}\theta\\ &=\int_{\mathbb{R}^n_\theta}\frac{1}{\varepsilon_0^n}\hat\chi\left(\frac{\theta-\xi}{\varepsilon_0}\right)\cdot D_\theta^\alpha\left(\chi(\varepsilon\theta)\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\right)\text{đ}\theta.\end{aligned}$$

We clearly may assume that $\hat\chi$ is even. Then we can send ε_0 to zero and get

$$\begin{aligned}&\lim_{\varepsilon_0\to+0}\iint_{\mathbb{R}^{2n}_{z,\theta}}\Big((-D_\theta)^\alpha e^{-\mathrm{i}z\cdot(\theta-\xi)}\Big)\cdot\chi(\varepsilon\theta)\chi(\varepsilon_0 z)\Big(\partial_y^\alpha a(x,y,\theta)\Big)\Big|_{y=x}\mathrm{d}z\,\text{đ}\theta\\ &=D_\xi^\alpha\left(\chi(\varepsilon\xi)\Big(\partial_y^\alpha a(x,y,\xi)\Big)\Big|_{y=x}\right).\end{aligned}$$

Sending now ε to zero in this item gives us $(D_\xi^\alpha\partial_y^\alpha a(x,y,\xi))_{|y=x}$.

Concerning the remainder term, we first remark that $R(x,z,\theta)$ will in general not be equal to zero for large $|z|$. But we can choose some large $k\in\mathbb{N}$ and observe that

$$\langle D_\theta\rangle^{2k}e^{-\mathrm{i}z\cdot(\theta-\xi)} = (1-\triangle_\theta)^k e^{-\mathrm{i}z\cdot(\theta-\xi)} = (1+|z|^2)^k e^{-\mathrm{i}z\cdot(\theta-\xi)},$$

which enables us to write

$$\begin{aligned}&\iint_{\mathbb{R}^{2n}_{z,\theta}}\Big((-D_\theta)^\alpha e^{-\mathrm{i}z\cdot(\theta-\xi)}\Big)\cdot\chi(\varepsilon\theta)\chi(\varepsilon_0 z)r_\alpha(x,z,\theta)\,\mathrm{d}z\,\text{đ}\theta\\ &=\iint_{\mathbb{R}^{2n}_{z,\theta}} e^{-\mathrm{i}z\cdot(\theta-\xi)}\frac{\chi(\varepsilon_0 z)}{\langle z\rangle^{2k}}\cdot\langle D_\theta\rangle^{2k}D_\theta^\alpha\Big(\chi(\varepsilon\theta)r_\alpha(x,z,\theta)\Big)\,\mathrm{d}z\,\text{đ}\theta.\end{aligned}$$

Now we can assume that k is so large that $2k>n$ and $2k+(M+1)-m>n$ to obtain an integral in which the limits $\varepsilon_0\to 0$ and $\varepsilon\to 0$ are easy to perform.

The final result then is

$$p(x,\xi) = \sum_{|\alpha|\le M} \frac{1}{\alpha!} \left(D_y^\alpha \partial_\xi^\alpha a(x,y,\xi)\right)\big|_{y=x} + \text{(Remainder)},$$

and the remainder (which does not seem to have a nice formula) indeed belongs to $S^{m-M-1}_{1,0}$, which we do not verify.

This way we can prove:

Proposition 4.2. *Let $a \in S^m_{1,0}(\mathbb{R}^n_x \times \mathbb{R}^n_y \times \mathbb{R}^n_\xi)$ with $a(x,y,\xi) = 0$ for $|x-y| \ge \delta > 0$, and let $\mathcal{A}$ be the associated ΨDO. Then this operator has a pseudodifferential symbol $p = \sigma_{\mathcal{A}} \in S^m_{1,0}(\mathbb{R}^n_x \times \mathbb{R}^n_\xi)$ given as*

$$p(x,\xi) = e^{-\mathrm{i}x\cdot\xi}(\mathcal{A}e_\xi)(x), \qquad (x,\xi) \in \mathbb{R}^{2n},$$

and we have the asymptotic expansion

$$p(x,\xi) \sim \sum_{j=0}^\infty \frac{1}{j!} \left(\langle \nabla_\xi, D_y\rangle^j \, a(x,y,\xi)\right)\Big|_{y=x},$$

to be understood in the sense that, for each $k \in \mathbb{N}_+$, we have $p - \sum_{j=0}^{k-1} \ldots \in S^{m-k}_{1,0}$. We may also write

$$p(x,\xi) \sim \exp(\langle \nabla_\xi, D_y\rangle) a(x,y,\xi)\big|_{y=x} \qquad \text{or} \qquad p(x,\xi) \sim \sum_{\alpha\in\mathbb{N}^n} \frac{1}{\alpha!} \left(\partial_\xi^\alpha D_y^\alpha a(x,y,\xi)\right)\big|_{y=x}.$$

4.2 Transposed Operators, Compositions, and Adjoints

Let $\mathcal{P}$ be a ΨDO with symbol $p \in S^m_{1,0}$, where $m \in \mathbb{R}$, hence

$$(\mathcal{P}u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} p(x,\xi)\hat{u}(\xi)\,\text{đ}\xi = \overset{\mathrm{Os}}{\iint_{\mathbb{R}^{2n}_{y,\xi}}} e^{\mathrm{i}(x-y)\cdot\xi} p(x,\xi) u(y)\,\mathrm{d}y\,\text{đ}\xi, \qquad u \in C_0^\infty(\mathbb{R}^n).$$

We do not assume $\mathcal{P}$ to be properly supported. By definition, the transposed operator $\mathcal{P}^t$ satisfies $\langle \mathcal{P}^t u, v\rangle = \langle u, \mathcal{P}v\rangle$, for $u, v \in C_0^\infty(\mathbb{R}^n)$, and $\langle f, g\rangle = \int_{\mathbb{R}^n} f(x)g(x)\,\mathrm{d}x$. The proof of Proposition 3.18 contains the representation

$$(\mathcal{P}^t u)(x) = \overset{\mathrm{Os}}{\iint_{\mathbb{R}^{2n}_{y,\xi}}} e^{-\mathrm{i}(x-y)\cdot\xi} p(y,\xi) u(y)\,\mathrm{d}y\,\text{đ}\xi.$$

To repair the sign in the phase function, we substitute $\xi \mapsto -\xi$, and this reveals the amplitude function for the operator $\mathcal{P}^t$ as

$$a(x,y,\xi) = p(y,-\xi).$$

Hence we have shown:

Proposition 4.3. *The symbol $\sigma_{\mathcal{P}^t}$ of the transposed operator $\mathcal{P}^t$ exists $\mod S^{-\infty}(\mathbb{R}^n_x \times \mathbb{R}^n_y \times \mathbb{R}^n_\xi)$, it belongs to $S^m_{1,0}$, too, and it is given by*

$$\sigma_{\mathcal{P}^t}(x,\xi) \sim \Big(\exp(-\langle\nabla_\xi, D_x\rangle)p\Big)(x,-\xi) \qquad \text{or} \qquad \sigma_{\mathcal{P}^t}(x,\xi) \sim \sum_\alpha \frac{1}{\alpha!}\partial_\xi^\alpha D_x^\alpha \Big(p(x,-\xi)\Big).$$

Now let us be given two ΨDOs $\mathcal{P}$ and $\mathcal{Q}$, and we ask for the composed operator $\mathcal{P}\circ\mathcal{Q}$. This is typically not definable, except additional assumptions hold:

- at least one of the operators has a Schwartz kernel that is properly supported (then also its transposed operator is properly supported),

- we operate in $\mathbb{R}^n$, and both operators have global symbol estimates. Then each of these operators maps from $\mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}(\mathbb{R}^n)$.

The first condition can be generalised to bounded domains Ω instead of $\mathbb{R}^n$ (see next section), but not the second.

Let the symbols of $\mathcal{P}$ and $\mathcal{Q}$ be p and q. To make things easier, we assume $\mathcal{Q}$ as properly supported (the other cases are nice homework problems). Then $u \in \mathcal{S}(\mathbb{R}^n)$ implies $\mathcal{Q}u \in \mathcal{S}(\mathbb{R}^n)$, hence

$$(\mathcal{P} \circ \mathcal{Q}u)(x) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} p(x,\xi) \cdot (\mathcal{Q}u)\hat{}(\xi)\, đ\xi,$$

and now we are interested in evaluating $(\mathcal{Q}u)\hat{}(\xi)$ without making a mess. To this end, we recall (2.4),

$$\langle (\mathcal{Q}u)\hat{}, v\rangle = \langle \mathcal{Q}u, \hat{v}\rangle = \left\langle u, \mathcal{Q}^t \hat{v}\right\rangle,$$

and therefore

$$\begin{aligned}
\int_{\mathbb{R}^n_\xi} (\mathcal{Q}u)\hat{}(\xi) \cdot v(\xi)\, \mathrm{d}\xi &= \int_{\mathbb{R}^n_x} u(x) \cdot (\mathcal{Q}^t \hat{v})(x)\, \mathrm{d}x \\
&= \int_{\mathbb{R}^n_x} u(x) \left(\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} \sigma_{\mathcal{Q}^t}(x,\xi) \hat{\hat{v}}(\xi)\, đ\xi \right) \mathrm{d}x && \Big|\ \text{now use } \hat{\hat{f}}(\xi) = (2\pi)^n f(-\xi) \\
&= \int_{\mathbb{R}^n_x} u(x) \left(\int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} \sigma_{\mathcal{Q}^t}(x,\xi) v(-\xi)\, \mathrm{d}\xi \right) \mathrm{d}x && \Big|\ \text{now substitute } \xi \to -\xi \\
&= \int_{\mathbb{R}^n_x} u(x) \int_{\mathbb{R}^n_\xi} e^{-\mathrm{i}x\cdot\xi} \sigma_{\mathcal{Q}^t}(x,-\xi) v(\xi)\, \mathrm{d}\xi\, \mathrm{d}x \\
&= \int_{\mathbb{R}^n_\xi} \left(\int_{\mathbb{R}^n_x} e^{-\mathrm{i}x\cdot\xi} \sigma_{\mathcal{Q}^t}(x,-\xi) u(x)\, \mathrm{d}x \right) v(\xi)\, \mathrm{d}\xi,
\end{aligned}$$

and after renaming $x \to y$ we therefore have (for properly supported $\mathcal{Q}$)

$$(\mathcal{Q}u)\hat{}(\xi) = \int_{\mathbb{R}^n_y} e^{-\mathrm{i}y\cdot\xi} \sigma_{\mathcal{Q}^t}(y,-\xi) u(y)\, \mathrm{d}y.$$

Now we introduce the notation $q^t := \sigma_{\mathcal{Q}^t}$, and the result is

$$\begin{aligned}
(\mathcal{P} \circ \mathcal{Q}u)(x) &= \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}x\cdot\xi} p(x,\xi) (\mathcal{Q}u)\hat{}(\xi)\, đ\xi \\
&= \int_{\mathbb{R}^n_\xi} \int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} p(x,\xi) q^t(y,-\xi) u(y)\, \mathrm{d}y\, đ\xi,
\end{aligned}$$

hence the amplitude function of $\mathcal{P} \circ \mathcal{Q}$ is $a(x,y,\xi) = p(x,\xi) q^t(y,-\xi)$, and its symbol then becomes

$$\begin{aligned}
(\sigma_{\mathcal{P}\circ\mathcal{Q}})(x,\xi) &\sim \sum_\alpha \frac{1}{\alpha!} \partial_\xi^\alpha D_y^\alpha a(x,y,\xi)\Big|_{y=x} \\
&= \exp\left(\langle \nabla_\xi, D_y \rangle\right) a(x,y,\xi)\Big|_{y=x} \\
&= \exp\left(\langle \nabla_\xi, D_y \rangle\right) \left(p(x,\xi) q^t(y,-\xi)\right)\Big|_{y=x} && \Big|\ \text{apply Leibniz formula / product rule} \\
&= \exp\left(\langle \nabla_\xi + \nabla_\eta, D_y \rangle\right) \left(p(x,\xi) q^t(y,-\eta)\right)\Big|_{y=x,\eta=\xi} \\
&= \exp\left(\langle \nabla_\xi + \nabla_\eta, D_y \rangle\right) \Big(p(x,\xi) \Big(\exp(-\langle \nabla_\eta, D_y \rangle) q\Big)(y,+\eta)\Big)\Big|_{y=x,\eta=\xi} \\
&= \exp\left(\langle \nabla_\xi + \nabla_\eta, D_y \rangle\right) \exp(-\langle \nabla_\eta, D_y \rangle) \Big(p(x,\xi) q(y,\eta)\Big)\Big|_{y=x,\eta=\xi} && \Big|\ \text{apply magic} \\
&= \exp\left(\langle \nabla_\xi, D_y \rangle\right) \Big(p(x,\xi) q(y,\eta)\Big)\Big|_{y=x,\eta=\xi} \\
&= \sum_{\alpha\in\mathbb{N}^n} \frac{1}{\alpha!} \left(\partial_\xi^\alpha p(x,\xi)\right) \left(D_x^\alpha q(x,\xi)\right).
\end{aligned}$$

This confirms the calculation (1.10) for PDOs. Hence we have shown:

Proposition 4.4. *Let $\mathcal{P}$ and $\mathcal{Q}$ be operators of the classes $\Psi^{m_p}_{1,0}(\mathbb{R}^n)$ and $\Psi^{m_q}_{1,0}(\mathbb{R}^n)$, and assume both to be properly supported. Then the composition $\mathcal{P}\circ\mathcal{Q}$ exists, it is a member of $\Psi^{m_p+m_q}_{1,0}(\mathbb{R}^n)$, and its symbol has the asymptotic expansion*

$$\sigma_{\mathcal{P}\circ\mathcal{Q}}(x,\xi) \sim \sum_{\alpha\in\mathbb{N}^n} \frac{1}{\alpha!}\left(\partial_\xi^\alpha p(x,\xi)\right)\left(D_x^\alpha q(x,\xi)\right).$$

Finally we have a look at the $L^2(\mathbb{R}^n)$ adjoint operator $\mathcal{A}^*$, which satisfies (by definition)

$$\langle \mathcal{A}u, v\rangle_{L^2} = \langle u, \mathcal{A}^* v\rangle_{L^2},$$

in the sense of

$$\int_{\mathbb{R}^n_x} (\mathcal{A}u)(x)\overline{v(x)}\,\mathrm{d}x = \int_{\mathbb{R}^n} u(x)\overline{(\mathcal{A}^*v)(x)}\,\mathrm{d}x, \qquad \forall u,v\in C_0^\infty(\mathbb{R}^n).$$

If now $\mathcal{A}$ is given by (4.1) (as always), and we suppose for sake of brevity $a\in S^m_{1,0}$ with $m<-n$, then all integral pairs can be swapped by appeal to Fubini, and it follows that

$$\begin{aligned}
\langle \mathcal{A}u, v\rangle_{L^2} &= \int_{\mathbb{R}^n_x}\int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)u(y)\overline{v(x)}\,\mathrm{d}y\,\text{đ}\xi\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n_y} u(y)\left(\int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_x} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)\overline{v(x)}\,\mathrm{d}x\,\text{đ}\xi\right)\mathrm{d}y \\
&= \int_{\mathbb{R}^n_y} u(y)\overline{\left(\int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_x} e^{\mathrm{i}(y-x)\cdot\xi}\overline{a(x,y,\xi)}v(x)\,\mathrm{d}x\,\text{đ}\xi\right)}\mathrm{d}y, \\
&= \int_{\mathbb{R}^n_y} u(y)\overline{(\mathcal{A}^*v)(y)}\,\mathrm{d}y,
\end{aligned}$$

and the consequence then is

$$(\mathcal{A}^*v)(x) = \int_{\mathbb{R}^n_\xi}\int_{\mathbb{R}^n_y} e^{\mathrm{i}(x-y)\cdot\xi}\overline{a(y,x,\xi)}v(y)\,\mathrm{d}y\,\text{đ}\xi.$$

The result is: if the $\mathcal{A}$ has the amplitude function $(x,y,\xi)\mapsto a(x,y,\xi)$, then its L^2 adjoint operator $\mathcal{A}^*$ has the amplitude function $(x,y,\xi)\mapsto\overline{a(y,x,\xi)}$.

Expressed in the language of symbols: if $\mathcal{A}$ has the symbol $p(x,\xi)$, then $\mathcal{A}^*$ has the symbol

$$\sigma_{\mathcal{A}^*}(x,\xi)\sim\sum_\alpha \frac{1}{\alpha!}\partial_\xi^\alpha D_x^\alpha\overline{p(x,\xi)}.$$

Ultimately, we have a look at the meaning of the variables x and y in the amplitude $a=a(x,y,\xi)$. Take for instance $a=a(x,y,\xi)=a_1(x)a_2(y)\xi^\alpha$, then we have

$$(\mathcal{A}u)(x) = a_1(x)D_x^\alpha\Big(a_2(x)u(x)\Big),$$

and we have here three operations (two multiplications and one differentiation) to be performed in a certain order.

Our desire to construct a symbol $p(x,\xi)$ from an amplitude $a(x,y,\xi)$ has mainly historical reasons. We also could have searched for a symbol $p(y,\xi)$. For understandable reasons, $p(x,\xi)$ is called the *left symbol* of $\mathcal{A}$, and $p(y,\xi)$ is called the *right symbol* of $\mathcal{A}$.

In the physics literature, a certain "compromise" has been found, the Weyl[2]-symbol

$$p^w = p^w\left(\frac{x+y}{2},\xi\right),$$

which has the key advantage that it becomes trivial to check whether an operator $\mathcal{A}$ is self-adjoint. Exactly in this case, its Weyl symbol is real-valued.

[2] Hermann Klaus Hugo Weyl, 1885–1955

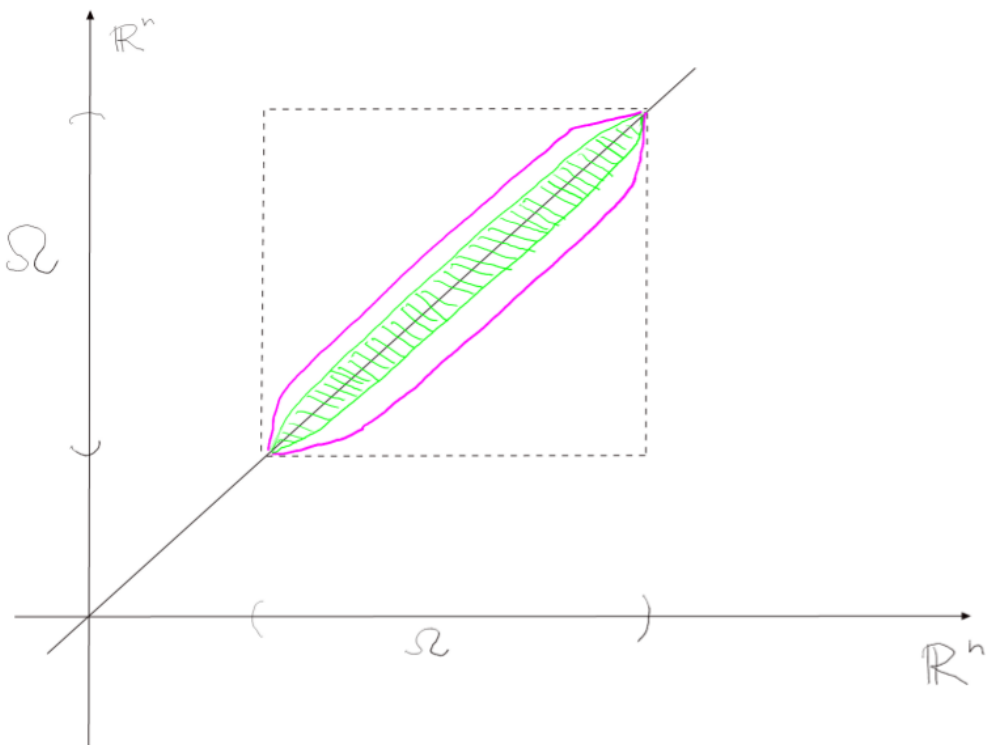

Figure 4.1: A kernel cutoff. In the green area, $\chi \equiv 1$. The pink line encloses $\operatorname{supp} \chi$.

4.3 Properly Supported Operators, Asymptotic Expansions, Algebras

Let $\mathcal{A}$ and $\mathcal{B}$ ΨDOs on a domain $\Omega \subset \mathbb{R}^n$. According to Proposition 3.18, both operators map from $\mathcal{D}(\Omega)$ into $\mathcal{E}(\Omega)$, and (after extension) from $\mathcal{E}'(\Omega)$ into $\mathcal{D}'(\Omega)$. But then the composition $\mathcal{A} \circ \mathcal{B}$ can not be defined directly, and (as a loophole), we split one operator, say $\mathcal{A}$:

$$\begin{aligned}(\mathcal{A}u)(x) &= \overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^n_\xi \times \Omega}} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)u(y)\,\mathrm{d}y\,\text{đ}\xi \qquad (4.2)\\ &= \overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^n_\xi \times \Omega}} e^{\mathrm{i}(x-y)\cdot\xi} \chi(x,y) a(x,y,\xi)u(y)\,\mathrm{d}y\,\text{đ}\xi \\ &\quad + \overset{\mathrm{Os}}{\iint\limits_{\mathbb{R}^n_\xi \times \Omega}} e^{\mathrm{i}(x-y)\cdot\xi}(1-\chi(x,y)) a(x,y,\xi)u(y)\,\mathrm{d}y\,\text{đ}\xi,\end{aligned}$$

where the function $\chi \in C^\infty(\Omega \times \Omega)$ is chosen in a clever way: the first oscillating integral on the RHS shall map $\mathcal{D}(\Omega)$ continuously into itself, also $\mathcal{E}(\Omega)$ continuously into itself, and the second oscillating integral on the RHS is a smoothing operator from $\mathcal{E}'(\Omega)$ into $\mathcal{E}(\Omega)$.

Definition 4.5 (*Proper* **map**). *A continuous map between two topological spaces is called* proper *if the pre-image of each compact set is again compact.*

Now we choose χ in such a way that $\chi \equiv 1$ in a neighbourhood of $\operatorname{diag}(\Omega \times \Omega)$. Then Proposition 3.16 gives us the claim concerning the second integral. And now $\operatorname{supp} \chi$ has to be chosen with the property that both canonical projections $(x,y) \mapsto x$ and $(x,y) \mapsto y$ are *proper* maps from $\operatorname{supp} \chi$ into Ω. This then implies the claim for the first integral, and the operator presented in this way is then called *properly supported*. See Figure 4.1. Note that χa will typically *not* satisfy global symbol estimates.

In formula (4.2), we have split $\mathcal{A}$ as $\mathcal{A} = \mathcal{A}_h + \mathcal{A}_r$, with $\mathcal{A}_h$ considered as main part and $\mathcal{A}_r$ as remainder. By Proposition 4.2, the operator $\mathcal{A}_h$ even has a symbol $\sigma_{\mathcal{A}_h}$ (not just an amplitude χa), and it has the form $\sigma_{\mathcal{A}_h}(x,\xi) = \exp(-\mathrm{i}x\cdot\xi)\mathcal{A}_h \exp(\mathrm{i}x\cdot\xi)$. Concerning $\mathcal{A}_r$, it may happen that a symbol can not be found, only an amplitude. That is the reason why $\Psi^{-\infty}$ has been defined as set of all operators with amplitude in $S^{-\infty}(\Omega\times\Omega\times\mathbb{R}^n)$, and not as set of all operators with symbol from $S^{-\infty}(\Omega\times\mathbb{R}^n)$.

We summarise:

Proposition 4.6. *Let $a = a(x,y,\xi) \in S^m_{\varrho,\delta}(\Omega\times\Omega\times\mathbb{R}^n)$ with $m\in\mathbb{R}$ and $0\le\delta<\varrho\le 1$, and let $\mathcal{A}$ be the associated ΨDO. Then $\mathcal{A}$ can be split as*

$$\mathcal{A} = \mathcal{A}_h + \mathcal{A}_r,$$

with $\mathcal{A}_h$ being a properly supported ΨDO with symbol $\sigma_{\mathcal{A}_h} \in S^m_{\varrho,\delta}(\Omega\times\mathbb{R}^n)$, whose symbol admits an asymptotic expansion

$$\sigma_{\mathcal{A}_h}(x,\xi) \sim \sum_\alpha \frac{1}{\alpha!}\left(\partial_\xi^\alpha D_y^\alpha a(x,y,\xi)\right)\Big|_{y=x}.$$

The operator $\mathcal{A}_r$ is a ΨDO with amplitude function in $S^{-\infty}(\Omega\times\Omega\times\mathbb{R}^n)$.

However, if $\mathcal{A}$ has a representation with a pseudodifferential symbol $a = a(x,\xi)$, then also $\mathcal{A}_r$ possesses a pseudodifferential symbol.

Here goes a relation between the amplitude function and the Schwartz kernel of the associated operator:

Lemma 4.7. *Let $a = a(x,y,\xi) \in S^m_{\varrho,\delta}(\Omega\times\Omega\times\mathbb{R}^n)$ with $m < -n$. Then the Schwartz kernel A of the operator $\mathcal{A}$ belonging to a is a function in $L^1_{\mathrm{loc}}(\Omega\times\Omega)$:*

$$A(x,y) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)\, đ\xi, \qquad (x,y)\in\Omega\times\Omega. \tag{4.3}$$

Proof. We have, for $u\in C_0^\infty(\Omega)$,

$$(\mathcal{A}u)(x) = \int_{\mathbb{R}^n_\xi}\int_\Omega e^{\mathrm{i}(x-y)\cdot\xi} a(x,y,\xi)u(y)\,\mathrm{d}y\, đ\xi,$$

and now $m<-n$ allows to swap the integrals. □

Proposition 4.8. *For an $\mathcal{A}$ on Ω, the following are equivalent:*

1. *$\mathcal{A}$ maps linearly and continuously from $\mathcal{E}'(\Omega)$ into $\mathcal{E}(\Omega)$,*
2. *$\mathcal{A}$ has a Schwartz kernel $A\in C^\infty(\Omega\times\Omega)$,*
3. *$\mathcal{A}$ it a ΨDO with amplitude $a\in S^{-\infty}(\Omega\times\Omega\times\mathbb{R}^n)$.*

Proof. The equivalence of 1. and 2. is part of the Schwartz kernel theorem.

Let $a\in S^{-\infty}(\Omega\times\Omega\times\mathbb{R}^n)$. Then the function A, given in (4.3), is a function from $C^\infty(\Omega\times\Omega)$.

Let us now be given $A\in C^\infty(\Omega\times\Omega)$, and we search $a\in S^{-\infty}(\Omega\times\Omega\times\mathbb{R}^n)$. Choose a function $\chi\in C_0^\infty(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n_\xi}\chi(\xi)\,đ\xi = 1$. Then we put

$$a(x,y,\xi) = e^{-\mathrm{i}(x-y)\cdot\xi}A(x,y)\chi(\xi),$$

which turns out to be an element of $S^{-\infty}(\Omega\times\Omega\times\mathbb{R}^n)$. If we construct a Schwartz kernel according to (4.3), we re-obtain the function A. □

From now on, most operators are assumed to be properly supported.

Let us talk about asymptotic expansions.

Our approach so far has been the following (compare the proof of Proposition 4.2): first we had a symbol $p \in S^m_{1,0}$; and from that, we then subsequently built a decomposition $p \sim \sum_{j=0}^{\infty} p_j$ with $p_j \in S^{m-j}_{1,0}$, in the sense of $p - \sum_{j=0}^{k-1} p_j \in S^{m-k}_{1,0}$ for all k.

Now our approach shall be the converse: let us be given p_0, p_1, ..., with $p_j \in S^{m-j}_{1,0}$. How can we then build some $p \in S^m_{1,0}$ with $p \sim \sum_{j=0}^{\infty} p_j$?

From the definition of asymptotic convergence we directly get that p (if it exists) is unique $\mod S^{-\infty}$.

Proposition 4.9. *Let p_0, p_1, ..., be given $p_j \in S^{m-j}_{1,0}(\Omega \times \mathbb{R}^n)$. Then there is a properly supported $p \in S^m_{1,0}(\Omega \times \mathbb{R}^n)$ with $p \sim \sum_{j=0}^{\infty} p_j$.*

Proof. For sake of simplicity, assume that the symbols p_j enjoy *global* symbol estimates:

$$\left|\partial_x^\alpha \partial_\xi^\beta p_j(x,\xi)\right| \le C_{\alpha\beta j} \langle\xi\rangle^{m-j-|\beta|}, \qquad \forall (x,\xi) \in \Omega \times \mathbb{R}^n, \quad \forall \alpha, \beta, j.$$

Now take come excision function $\varphi \in C^\infty(\mathbb{R}^n)$ with

$$\varphi(\xi) = \begin{cases} 0 & : |\xi| \le 1, \\ 1 & : |\xi| \ge 2, \end{cases}$$

and let us hope that the following *ansatz* works:

$$p(x,\xi) := \sum_{j=0}^{\infty} \varphi\left(\frac{\xi}{t_j}\right) p_j(x,\xi).$$

Here, the numbers $t_0, t_1, \dots$ are still available for choice. We assume that they form a sequence of positive numbers that diverges strictly monotonically to $+\infty$. Then, for each fixed (x,ξ), this series contains only a finite number of non-zero items. Moreover, we have $\varphi(\frac{\cdot}{t_j}) \in S^0_{1,0}$, with symbol estimates that do not depend on j:

$$\left|\partial_\xi^\beta \varphi\left(\frac{\xi}{t_j}\right)\right| \le C_\beta \langle\xi\rangle^{-|\beta|}, \qquad \xi \in \mathbb{R}^n,$$

with C_β independent of j. Our goal is to choose the parameters t_j in such a way that

$$\left|\partial_x^\alpha \partial_\xi^\beta \left(\varphi\left(\frac{\xi}{t_j}\right) p_j(x,\xi)\right)\right| \le 2^{-j} \langle\xi\rangle^{m+1-j-|\beta|}, \qquad \xi \in \mathbb{R}^n, \qquad \forall (x,\xi) \in \Omega \times \mathbb{R}^n,$$

for all combinations of α, β, j for which $|\alpha| + |\beta| \le j$. This is possible, because (employing a generic constant C_j)

$$\begin{aligned}
\left|\partial_x^\alpha \partial_\xi^\beta \left(\varphi\left(\frac{\xi}{t_j}\right) p_j(x,\xi)\right)\right| &\le \sum_{\beta'+\beta''=\beta} \binom{\beta}{\beta'} \left|\partial_\xi^{\beta'} \varphi\left(\frac{\xi}{t_j}\right)\right| \cdot \left|\partial_x^\alpha \partial_\xi^{\beta''} p_j(x,\xi)\right| \\
&\le C_j \sum_{\beta'+\beta''=\beta} \langle\xi\rangle^{-|\beta'|} \langle\xi\rangle^{m-j-|\beta''|} \\
&\le C_j \langle\xi\rangle^{m-j-|\beta|} \\
&= \frac{C_j}{\langle\xi\rangle} \langle\xi\rangle^{m+1-j-|\beta|}.
\end{aligned}$$

Now we can assume $\frac{|\xi|}{t_j} > 1$, because otherwise the factor $\varphi(\dots)$ vanishes identically anyway, making the claimed inequality a triviality. Now we choose t_j so large that $\frac{C_j}{\langle t_j\rangle} \le 2^{-j}$.

Next we will prove that

$$\left|\partial_x^\alpha \partial_\xi^\beta p(x,\xi)\right| \le C_{\alpha\beta} \langle\xi\rangle^{m+1-|\beta|},$$

i.e. $p \in S^{m+1}_{1,0}$. To this end, let α and β be given, and choose $N \geq |\alpha| + |\beta|$. Then we write

$$\begin{aligned}\left|\partial_x^\alpha \partial_\xi^\beta p(x,\xi)\right| &\leq \sum_{j=0}^{N} \left|\partial_x^\alpha \partial_\xi^\beta (\chi(\xi/t_j) p_j(x,\xi))\right| + \sum_{j=N+1}^{\infty} \left|\partial_x^\alpha \partial_\xi^\beta (\chi(\xi/t_j) p_j(x,\xi))\right| \\ &\leq C_N \sum_{j=0}^{N} \langle\xi\rangle^{m-j-|\beta|} + \sum_{j=N+1}^{\infty} 2^{-j} \langle\xi\rangle^{m+1-j-|\beta|} \\ &\leq \langle\xi\rangle^{m+1-|\beta|} \left(N \cdot C_N + \sum_{j=N+1}^{\infty} 2^{-j} \right),\end{aligned}$$

which indeed gives $p \in S^{m+1}_{1,0}$, and this is one order too high. But this is no problem: just temporarily omit the item with $j = 0$ in the sum defining p, and go through the above calculation once more. Then we get $p \in S^m_{1,0}$.

It remains to show that, for all k, we have $p - \sum_{j=0}^{k-1} \chi(\dots) p_j \in S^{m-k}_{1,0}$, whose proof is analogous. □

The next result is useful for the following: suppose someone has given you some symbol p and other symbols p_j, with the claim $p \sim \sum_j p_j$. How can you verify this claim easily ? In particular we prefer to avoid estimating all the derivatives of all the truncation errors

$$\left|\partial_x^\alpha \partial_\xi^\beta \left(p - \sum_{j=0}^{k-1} p_j(x,\xi) \right)\right|$$

for all α, for all β, for all k, for all x, for all ξ (because this gets tedious quickly).

Proposition 4.10. *Let $p_j \in S^{m_j}_{1,0}(\Omega \times \mathbb{R}^n)$ be given with $m_0 > m_1 > m_2 > \dots \to -\infty$. Further, let us be given $p \in C^\infty(\Omega \times \mathbb{R}^n)$, and suppose that for all $\alpha, \beta \in \mathbb{N}^n$, and for all compact sets $K \Subset \Omega$ there are numbers $\mu = \mu(\alpha,\beta,K)$ and $C = C(\alpha,\beta,K)$ such that*

$$\left|\partial_x^\alpha \partial_\xi^\beta p(x,\xi)\right| \leq C \langle\xi\rangle^{\mu}, \qquad \forall (x,\xi) \in K \times \mathbb{R}^n.$$

Additionally, we assume that for each compact set $K \Subset \Omega$ and for each $k \in \mathbb{N}$ there are numbers $\mu = \mu(k,K)$ and $C_{k,K}$ with

$$\left|p(x,\xi) - \sum_{j=0}^{k-1} p_j(x,\xi)\right| \leq C_{k,K} \langle\xi\rangle^{\mu(k,K)}, \qquad \forall (x,\xi) \in K \times \mathbb{R}^n,$$

and $\lim_{k\to\infty} \mu(k,K) = -\infty$ for each K.

Then we have $p \sim \sum_j p_j$.

Proof. See [22], Proposition 3.6. □

Let us now have a look at algebra properties. We have the following:

- operators from $\Psi^m_{1,0}$, properly supported, with $m \in \mathbb{R}$,
- their associated symbols from $S^m_{1,0}$,
- actions:
  - adding two operators: $\Psi^{m_1}_{1,0} \times \Psi^{m_2}_{1,0} \to \Psi^{\max(m_1,m_2)}_{1,0}$
  - scalar multiplication: $\mathbb{C} \times \Psi^m_{1,0} \to \Psi^m_{1,0}$
  - composition of two operators: $\Psi^{m_1}_{1,0} \times \Psi^{m_2}_{1,0} \to \Psi^{m_1+m_2}_{1,0}$

– adjoining an operator: $\Psi^m_{1,0} \to \Psi^m_{1,0}$,
– transposing an operators: $\Psi^m_{1,0} \to \Psi^m_{1,0}$.

We summarise this in the statement that the set $\Psi^\infty_{1,0} = \cup_{m\in\mathbb{R}}\Psi^m_{1,0}$ of the properly supported pseudodifferential operators forms an *algebra*.

An associated algebra comes from the set of their pseudodifferential symbols $S^\infty_{1,0} = \cup_{m\in\mathbb{R}} S^m_{1,0}$. You may wish to devise an *algebra isomorphism* yourself.

4.4 Classical Symbols

Up to now, we have had a look at the symbol classes $S^m_{1,0}$ and $S^m_{\varrho,\delta}$. Their elements $p = p(x,\xi)$ have, as only exploitable property, estimates for all derivatives of the form $\partial^\alpha_x \partial^\beta_\xi p(x,\xi)$.

Now we want more:

Definition 4.11 (**Classical symbols, homogeneous symbols**). *We say that a symbol $p \in S^m_{1,0}(\Omega\times\mathbb{R}^n)$ is a* classical symbol *(and write $p \in S^m_{\mathrm{cl}}(\Omega\times\mathbb{R}^n)$) if there is an excision function $\varphi \in C^\infty(\mathbb{R}^n)$ and functions $p_j \in C^\infty(\Omega \times (\mathbb{R}^n \setminus 0))$ such that each p_j is positive homogeneous of order $m-j$ in the variable ξ:*

$$p_j(x,t\xi) = t^{m-j}p_j(x,\xi), \qquad \forall (x,\xi,t) \in \Omega\times\mathbb{R}^n\times\mathbb{R}_+,$$

and additionally

$$p(x,\xi) \sim \sum_{j=0}^{\infty} \varphi(\xi)p_j(x,\xi),$$

in the usual sense of $p - \sum_{j=0}^{k-1}\varphi p_j \in S^{m-k}$.

The functions p_j are called homogeneous symbols*, and we write $p_j \in S^{m-j}_{\mathrm{Hom}}(\Omega\times(\mathbb{R}^n\setminus 0))$.*

Note that homogeneous symbols with $m-j<0$ will typically have poles at $\xi = 0$, and symbols from $S^m_{\varrho,\delta}$ with $(\varrho,\delta) \neq (1,0)$ will almost never be classical.

Definition 4.12 (**Classical operators**). *Let $p \in S^m_{\mathrm{cl}}(\Omega\times\mathbb{R}^n)$. The the associated ΨDO $\mathcal{P}$, defined via*

$$(\mathcal{P}u)(x) = \int_{\mathbb{R}^n_\xi}\int_\Omega e^{\mathrm{i}(x-y)\cdot\xi}p(x,\xi)u(y)\,\mathrm{d}y\,đ\xi, \qquad u \in C_0^\infty(\Omega),$$

is called a classical pseudodifferential operator*, and we write this as $\mathcal{P} \in \Psi^m_{\mathrm{cl}}(\Omega)$.*

Examples are

- every PDO $\mathcal{P} = \sum_{|\alpha|\le m} a_\alpha(x)D^\alpha_x$ with symbol $p(x,\xi) = \sum_{|\alpha|\le m} a_\alpha(x)\xi^\alpha$,
- the operator $\langle D\rangle$ with symbol $\langle\xi\rangle = \sqrt{1+|\xi|^2}$.

Using the methods of Proposition 4.6, we can show:

Proposition 4.13. *Let $\mathcal{A} \in \Psi^m_{\mathrm{cl}}(\Omega)$. Then $\mathcal{A}$ can be split as $\mathcal{A} = \mathcal{A}_h + \mathcal{A}_r$, where $\mathcal{A}_h$ is properly supported, belongs to $\Psi^m_{\mathrm{cl}}(\Omega)$, possesses the same asymptotic expansion, and $\mathcal{A}_r \in \Psi^{-\infty}(\Omega)$.*

We quickly check that the classical ΨDO with symbols from $\bigcup_{m\in\mathbb{Z}} S^m_{\mathrm{cl}}$ also form an algebra (assume that everybody is properly supported).

A **first advantage** of the classical symbols is that only now we are in a position to speak about the *principal symbol* of an operator.

A **second advantage** are *tensor product representations* of the form

$$(\mathcal{P}u)(x) \sim \sum_{j,l,m} a_{jlm}(x)\cdot(\mathcal{P}_{jlm}(D)u)(x),$$

modulo smoothing operators, $\mathcal{P} \in S^M_{\mathrm{cl}}$, and $\mathcal{P}_{jlm} \in S^{M-j}_{\mathrm{Hom}}$ possesses "constant coefficients" (which means that they do not depend on x). Even better, the convergence of the series $\sum_{l,m}$ is totally harmless.

Let us have a closer look. Take

$$p(x,\xi) \sim \sum_{j=0}^{\infty} \varphi(\xi) p_j(x,\xi) \in S^M_{\mathrm{cl}}(\Omega \times \mathbb{R}^n),$$

with some excision function φ. Then we have

$$p_j(x,\xi) = p_j\left(x, |\xi| \frac{\xi}{|\xi|}\right) = |\xi|^{M-j} \cdot p_j\left(x, \frac{\xi}{|\xi|}\right),$$

hence p_j is uniquely determined by its values on $\Omega \times S^{n-1}$, with S^{n-1} being the unit sphere in $\mathbb{R}^n$. Take $n = 2$, for instance, then we can write

$$\frac{\xi}{|\xi|} = \omega = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \in S^1, \qquad 0 \le t \le 2\pi,$$

hence also

$$p_j(x,\xi) = |\xi|^{M-j} p_j(x_1, x_2, \cos t, \sin t), \qquad t = t(\xi).$$

Now keep x and $|\xi|$ fixed, then p_j becomes a 2π-period function in t, hence we have the Fourier series decomposition

$$p_j(x,\xi) = |\xi|^{M-j} \left(\frac{a_0(x)}{2} + \sum_{l=1}^{\infty} \Big(a_l(x) \cos(l \cdot t(\xi)) + b_l(x) \sin(l \cdot t(\xi)) \Big) \right),$$

and this looks very much like the above *tensor product representation*.

The general case needs the *spheric harmonics*. We follow the representation in [10, Volume 2] and their choice of variables.

Let $x = (x_1, x_2, \ldots, x_{p+2}) \in \mathbb{R}^{p+2}$ and $p \ge 1$. A polynomial $H_n = H_n(x)$ is called *harmonic polynomial of degree* n if $\triangle H_n(x) = 0$ for each $x \in \mathbb{R}^{p+2}$, and additionally

$$H_n(\lambda x) = \lambda^n H_n(x), \qquad \forall (x,\lambda) \in \mathbb{R}^n \times \mathbb{R}.$$

Here we assume $n \ge 0$. There are

$$h(h,p) = (2n+p) \frac{(n+p-1)!}{p!n!} = \mathcal{O}(\langle n \rangle^p)$$

linearly independent homogeneous polynomials of degree n, see Section 11.2 in [10]. They have (compare Theorem 1 in Section 11.2 in [10]) the following form:

Let $m_0, \ldots, m_p$ be integers with

$$n = m_0 \ge m_1 \ge m_2 \ge \cdots \ge m_{p-1} \ge |m_p|,$$

and let r_k be defined by

$$r_k = \left(x_{k+1}^2 + \cdots + x_{p+2}^2\right)^{1/2}, \quad k = 0, \ldots, p, \quad r_0 = r.$$

Then the functions $H(\{m_k\}; \cdot)$ with

$$H(\{m_k\}, x) := H(n, m_1, \ldots, m_p, x) = \\ = \left(\frac{x_{p+1}}{r_p} + \mathrm{i} \frac{x_{p+2}}{r_p} \right)^{m_p} r_p^{m_p} \prod_{k=0}^{p-1} r_k^{m_k - m_{k+1}} C^{m_{k+1} + (p-k)/2}_{m_k - m_{k+1}} \left(\frac{x_{k+1}}{r_k} \right)$$

are a complete system of $h(n,p)$ linearly independent harmonic polynomials of degree n. The functions $C^{m_{k+1}+(p-k)/2}_{m_k-m_{k+1}}$ are the GEGENBAUER[3] polynomials (see Section 3.5 in [10, Volume 1]), and they can be written in terms of the hypergeometric functions ${}_2\mathrm{F}_1(a,b;c;z)$ like this:

$$\begin{aligned} n!C_n^\lambda(x) &:= (2\lambda)_n\, {}_2\mathrm{F}_1\left(-n, n+2\lambda; \lambda+\frac{1}{2}; \frac{1-x}{2}\right), \\ {}_2\mathrm{F}_1(a,b;c;z) &:= \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k k!} z^k, \\ (a)_n &:= \frac{\Gamma(a+n)}{\Gamma(a)}. \end{aligned}$$

The restriction of these functions to the unit sphere $\{|x|=1\}$ then is a complete system of orthogonal functions in the sense of an orthogonal basis of $L^2(\{|x|=1\})$. Those functions are then called *spheric harmonics*, and they are denoted by $Y(\{m_k\},\theta,\varphi)$ with

$$Y(\{m_k\},\theta,\varphi) := r^{-n} H(\{m_k\},x),$$

where $(r,\theta,\varphi) = (r,\theta_1,\dots,\theta_p,\varphi)$ are the polar coordinates in $\mathbb{R}^{p+2}$:

$$\begin{aligned} x_1 &= r\cos\theta_1, \\ x_2 &= r\sin\theta_1\cos\theta_2, \\ x_3 &= r\sin\theta_1\sin\theta_2\cos\theta_3, \\ &\dots, \\ x_p &= r\sin\theta_1\sin\theta_2\dots\sin\theta_{p-1}\cos\theta_p, \\ x_{p+1} &= r\sin\theta_1\sin\theta_2\dots\sin\theta_{p-1}\sin\theta_p\cos\varphi, \\ x_{p+2} &= r\sin\theta_1\sin\theta_2\dots\sin\theta_{p-1}\sin\theta_p\sin\varphi, \\ 0 \le r &< \infty, \quad 0 \le \theta_1,\dots,\theta_p \le \pi, \quad 0 \le \varphi \le 2\pi. \end{aligned}$$

The volume element then is

$$\mathrm{d}V = r^{p+1}(\sin\theta_1)^p(\sin\theta_2)^{p-1}\dots(\sin\theta_p)\,\mathrm{d}r\,\mathrm{d}\theta_1\dots\mathrm{d}\theta_p\,\mathrm{d}\varphi,$$

and the surface element is

$$\mathrm{d}\sigma = r^{p+1}(\sin\theta_1)^p(\sin\theta_2)^{p-1}\dots(\sin\theta_p)\,\mathrm{d}\theta_1\dots\mathrm{d}\theta_p\,\mathrm{d}\varphi.$$

The Laplacian in $\mathbb{R}^{p+2}$ then becomes

$$\begin{aligned} \triangle &= \triangle_r + \frac{1}{r^2}\triangle_S \qquad (4.4) \\ &= r^{-p-1}\frac{\partial}{\partial r}\left(r^{p+1}\frac{\partial}{\partial r}\right) \\ &\quad + \frac{1}{r^2}\frac{1}{(\sin\theta_1)^p}\frac{\partial}{\partial\theta_1}\left((\sin\theta_1)^p\frac{\partial}{\partial\theta_1}\right) \\ &\quad + \frac{1}{r^2}\frac{1}{(\sin\theta_1)^2}\frac{1}{(\sin\theta_2)^{p-1}}\frac{\partial}{\partial\theta_2}\left((\sin\theta_2)^{p-1}\frac{\partial}{\partial\theta_2}\right) \\ &\quad + \frac{1}{r^2}\frac{1}{(\sin\theta_1\sin\theta_2)^2}\frac{1}{(\sin\theta_3)^{p-2}}\frac{\partial}{\partial\theta_3}\left((\sin\theta_3)^{p-2}\frac{\partial}{\partial\theta_3}\right) \\ &\quad \dots \\ &\quad + \frac{1}{r^2}\frac{1}{(\sin\theta_1\sin\theta_2\dots\sin\theta_{p-1})^2}\frac{1}{(\sin\theta_p)^1}\frac{\partial}{\partial\theta_p}\left((\sin\theta_p)^1\frac{\partial}{\partial\theta_p}\right) \\ &\quad + \frac{1}{r^2}\frac{1}{(\sin\theta_1\sin\theta_2\dots\sin\theta_p)^2}\frac{\partial^2}{\partial\varphi^2}. \end{aligned}$$

[3]LEOPOLD GEGENBAUER, 1849–1903

Here $\triangle_r$ operates in radial direction only, and the operator $\triangle_S$ acts only upon the angles. We call $\triangle_S$ also the *Laplace*[4]*–Beltrami*[5] *operator on the unit sphere.*

We remark that the spheric harmonics $Y(\{m_k\}, \theta, \varphi)$ are much more than only a basis of the L^2 on the sphere — they are even eigenfunctions of $\triangle_S$:

Proposition 4.14. *Let $Y(\{m_k\}, \cdot, \cdot)$ be a spheric harmonic of degree n. Then*

$$\triangle_S Y(\{m_k\}, \theta, \varphi) = -n(n+p)Y(\{m_k\}, \theta, \varphi).$$

Proof. The function $H(\{m_k\}, x) = r^n Y(\{m_k\}, \theta, \varphi)$ is a harmonic polynomial of degree n, hence we have

$$0 = \triangle H = \triangle(r^n Y) = (\triangle_r r^n)Y + r^{n-2} \triangle_S Y = r^{n-2}(n(n+p) + \triangle_S)Y.$$

□

We return to the traditional notation:

$$\begin{aligned} p+2 &\longmapsto n, \\ (n, m_1, \ldots, m_p) &\longmapsto (l, m), \quad 1 \le m \le h(l, n-2) = \mathcal{O}(\langle l \rangle^{n-2}), \\ (x_1, \ldots, x_{p+2}) &\longmapsto \xi \in \mathbb{R}^n, \\ (\theta_1, \ldots, \theta_p, \varphi) &\longmapsto \frac{\xi}{|\xi|} \in S^{n-1}, \\ Y(\{m_k\}, \theta, \varphi) &\longmapsto Y_{lm}(\xi) \text{ or rather } Y_{lm}(\theta, \varphi). \end{aligned}$$

Remark 4.15. *Pedantically, we should rather write $Y_{lm}(\xi/|\xi|)$. But for reasons of simplicity of notation we make the agreement that each argument ξ of Y_{lm} is being tacitly normalised.*

Our results become in the new notation:

$$\begin{aligned} &\{Y_{lm}\}_{l=0,\ldots,\infty,\ m=1,\ldots,h(l,n-2)} \text{ is an orthogonal basis of } L^2(S^{n-1}), \\ &h(l, n-2) = \mathcal{O}(\langle l \rangle^{n-2}), \\ &Y_{lm} \in C^\infty(S^{n-1}) \quad \forall l, m, \\ &-\triangle_S Y_{lm}(\xi) = l(l+n-2)Y_{lm}(\xi). \end{aligned}$$

We may assume (wlog) that $\|Y_{lm}\|_{L^2(S^{n-1})} = 1$ for all l, m.

Let us now be given $p_j \in S^{M-j}_{\mathrm{Hom}}$, then we have

$$\begin{aligned} p_j(x, \xi) &= |\xi|^{M-j} p_j\left(x, \frac{\xi}{|\xi|}\right) = |\xi|^{M-j} \sum_{l,m} a_{jlm}(x) Y_{lm}(\xi), \\ a_{jlm}(x) &= \left\langle p_j\left(x, \frac{\xi}{|\xi|}\right), Y_{lm}(\xi)\right\rangle_{L^2(S^{n-1}_\xi)}. \end{aligned}$$

This Fourier series converges obviously in $L^2(S^{n-1})$, for each fixed $x \in \Omega$. But in reality, this convergence is much faster, because for each $l \ge 1$ and each $N \in \mathbb{N}$, we have

$$\begin{aligned} |a_{jlm}(x)| &= \left|\left\langle p_j, (-\triangle_S)^N (l(l+n-2))^{-N} Y_{lm}\right\rangle\right| \\ &= \frac{1}{(l(l+n-2))^N} \left|\left\langle (-\triangle_S)^N p_j, Y_{lm}\right\rangle\right| \\ &\le \frac{C_N}{l^{2N}} \left\|\triangle_S^N p_j(x, \cdot)\right\|_{L^2(S^{n-1})}, \end{aligned}$$

since the operator $\triangle_S$ is self-adjoint. Here we crucially exploit that the sphere S^{n-1} is a compact manifold without boundary.

The number of admissible m depends polynomially on l, because of $h(l, n-2) = \mathcal{O}(\langle l \rangle^{n-2})$. But this is no problem because N can be chosen as large as we want, which then reveals us that the sequence $(\sum_m |a_{jlm}|)_{l=0,1,2,\ldots}$ decays faster than every power of l^{-1}.

[4] Pierre–Simon Laplace, 1749–1827
[5] Eugenio Beltrami, 1835–1900

4.5 Elliptic ΨDOs

Now our goal shall be to "invert" a ΨDO. We observe quickly that this will not be possible for all operators, because: let $p \in S^m_{\mathrm{cl}}$ and $p \equiv 0$ in a conic neighbourhood of (x_0, ξ_0) with $\xi_0 \neq 0$. Now let $u \in L^2$ be a function whose Fourier transform is identically equal to zero outside a tiny conic neighbourhood of ξ_0. Then we have $\mathcal{P}u \equiv 0$, but u is not the zero function. This means that the operator $\mathcal{P}$ has a null space which contains functions from L^2. This annoys us. Instead we wish the following: if it is unavoidable to have a non-trivial null space of the operator $\mathcal{P}$, then this null space shall only contain smooth functions.

We have seen that the symbol of such an "invertible" operator cannot vanish in any conic neighbourhood. Actually, we will now consider pseudodifferential symbols that can be "estimated from below":

Definition 4.16 (**Elliptic symbols**). *A symbol $p \in S^m_{\varrho,\delta}(\Omega \times \mathbb{R}^n)$ with $0 \le \delta, \varrho \le 1$ is called* elliptic *if there is some R and some $c > 0$ with*

$$|p(x, \xi)| \ge c \langle\xi\rangle^m, \qquad \forall x \in \Omega, \quad |\xi| > R.$$

In case of a classical symbol $p \in S^m_{\mathrm{cl}}$ with principal symbol p_m, this means $|p_m(x, \xi)| \ge c|\xi|^m$ for all $x \in \Omega$ and all $\xi \in \mathbb{R}^n \setminus 0$.

A generalisation are hypo-elliptic symbols:

Definition 4.17 (**Hypo-elliptic symbols**). *A symbol $p \in S^m_{\varrho,\delta}(\Omega \times \mathbb{R}^n)$ with $0 \le \delta, \varrho \le 1$ is called* hypo-elliptic *if there is some R and some $c > 0$, and also some $m_0 \in \mathbb{R}$ with*

$$|p(x, \xi)| \ge c \langle\xi\rangle^{m_0}, \qquad \forall x \in \Omega, \quad |\xi| > R.$$

Furthermore, we require the estimates

$$|\partial_x^\alpha \partial_\xi^\beta p(x, \xi)| \le C_{\alpha,\beta,K} |p(x, \xi)| \langle\xi\rangle^{-\varrho|\beta|+\delta|\alpha|}, \qquad \forall x \in K \Subset \Omega, \quad |\xi| > R, \quad \alpha, \beta \in \mathbb{N}^n.$$

We quickly find that $m_0 \le m$.

Example: *Let $\mathcal{P} = \mathcal{P}(x, D_x) = \sum_{|\alpha| \le m} a_\alpha(x) D_x^\alpha$ be a differential operator with principal symbol $p_m(x, \xi) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha$, for which $|p(x, \xi)| \ge c|\xi|^m$, for all (x, ξ).*

Example: *The $\partial_t - \triangle_x$ is hypo-elliptic in $\mathbb{R}^{1+n}$, but not elliptic.*

Example: *The symbol $\sqrt{1 + \xi_1^2 + \xi_2^4}$ is hypo-elliptic, but not elliptic.*

From now on let $\mathcal{P}$ be elliptic, properly supported in Ω, $\mathcal{P}u = f$ with $u \in C_0^\infty(\Omega)$ and $f \in C^\infty(\Omega)$.

For simplicity, we suppose $\mathcal{P} \in \Psi^m_{\mathrm{cl}}(\Omega)$, but hypo-elliptic symbols $\mathcal{P} \in \Psi^m_{\varrho,\delta}(\Omega)$ with $0 \le \delta < \varrho \le 1$ would also be possible (the calculations would just become longer).

Our objective is to find an operator $\mathcal{Q}$ with $\mathcal{Q} \circ \mathcal{P} = I$. We make for $\mathcal{Q}$ an *ansatz* as a ΨDO with symbol q, and we obtain then

$$1 \overset{!}{\sim} \sum_{|\alpha|=0}^{\infty} \frac{1}{\alpha!} \left(\partial_\xi^\alpha q(x, \xi)\right) \left(D_x^\alpha p(x, \xi)\right). \tag{4.5}$$

This brings us to $q \in S^{-m}_{\mathrm{cl}}(\Omega \times \mathbb{R}^n)$, thus also

$$q(x, \xi) \sim \sum_{j=0}^{\infty} \varphi(\xi) q_j(x, \xi), \qquad q_j \in S^{-m-j}_{\mathrm{Hom}}.$$

We have a corresponding decomposition p. An exercise shows that asymptotic series may be differentiated term-wise, hence we find

$$1 \overset{!}{\sim} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{|\alpha|=0}^{\infty} \frac{1}{\alpha!} \left(\partial_\xi^\alpha q_j(x, \xi)\right) \left(D_x^\alpha p_k(x, \xi)\right).$$

The terms of order $-l$ with $l = 0, 1, 2, \dots$ are

$$\sum_{j+k+|\alpha|=l} \frac{1}{\alpha!} \left(\partial_\xi^\alpha q_j(x, \xi)\right) \left(D_x^\alpha p_k(x, \xi)\right).$$

For $l = 0$ this shall be equal to 1, which means $1 = q_0(x,\xi)p_0(x,\xi)$, and therefore

$$q_0 = \frac{1}{p_0} \in S^{-m}_{\mathrm{Hom}}.$$

Here the assumed ellipticity of p becomes crucial.

Next we consider $l = 1$, resulting in

$$0 \stackrel{!}{=} \sum_{|\alpha|=1} \frac{1}{\alpha!} \left(\partial_\xi^\alpha q_0(x,\xi)\right)\left(D_x^\alpha p_0(x,\xi)\right) + q_1 p_0 + q_0 p_1,$$

which brings us to

$$q_1 = -\frac{1}{p_0}\left(\sum_{|\alpha|=1} \frac{1}{\alpha!} \left(\partial_\xi^\alpha q_0(x,\xi)\right)\left(D_x^\alpha p_0(x,\xi)\right) + q_0 p_1\right) \in S^{-m-1}_{\mathrm{Hom}}.$$

The remaining terms are determined similarly, and we get $q_j \in S^{-m-j}_{\mathrm{Hom}}$. Pick now some excision function φ, and we get some $q \in S^{-m}_{\mathrm{cl}}$, for which

$$q \sim \sum_{j=0}^{\infty} \varphi q_j.$$

It is easy to check that this construction then indeed satisfies (4.5).

We summarise:

Proposition 4.18. *Let $p \in S^m_{\mathrm{cl}}(\Omega \times \mathbb{R}^n)$ with $m \in \mathbb{R}$, elliptic, not necessarily properly supported. Then there is at least one $q \in S^{-m}_{\mathrm{cl}}(\Omega \times \mathbb{R}^n)$, elliptic and properly supported, such that the associated operators $\mathcal{P}$ and $\mathcal{Q}$ satisfy:*

$$\mathcal{Q}(x, D_x) \circ \mathcal{P}(x, D_x) = \mathrm{id} + \mathcal{R}(x, D_x),$$

where $\mathcal{R} \in \Psi^{-\infty}$.

Each such $\mathcal{Q}$ is called *left parametrix* for $\mathcal{P}$. The proposition remains true for $S^m_{1,0}$ instead of S^m_{cl}, but then the proof gets more involved. For details, see [22].

An open question is about uniqueness of $\mathcal{Q}$, and whether $\mathcal{Q}$ were a *right parametrix*, too.

Definition 4.19. *Let p_1 and p_2 be from $S^m_{1,0}$, for $m \in \mathbb{R}$. We say that p_1 and p_2 are equivalent, $p_1 \sim p_2$, whenever $p_1 - p_2 \in S^{-\infty}$.*

Obviously, this is an equivalence relation. Equivalent classical symbols have the same asymptotic expansion. And we also see: if p_1 is elliptic and $p_1 \sim p_2$, then also p_2 is elliptic.

For each symbol, we can find an equivalent symbol with proper support. In other words: each equivalence class of symbols contains a properly supported representative. All the following algebraic considerations refer to such a representative.

Let us define an operation $\sharp$ for such equivalence classes,

$$p \sharp q = \sigma(\mathcal{P} \circ \mathcal{Q}),$$

where σ is that symbol which maps a ΨDO to its uniquely determined symbol.

This operation $\sharp$ is a binary operation on following set of residue classes

$$S^{\mathrm{ellipt}} := \bigcup_{m\in\mathbb{R}} S^{m,\mathrm{ellipt}}_{1,0} \Big/ \sim .$$

The operation $\sharp$ is associative (exercise) and has at least one left-neutral element, namely $e = e(x,\xi) \equiv 1$, the symbol of the identity operator. To each elliptic symbol $p \in S^{\mathrm{ellipt}}$, there is at least one elliptic symbol $q \in S^{\mathrm{ellipt}}$ with $q \sharp p = e$, *i.e.* each element of S^{ellipt} possesses at least one left-inverse element. According to a known theorem from algebra, then $(S^{\text{ellipt}}, \sharp)$ is a group. Consequently, the left-inverse element is unique, and it is also right-inverse.

We come to applications.

Let $\mathcal{P}u = f$ with $u \in \mathcal{E}'(\Omega)$ and $\mathcal{P} \in \Psi_{1,0}^{m,\text{ellipt}}$ with proper support. Then there is a $\mathcal{Q}$ with $\mathcal{Q} \circ \mathcal{P} = \mathrm{id} + \mathcal{R}$, where $\mathcal{Q} \in \Psi_{1,0}^{-m,\text{ellipt}}$, and $\mathcal{R}$ maps from $\mathcal{E}'(\Omega)$ into $\mathcal{E}(\Omega)$. We then have

$$u + \mathcal{R}u = (\mathrm{id} + \mathcal{R})u = \mathcal{Q} \circ \mathcal{P}u = \mathcal{Q}f,$$

hence $u = \mathcal{Q}f - \mathcal{R}u$, where $\mathcal{R}u \in C^\infty(\Omega)$. Now we apply 3.25 once for $\mathcal{Q}$ and once for $\mathcal{P}$ (note that $S_\Phi = \mathrm{diag}(\Omega \times \Omega)$):

$$\text{sing-supp}(u) = \text{sing-supp}(\mathcal{Q}f) \subset \text{sing-supp}(f) = \text{sing-supp}(\mathcal{P}u) \subset \text{sing-supp}(u).$$

This then can be summarised like this:

Proposition 4.20. *Let $\mathcal{P} \in \Psi_{1,0}^m(\Omega)$ elliptic and properly supported. Then we have, for each $u \in \mathcal{E}'(\Omega)$,*

$$\text{sing-supp}(u) = \text{sing-supp}(\mathcal{P}u).$$

The same holds for elliptic operators $\mathcal{P} \in \Psi_{1,0}^m(\mathbb{R}^n)$ (with global symbol estimates) and $u \in \mathcal{S}'(\mathbb{R}^n)$.

A consequence then is that the null space of $\mathcal{P}$ contains only smooth functions, as desired.

Chapter 5

Mapping Properties

5.1 The Case of Ψ^m_{cl} on $\mathbb{R}^n$

We define Sobolev spaces on $\mathbb{R}^n$:

Definition 5.1. *For $s \in \mathbb{R}$, we define the* SOBOLEV[1] *space*

$$H^s(\mathbb{R}^n) := \left\{ u \in \mathcal{S}'(\mathbb{R}^n) \colon \hat{u} \in L^2_{\mathrm{loc}}(\mathbb{R}^n_\xi) \text{ and } \int_{\mathbb{R}^n_\xi} \langle\xi\rangle^{2s} |\hat{u}(\xi)|^2 \,\mathrm{d}\xi < \infty \right\}$$

and its norm

$$\|u\|_{H^s(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^n_\xi} \langle\xi\rangle^{2s} |\hat{u}(\xi)|^2 \,\mathrm{d}\xi \right)^{1/2}.$$

Exercise: *For which $s \in \mathbb{R}$ is $\delta \in H^s(\mathbb{R}^n)$?*

Theorem 5.2. *Every operator from Ψ^r_{cl} with $r \in \mathbb{R}$ and global symbol estimates maps $H^s(\mathbb{R}^n)$ continuously into $H^{s-r}(\mathbb{R}^n)$, for every $s \in \mathbb{R}$.*

The proof is quite easy, because of the tensor product structure of classical operators.

Proof. By definition of $H^s(\mathbb{R}^n)$, the operator $\langle D\rangle^r$ is a continuous isomorphism from $H^t(\mathbb{R}^n)$ onto $H^{t-r}(\mathbb{R}^n)$, for each $t \in \mathbb{R}^n$. Now take $p \in S^r_{\mathrm{cl}}$. Then we can write

$$p(x,\xi) = \Big(p(x,\xi) \langle\xi\rangle^{-r} \Big) \cdot \langle\xi\rangle^r, \qquad \mathcal{P}(x, D_x) = \mathrm{Op}\{p(x,\xi)\langle\xi\rangle^{-r}\} \circ \langle D\rangle^r,$$

with $\mathrm{Op}\{p(x,\xi)\langle\xi\rangle^{-r}\}$ being a classical operator of order zero.

This allows us to consider the case $r = 0$ only. Now construct another classical ΨDO $\tilde{\mathcal{P}}$ of order zero according to

$$\tilde{\mathcal{P}}(x,D) := \langle D\rangle^s \circ \mathcal{P}(x,D) \circ \langle D\rangle^{-s}, \qquad \mathcal{P}(x,D) = \langle D\rangle^{-s} \circ \tilde{\mathcal{P}}(x,D) \circ \langle D\rangle^s.$$

We are done if we succeed in showing that $\tilde{\mathcal{P}}$ maps $L^2(\mathbb{R}^n)$ continuously into itself. Hence it suffices to consider $s = 0$ only. We drop now the tilde.

For $p \in S^0_{\mathrm{cl}}$, we have the expansion

$$p(x,\xi) = \sum_{j=0}^{J} \varphi(\xi) p_j(x,\xi) + r_J(x,\xi),$$

with χ as excision function (which vanishes for $|\xi| \le 1$ and is identically equal to 1 for $|\xi| > 2$) and $p_j \in S^{-j}_{\mathrm{Hom}}(\mathbb{R}^n \times (\mathbb{R}^n \setminus 0))$. And the remainder r_J is from $S^{-J-1}_{1,0}$.

[1] SERGEI LVOVICH SOBOLEV, 1908–1989

The homogeneous components p_j admit tensor product decompositions

$$\varphi(\xi)p_j(x,\xi) = \sum_{l,m} a_{jlm}(x) \cdot Y_{lm}(\xi)\varphi(\xi)|\xi|^{-j}.$$

We wish to show, for all $u \in \mathcal{S}(\mathbb{R}^n)$, that

$$\|\mathcal{P}u\|_{L^2(\mathbb{R}^n)} \le C\,\|u\|_{L^2(\mathbb{R}^n)}\,,$$

with some constant C that depends only on a finite number of derivatives of the p_j and the remainder r_J. Then the density $\mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ allows to conclude the same inequality for $u \in L^2(\mathbb{R}^n)$.

Now we calculate like this:

$$\begin{aligned}
\|\mathcal{P}u\|_{L^2(\mathbb{R}^n)} &\le \sum_{j=0}^{J} \|\mathrm{Op}\{\varphi(\xi)p_j(x,\xi)\}u\|_{L^2(\mathbb{R}^n)} + \|\mathcal{R}_J(x,D)u\|_{L^2(\mathbb{R}^n)}\,, \\
\|\mathrm{Op}\{\varphi(\xi)p_j(x,\xi)\}u\|_{L^2(\mathbb{R}^n)} &\le \sum_{l,m} \left\| a_{jlm} \cdot \mathrm{Op}\{Y_{lm}(\xi)\varphi(\xi)|\xi|^{-j}\}u \right\|_{L^2(\mathbb{R}^n)} \\
&\le \sum_{l,m} \|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \cdot \left\|\mathrm{Op}\{Y_{lm}(\xi)\varphi(\xi)|\xi|^{-j}\}u\right\|_{L^2(\mathbb{R}^n_x)} \\
&= \sum_{l,m} \|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \cdot \left\|Y_{lm}(\cdot)\varphi(\cdot)|\cdot|^{-j}\hat{u}(\cdot)\right\|_{L^2(\mathbb{R}^n_\xi)} \qquad \Big| \quad \text{(Plancherel used)} \\
&\le \sum_{l,m} \|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \cdot \|Y_{lm}\|_{L^\infty(S^{n-1})}\,\|\hat{u}\|_{L^2(\mathbb{R}^n_\xi)} \\
&= \left(\sum_{l,m} \|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \cdot \|Y_{lm}\|_{L^\infty(S^{n-1})}\right) \|u\|_{L^2(\mathbb{R}^n_x)}\,.
\end{aligned}$$

We recall that $-\triangle_S Y_{lm} = l(l+n-2)Y_{lm}$ and $\|Y_{lm}\|_{L^2(S^{n-1})} = 1$. Now choose some even number $\sigma > n/2$. Then, by Sobolev's embedding theorem and elliptic regularity,

$$\|Y_{lm}\|_{L^\infty(S^{n-1})} \le C\,\|Y_{lm}\|_{H^\sigma(S^{n-1})} \le C\left(\left\|\triangle_S^{\sigma/2} Y_{lm}\right\|_{L^2(S^{n-1})} + \|Y_{lm}\|_{L^2(S^{n-1})}\right) \le C\,\langle l\rangle^\sigma\,.$$

And we have, for each positive integer N,

$$\|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \le \frac{C_N}{\langle l\rangle^{2N}} \left\|\triangle_S^N\, p_j(x,\cdot)\right\|_{L^2(S^{n-1})}.$$

Then we have, using $h(l,n-2) = \mathcal{O}(\langle l\rangle^{n-2})$,

$$\begin{aligned}
\sum_{l,m} \|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \cdot \|Y_{lm}\|_{L^\infty(S^{n-1})} &= \sum_{l=0}^{\infty} \sum_{m=1}^{h(l,n-2)} \|a_{jlm}\|_{L^\infty(\mathbb{R}^n)} \cdot \|Y_{lm}\|_{L^\infty(S^{n-1})} \\
&\le \sum_{l=0}^{\infty} h(l,n-2)\frac{C_N}{\langle l\rangle^{2N}} \left\|\triangle_S^N\, p_j(x,\cdot)\right\|_{L^2(S^{n-1})} \cdot C\,\langle l\rangle^\sigma \\
&\le C_N \left\|\triangle_S^N\, p_j(x,\cdot)\right\|_{L^2(S^{n-1})} \sum_{l=0}^{\infty} C\,\langle l\rangle^{n-2} \cdot \frac{1}{\langle l\rangle^{2N}} \cdot \langle l\rangle^\sigma\,.
\end{aligned}$$

Now choose N so large that $2N > n - 2 + \sigma + 1$.

It remains to discuss $\mathcal{R}_J$. We may assume $J > n$, in which case $\mathcal{R}_J$ has a Schwartz kernel $R_J(x,y)$ given by the convergent integral

$$R_J(x,y) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi} r_J(x,\xi)\,đ\xi.$$

Pick some positive integer M. Then we quickly check

$$\langle x-y\rangle^{2M} R_J(x,y) = \int_{\mathbb{R}^n_\xi} e^{\mathrm{i}(x-y)\cdot\xi}\,\langle D_\xi\rangle^{2M}\,r_J(x,\xi)\,đ\xi,$$

which allows for $|R_J(x,y)| \le \frac{C}{\langle x-y\rangle^{2M}}$, where C only depends on a finite number of symbol semi-norms of r_J. The desired estimate $\|\mathcal{R}_J u\|_{L^2(\mathbb{R}^n)} \le C\,\|u\|_{L^2(\mathbb{R}^n)}$ then follows[2] after appropriate choice of M. $\square$

Corollary 5.3. *If $p \in S^r_{\mathrm{cl}}$ is elliptic, then there is (for all s, $t \in \mathbb{R}$) a constant C such that*

$$\|u\|_{H^{s+r}(\mathbb{R}^n)} \le C\left(\|\mathcal{P}u\|_{H^s(\mathbb{R}^n)} + \|u\|_{H^t(\mathbb{R}^n)}\right).$$

If $\mathcal{P}$ has no null space, then the term $\|u\|_{H^t(\mathbb{R}^n)}$ can be dropped.

Proof. We have a parametrix $\mathcal{Q} \in \Psi^{-r}_{\mathrm{cl}}$ and a smoothing remainder $\mathcal{R}$ such that $\mathcal{Q}\circ\mathcal{P} = \mathrm{id} + \mathcal{R}$, hence

$$u = \mathcal{Q}\circ\mathcal{P}u - \mathcal{R}u.$$

$\square$

Useful is of course the choice of a t that is much smaller than $s+r$ and s.

5.2 The Case of $\Psi^m_{0,0}$ on $\mathbb{R}^n$

We present here a famous result of Calderón[3] and Vaillancourt[4] from 1970 [8], because its proof is exceptionally beautiful (meaning that it incorporates many ideas).

Theorem 5.4. *Let $p \in S^0_{0,0}$ be a pseudodifferential symbol on $\mathbb{R}^n$ with global symbol estimates. Then p generates a ΨDO $\mathcal{P}$ that maps continuously from $L^2(\mathbb{R}^n)$ into $L^2(\mathbb{R}^n)$.*

The proof combines many ideas from various branches of mathematics. We begin with functional analysis.

Lemma 5.5. *Let $(Z, \Sigma, \mathrm{d}z)$ be a measure space, and let us be given a family $\{A_z : z \in Z\}$ of operators A_z that map from $L^2(\mathbb{R}^n)$ continuously into $L^2(\mathbb{R}^n)$, with uniformly bound on the the operator norm:*

$$\exists M_0 > 0 : \ \forall z \in Z : \ \|A_z\|_{\mathrm{op}} \le M_0.$$

We further assume that the map $z \mapsto A_z$ is weakly measurable, which means that for all φ, $\psi \in L^2(\mathbb{R}^n)$, the function $z \mapsto \langle A_z\varphi, \psi\rangle$ is measurable.

Now suppose that there is a function $H\colon Z\times Z \to [0,\infty)$ with

- $\|A_z A^*_{z'}\|_{\mathrm{op}} \le H^2(z,z')$ *and* $\|A^*_z A_{z'}\|_{\mathrm{op}} \le H^2(z,z')$,
- $(\mathcal{H}w)(z) := \int_{z'\in Z} H(z,z')w(z')\,\mathrm{d}z'$ *is an operator $\mathcal{H}$ that maps from $L^2(Z)$ into itself with norm M.*

Then $\mathcal{A} := \int_{z\in Z} A_z\,\mathrm{d}z$ is an operator from $L^2(\mathbb{R}^n)$ into itself with norm $\|\mathcal{A}\|_{\mathrm{op}} \le M$.

Proof. We note that the operator $\mathcal{H}^2$ has integral kernel

$$H_2(z,z') := \int_{z_1\in Z} H(z,z_1)H(z_1,z')\,\mathrm{d}z_1,$$

and analogously, then the operator $\mathcal{H}^j$ (with $j \ge 2$) has integral kernel

$$H_j(z,z') := \int_{z_1\in Z}\cdots\int_{z_{j-1}\in Z} H(z,z_1)H(z_1,z_2)\cdots H(z_{j-2},z_{j-1})H(z_{j-1},z')\,\mathrm{d}z_{j-1}\,\mathrm{d}z_{j-2}\cdots\mathrm{d}z_1.$$

Now pick some $m \in \mathbb{N}$ with $m \ge 2$, and points $z_1, \ldots, z_{2m} \in Z$. Define a number

$$T_m := \left\|A_{z_1}A^*_{z_2}A_{z_3}A^*_{z_4}\ldots A_{z_{2m-1}}A^*_{z_{2m}}\right\|_{\mathrm{op}},$$

[2] Figuring out the details is an exercise everybody should have done once in life. I have already, so ...
[3] Alberto Calderón, 1920–1998
[4] Rémi Vaillancourt, 1934–2015

which we now estimate twice:

$$T_m \leq \left\|A_{z_1}A^*_{z_2}\right\|_{\mathrm{op}} \left\|A_{z_3}A^*_{z_4}\right\|_{\mathrm{op}} \left\|A_{z_5}A^*_{z_6}\right\|_{\mathrm{op}} \cdots \left\|A_{z_{2m-1}}A^*_{z_{2m}}\right\|_{\mathrm{op}},$$
$$T_m \leq \left\|A_{z_1}\right\|_{\mathrm{op}} \left\|A^*_{z_2}A_{z_3}\right\|_{\mathrm{op}} \left\|A^*_{z_4}A_{z_5}\right\|_{\mathrm{op}} \left\|A^*_{z_6}A_{z_7}\right\|_{\mathrm{op}} \cdots \left\|A^*_{z_{2m-2}}A_{z_{2m-1}}\right\|_{\mathrm{op}} \left\|A^*_{z_{2m}}\right\|_{\mathrm{op}}.$$

Multiplying both estimates, we then obtain

$$\begin{aligned} T_m^2 &\leq \left\|A_{z_1}\right\|_{\mathrm{op}} \left\|A_{z_1}A^*_{z_2}\right\|_{\mathrm{op}} \left\|A^*_{z_2}A_{z_3}\right\|_{\mathrm{op}} \left\|A_{z_3}A^*_{z_4}\right\|_{\mathrm{op}} \left\|A^*_{z_4}A_{z_5}\right\|_{\mathrm{op}} \cdots \left\|A_{z_{2m-1}}A^*_{z_{2m}}\right\|_{\mathrm{op}} \left\|A^*_{z_{2m}}\right\|_{\mathrm{op}} \\ &\leq M_0 H^2(z_1,z_2) H^2(z_2,z_3) H^2(z_3,z_4) H^2(z_4,z_5) \cdots H^2(z_{2m-1},z_{2m}) M_0. \end{aligned}$$

Now pick a measurable set $\Omega \subset Z$ with characteristic function χ_Ω and finite measure $S_\Omega := \int_\Omega 1\,\mathrm{d}z$. Then we consider

$$\begin{aligned} &\left\| \left(\left(\int_\Omega A_z \,\mathrm{d}z \right) \left(\int_\Omega A_z \,\mathrm{d}z \right)^* \right)^m \right\|_{\mathrm{op}}^{1/m} \\ &\leq \left(\int_{z_1\in\Omega} \int_{z_2\in\Omega} \cdots \int_{z_{2m}\in\Omega} \left\|A_{z_1}A^*_{z_2}A_{z_3}A^*_{z_4}\cdots A^*_{z_{2m}}\right\|_{\mathrm{op}} \,\mathrm{d}z_{2m}\,\mathrm{d}z_{2m-1}\ldots\,\mathrm{d}z_1 \right)^{1/m} \\ &\leq \left(M_0 \int_{z_1\in\Omega} \int_{z_{2m}\in\Omega} \left(\int_{(z_2,\ldots,z_{2m-1})\in\Omega^{2m-2}} H(z_1,z_2)\ldots H(z_{2m-1},z_{2m}) \,\mathrm{d}(z_2,\ldots,z_{2m-1}) \right) \mathrm{d}z_{2m}\,\mathrm{d}z_1 \right)^{1/m} \\ &= \left(M_0 \int_{z_1\in\Omega} \int_{z_{2m}\in\Omega} H_{2m-1}(z_1,z_{2m}) \,\mathrm{d}z_1\,\mathrm{d}z_{2m} \right)^{1/m} \\ &= \left(M_0 \int_{Z_{z_1}} \int_{Z_{z_{2m}}} \chi_\Omega(z_1) H_{2m-1}(z_1,z_{2m}) \chi_\Omega(z_{2m}) \,\mathrm{d}z_{2m}\,\mathrm{d}z_1 \right)^{1/m} \\ &= \left(M_0 \int_{Z_{z_1}} \chi_\Omega(z_1) \left(\mathcal{H}^{2m-1}\chi_\Omega\right)(z_1) \,\mathrm{d}z_1 \right)^{1/m} \\ &\leq \left(M_0 S_\Omega M^{2m-1} S_\Omega \right)^{1/m}. \end{aligned}$$

Recall that for each bounded self-adjoint operator $\mathcal{B}$, we have $\|\mathcal{B}\|_{\mathrm{op}} = \lim_{m\to\infty} \|\mathcal{B}^m\|_{\mathrm{op}}^{1/m}$. This gives

$$\left\| \left(\int_\Omega A_z \,\mathrm{d}z \right) \left(\int_\Omega A_z \,\mathrm{d}z \right)^* \right\|_{\mathrm{op}} \leq M^2.$$

A consequence then is

$$\left\| \int_\Omega A_z \,\mathrm{d}z \right\|_{\mathrm{op}} \leq M.$$

But the RHS is independent of Ω, which finally allows us to deduce that

$$\left\| \int_Z A_z \,\mathrm{d}z \right\|_{\mathrm{op}} \leq M.$$

□

The key idea was to decompose $\mathcal{A}$ into components A_z in the sense of $\mathcal{A} = \int_Z A_z \,\mathrm{d}z$, such that A_z and $A^*_{z'}$ are "basically orthogonal" for z and z' "far away from each other", in the sense that $A_z A^*_{z'}$ and $A^*_z A_{z'}$ have "small norms". The proof of the main result will make this clearer.

Proof of the Calderón–Vaillancourt Theorem. Consider first the case $n = 1$. We know that

$$\forall(\alpha,\beta)\colon\ \exists C_{\alpha\beta}\colon\ \forall(x,\xi)\in\mathbb{R}^2\colon\ |\partial_x^\alpha \partial_\xi^\beta p(x,\xi)| \leq C_{\alpha\beta}.$$

We define an auxiliary function ψ on $\mathbb{R}$,

$$\psi(t) = \begin{cases} \frac{1}{2}t^2 e^{-t} & : t \geq 0, \\ 0 & t < 0. \end{cases}$$

Observe that this can be written as $\psi(t) = H(t) \cdot \frac{1}{2}t^2 e^{-t}$, with $H(t)$ being the Heaviside function, and a distribution H and a smooth function have been multiplied, resulting in a distribution again. In such a setting, the product rule of differentiation is valid[5], and therefore

$$\begin{aligned}(1+\partial_t)\psi(t) &= H(t)\cdot(1+\partial_t)\frac{1}{2}t^2e^{-t} + \Big(\partial_t H(t)\Big)\cdot\frac{1}{2}t^2e^{-t} \\ &= H(t)\cdot te^{-t} + \delta(t)\cdot\frac{1}{2}t^2e^{-t} \\ &= H(t)\cdot te^{-t} + 0.\end{aligned}$$

And we apply this reasoning once more,

$$\begin{aligned}(1+\partial_t)^2\psi(t) &= H(t)\cdot(1+\partial_t)te^{-t} + \Big(\partial_t H(t)\Big)\cdot te^{-t} \\ &= H(t)\cdot e^{-t} + \delta(t)\cdot te^{-t} \\ &= H(t)\cdot e^{-t} + 0.\end{aligned}$$

And once more:

$$\begin{aligned}(1+\partial_t)^3\psi(t) &= H(t)\cdot(1+\partial_t)e^{-t} + \Big(\partial_t H(t)\Big)\cdot e^{-t} \\ &= 0 + \delta(t)\cdot e^{-t} \\ &= \delta(t).\end{aligned}$$

Now we define another symbol $g = g(x,\xi)$ on $\mathbb{R}^2$:

$$g(x,\xi) := (1+\partial_x)^3(1+\partial_\xi)^3 p(x,\xi).$$

Then we have $g \in S^0_{0,0}$, and by appealing to a variant of (2.6), we have

$$p(x,\xi) = \iint_{\mathbb{R}^2_{y,\eta}} g(y,\eta)\psi(x-y)\psi(\xi-\eta)\,\mathrm{d}y\,\mathrm{d}\eta.$$

Now we plug this representation of p into the formula for $\mathcal{P}$:

$$\begin{aligned}(\mathcal{P}f)(x) &= \int_{\mathbb{R}_\xi} e^{\mathrm{i}x\xi} p(x,\xi)\hat{f}(\xi)\,đ\xi \\ &= \iint_{\mathbb{R}^2_{y,\eta}} g(y,\eta)\left(\int_{\mathbb{R}_\xi} e^{\mathrm{i}x\xi}\psi(x-y)\psi(\xi-\eta)\hat{f}(\xi)\,đ\xi\right)\mathrm{d}y\,\mathrm{d}\eta \\ &= \iint_{\mathbb{R}^2_{y,\eta}} g(y,\eta)\cdot(A_{y\eta}f)(x)\,\mathrm{d}y\,\mathrm{d}\eta,\end{aligned}$$

where we have defined

$$(A_{y\eta}f)(x) := \int_{\mathbb{R}_\xi} e^{\mathrm{i}x\xi}\psi(x-y)\psi(\xi-\eta)\hat{f}(\xi)\,đ\xi.$$

We intend to apply the previous Lemma with $Z = \mathbb{R}^2_{y\eta}$ and $z = (y,\eta)$.
First we remark $0 \leq \psi(t) \leq 2e^{-2}$, and therefore

$$\|A_{y\eta}f\|_{L^2(\mathbb{R})} \leq 4e^{-4}\|f\|_{L^2(\mathbb{R})}, \quad \text{hence} \quad \|A_z\|_{\mathrm{op}} \leq 4e^{-4} =: M_0.$$

[5] The reason being: multiplying by a fixed smooth function is a continuous map $\mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(\mathbb{R})$, and also taking the derivative is a continuous map $\mathcal{D}'(\mathbb{R}) \to \mathcal{D}'(\mathbb{R})$. Now the space $\mathcal{D}(\mathbb{R})$ of smooth functions is dense in the distribution space $\mathcal{D}'(\mathbb{R})$. It remains to observe that the product rule obviously holds for both factors coming from $\mathcal{D}(\mathbb{R})$. Another way consists in stoical application of the definition of the two operations.

Second we have to estimate $\left\|A_{y\eta}A^*_{y'\eta'}\right\|_{\mathrm{op}}$ and $\left\|A^*_{y\eta}A_{y'\eta'}\right\|_{\mathrm{op}}$.

The operator $A_{y\eta}$ is a ΨDO with non-smooth symbol $(x,\xi)\mapsto\psi(x-y)\psi(\xi-\eta)$. Therefore, $A^*_{y'\eta'}$ is again a ΨDO which we can express using the amplitude $(x,t,\xi)\mapsto\overline{\psi(t-y')\psi(\xi-\eta')}$. But ψ is real, hence

$$\begin{aligned}(A^*_{y'\eta'}f)(x) &= \int_{\mathbb{R}_\xi}\int_{\mathbb{R}^t} e^{\mathrm{i}(x-t)\xi}\psi(t-y')\psi(\xi-\eta')f(t)\,\mathrm{d}t\,\text{đ}\xi\\ &= \int_{\mathbb{R}_\xi} e^{\mathrm{i}x\xi}\psi(\xi-\eta')\left(\int_{\mathbb{R}_t} e^{-\mathrm{i}t\xi}\psi(t-y')f(t)\,\mathrm{d}t\right)\text{đ}\xi\\ &= \mathcal{F}^{-1}_{\xi\to x}\Big\{\psi(\xi-\eta')\cdot\mathcal{F}_{t\to\xi}\{\psi(t-y')f(t)\}(\xi)\Big\}(x),\end{aligned}$$

in particular $A^*_{y'\eta'}f$ is the inverse Fourier transform of something for which we have a formula. Now we build $A_{y\eta}A^*_{y'\eta'}f$:

$$\begin{aligned}\left(A_{y\eta}A^*_{y'\eta'}f\right)(x) &= \int_{\mathbb{R}_\xi} e^{\mathrm{i}x\xi}\psi(x-y)\psi(\xi-\eta)\widehat{\left(A^*_{y'\eta'}f\right)}(\xi)\,\text{đ}\xi\\ &= \int_{\mathbb{R}_\xi} e^{\mathrm{i}x\xi}\psi(x-y)\psi(\xi-\eta)\psi(\xi-\eta')\mathcal{F}_{t\to\xi}\left\{\psi(t-y')f(t)\right\}(\xi)\,\text{đ}\xi\\ &= \int_{\mathbb{R}_\xi}\int_{\mathbb{R}_t} e^{\mathrm{i}(x-t)\xi}\psi(x-y)\psi(\xi-\eta)\psi(\xi-\eta')\psi(t-y')f(t)\,\mathrm{d}t\,\text{đ}\xi\\ &= \int_{\mathbb{R}^t} K(x,t)f(t)\,\mathrm{d}t,\end{aligned}$$

with a kernel function K given by

$$\begin{aligned}K(x,t) &:= \int_{\mathbb{R}_\xi} e^{\mathrm{i}(x-t)\xi}\psi(x-y)\psi(\xi-\eta)\psi(\xi-\eta')\psi(t-y')\,\text{đ}\xi\\ &= \psi(x-y)\psi(t-y')\int_{\mathbb{R}_\xi} e^{\mathrm{i}(x-t)\xi}\psi(\xi-\eta)\psi(\xi-\eta')\,\text{đ}\xi.\end{aligned}$$

This kernel function K gives rise to an operator $\mathcal{K}\colon L^2(\mathbb{R})\to L^2(\mathbb{R})$ which we then will estimate by a function $H^2(z,z')$, which then in turn will give rise to one more operator $\mathcal{H}\colon L^2(\mathbb{R}^2_{z'})\to L^2(\mathbb{R}^2_z)$, for which we will then have to find again a norm (ending the proof).

We begin our work with $K(x,t)$. Note that $\psi(\sigma)=0$ for $\sigma\le 0$. Define $s:=x-t$ and $\sigma:=\xi-\eta$. Then $\xi=\sigma+\eta$ and

$$K(x,t)=\psi(x-y)\psi(t-y')e^{\mathrm{i}s\eta}\int_{\sigma=0}^{\infty} e^{\mathrm{i}s\sigma}\psi(\sigma)\psi(\sigma+\eta-\eta')\,\text{đ}\sigma.$$

Case 1: $\eta\ge\eta'$. Then we get

$$\begin{aligned}K(x,t) &= \psi(x-y)\psi(t-y')e^{\mathrm{i}s\eta}\int_{\sigma=0}^{\infty} e^{\mathrm{i}s\eta}\frac{\sigma^2(\sigma+\eta-\eta')^2}{4}e^{-2\sigma}e^{-(\eta-\eta)'}\,\text{đ}\sigma\\ &= \psi(x-y)\psi(t-y')\frac{e^{\mathrm{i}s\eta}e^{-|\eta-\eta'|}}{4}\int_{\sigma=0}^{\infty} e^{(\mathrm{i}s-2)\sigma}\sigma^2(\sigma+\eta-\eta')^2\,\text{đ}\sigma\\ &= \psi(x-y)\psi(t-y')\frac{e^{\mathrm{i}s\eta}e^{-|\eta-\eta'|}}{4}D_s^2(D_s+\eta-\eta')^2\int_{\sigma=0}^{\infty} e^{(\mathrm{i}s-2)\sigma}\,\text{đ}\sigma\\ &= \psi(x-y)\psi(t-y')\frac{e^{\mathrm{i}s\eta}e^{-|\eta-\eta'|}}{4}D_s^2(D_s+\eta-\eta')^2\frac{1}{2-\mathrm{i}s},\end{aligned}$$

from which we then find

$$\begin{aligned}|K(x,t)| &\le \text{const.}\,\psi(x-y)\psi(t-y')e^{-|\eta-\eta'|}(1+|x-t|)^{-3}(|\eta-\eta'|+1)^2\\ &\le C\psi(x-y)\psi(t-y')e^{-|\eta-\eta'|/2}(1+|x-t|)^{-3}.\end{aligned}$$

Case 2: $\eta<\eta'$. Then a similar calculation leads to the same estimate.

This estimate of K is strong enough to get the operator norm of $\mathcal{K} = A_{y\eta}A^*_{y'\eta'}$ under control:

$$\|(A_{y\eta}A_{y'\eta'}f)(x)\|^2_{L^2(\mathbb{R}_x)} \leq \left(\int_{\mathbb{R}^2_{xt}} |K(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t\right) \|f\|^2_{L^2(\mathbb{R}_x)},$$

where we have used Lemma 2.61. Now we consider

$$\begin{aligned}\int_{\mathbb{R}^2_{xt}} |K(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t &\leq Ce^{-|\eta-\eta'|}\int_{\mathbb{R}^2_{xt}} \frac{\psi^2(x-y)\psi^2(t-y')}{(1+|x-t|^2)^3}\,\mathrm{d}x\,\mathrm{d}t \\ &= Ce^{-|\eta-\eta'|}\int_{p=0}^\infty\int_{q=0}^\infty \frac{\psi^2(p)\psi^2(q)}{(1+|(y-y')+(p-q)|^2)^3}\,\mathrm{d}p\,\mathrm{d}q,\end{aligned}$$

where we have substituted $x = y+p$ and $t = y'+q$. Note that $\psi \equiv 0$ for negative arguments.

Now is a good moment to present PEETRE's inequality: for all a, $b \in \mathbb{R}^n$, and all $t \in \mathbb{R}$, it holds

$$\left(\frac{1+|a|^2}{1+|b|^2}\right)^t \leq 2^{|t|}(1+|a-b|^2)^{|t|}.$$

We then have

$$\frac{1}{(1+|b|^2)^3} \leq \frac{8}{(1+|a|^2)^3}\cdot(1+|a-b|^2)^3,$$

hence

$$\frac{1}{(1+|(y-y')+(p-q)|^2)^3} \leq \frac{8}{(1+|y-y'|^2)^3}(1+|p-q|^2)^3,$$

which then has the consequence

$$\begin{aligned}\int_{\mathbb{R}^2_{xt}} |K(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t &\leq Ce^{-|\eta-\eta'|}(1+|y-y'|^2)^{-3}\int_{p=0}^\infty\int_{q=0}^\infty (1+|p-q|^2)^3\psi^2(p)\psi^2(q)\,\mathrm{d}p\,\mathrm{d}q \\ &\leq Ce^{-|\eta-\eta'|}(1+|y-y'|)^{-6} \\ &\leq C(1+|(y,\eta)-(y',\eta')|)^{-6}.\end{aligned}$$

With $z = (y,\eta)$ and $z' = (y',\eta')$ we therefore have

$$\|A_zA^*_z\|_{\mathrm{op}} \leq C(1+|z-z'|)^{-3} =: H(z,z').$$

Now we should consider the operator norm $\|A^*_zA_{z'}\|_{\mathrm{op}}$. It turns out that the above calculation can be replicated to a large extent, but there is a twist: we should translate from the x-world into the ξ-world, using Plancherel's identity. The key step is to verify that the operator that maps $\hat{f} = \hat{\xi}$ into $(A^*_{y\eta}A_{y'\eta'}f)\hat{}(\xi)$ has integral kernel

$$\tilde{K}(\xi,\tau) = \psi(\xi-\eta)\psi(\tau-\eta')\int_{\mathbb{R}_x} e^{-\mathrm{i}x(\xi-\tau)}\psi(x-y)\psi(x-y')\,\mathrm{d}x,$$

hence $\tilde{K}$ can be obtained from K by swapping the arguments and taking the complex conjugate.

The final step is to consider the operator $\mathcal{H}$ with the kernel function H. This can be handled by either a variant of Lemma 2.48 with $k=2$, $p_1 = 1$ and $p_2 = q = 2$, or (2.10). The Theorem of Calderón and Vaillancourt is proved, at least for $n = 1$. And for general n, we build g according to

$$g(x,\xi) := (1+\partial_{x_1})^3(1+\partial_{\xi_1})^3\dots(1+\partial_{x_n})^3(1+\partial_{\xi_n})^3 p(x,\xi),$$

and now we observe that the different directions along the various coordinate axes do not interact. □

5.3 Summary

5.3.1 What has been Obtained So Far

Suppose $u \in \mathcal{E}'(\mathbb{R}^n)$ solves $\mathcal{P}u = f$, with $\mathcal{P}$ being an elliptic ΨDO, whose symbol p has global symbol estimates in $S^m_{1,0}$. Then regularity of f can be transferred into regularity of u. We explain this as follows.

The operator $\mathcal{P}$ has a parametrix $\mathcal{P}^\sharp$ which has global symbol estimates in $S^{-m}_{1,0}$, and $\mathcal{P}^\sharp$ is properly supported. Being a parametrix means $\mathcal{P}^\sharp \circ \mathcal{P} = \mathrm{id} - \mathcal{R}$, with $\mathcal{R}$ being a ΨDO with symbol in $S^{-\infty}$, and $\mathcal{R}$ maps each Sobolev space into any Sobolev space.

Now what can be said about regularities ? From $u \in \mathcal{E}'(\mathbb{R}^n)$ we have $f \in \mathcal{D}'(\mathbb{R}^n)$. It is an easy exercise to show $\mathcal{E}'(\mathbb{R}^n) \subset \cup_{t\in\mathbb{R}} H^t(\mathbb{R}^n)$, so u belongs to *some* Sobolev space $H^t(\mathbb{R}^n)$; perhaps $t \ll -1$. Suppose $f \in H^s(\mathbb{R}^n)$, for some $s \in \mathbb{R}$. Then we find

$$u = \mathrm{id}\, u = \mathcal{P}^\sharp \mathcal{P} u + \mathcal{R} u = \mathcal{P}^\sharp f + \mathcal{R} u,$$

and now we have $\mathcal{P}^\sharp f \in H^{s+m}(\mathbb{R}^n)$, whereas $\mathcal{R}u$ belongs to any Sobolev space we want. Therefore $u \in H^{s+m}(\mathbb{R}^n)$, and if the regularity of f improves, then also the regularity of u.

Microlocal analysis enables us to prove the regularity of a solution to an elliptic problem, provided the existence of that solution has already been established somehow, together with some starting regularity of that solution, as for instance $u \in \mathcal{E}'(\mathbb{R}^n)$.

We have no method of proving that the problem $\mathcal{P}u = f$ possesses at least one solution. And the condition that $\mathcal{P}$ shall be elliptic seems restrictive, so it looks like we have spent 60 pages of work for little output; but in fact we have obtained more, and we explain it in the next section.

5.3.2 What does "Micro-Local" Mean ?

The mapping properties of a ΨDO can be **localised**, which means that we disregard all those points x we are not interested in.

Definition 5.6. *Let $u \in \mathcal{D}'(\Omega)$ and $x_0 \in \Omega$, and $s \in \mathbb{R}$. We say that $u \in H^s_{x_0}$ provided that there is a neighbourhood $U \Subset \Omega$ of x_0 such that $\varphi u \in H^s(\Omega)$ for each $\varphi \in C_0^\infty(U)$. We read this as "u has regularity H^s at the point x_0".*

Then the mapping properties **localise** as well:

Lemma 5.7. *Let $u \in \mathcal{D}'(\Omega)$ belong to $H^s_{x_0}$ for some $x_0 \in \Omega$ and some $s \in \mathbb{R}$, and let $\mathcal{P}$ be a classical operator with symbol from $S^m_{\mathrm{cl}}(\Omega \times \mathbb{R}^n)$, with proper support.*
Then $\mathcal{P}u \in H^{s-m}_{x_0}$.

Proof. We know that there is some neighbourhood $U \Subset \Omega$ of x_0 such that $\varphi u \in H^s(\Omega)$ for all $\varphi \in C_0^\infty(U)$. Choose some neighbourhood $V \Subset U$ of x_0. We wish to show $\psi \mathcal{P} u \in H^{s-m}(\Omega)$, for all $\psi \in C_0^\infty(V)$.

Now $\mathcal{P}$ has proper support, which means that there is some set $K \Subset \Omega$ such that $\mathcal{P}u$ (when we evaluate this inside V) does not depend on values $u(x)$ with $x \notin K$. Choose some $\chi \in C_0^\infty(\Omega)$ with $\chi \equiv 1$ on K. Then we have $\mathcal{P}u = \mathcal{P}(\chi u)$ as identity inside V, and now $\chi u \in \mathcal{E}'(\Omega)$. Next we choose one function $\varphi \in C_0^\infty(U)$ with $\varphi \equiv 1$ on V. We write

$$\psi \mathcal{P} u = \psi \mathcal{P}(\chi u) = \psi \mathcal{P}(\chi \varphi u) + \psi \mathcal{P}(\chi \cdot (1-\varphi) u), \quad \text{in V}.$$

By assumption $u \in H^s_{x_0}$ we have $\varphi u \in H^s(\Omega)$, then also $\chi \varphi u \in H^s(\Omega)$ with compact support, hence $\mathcal{P}(\chi \varphi u) \in H^{s-m}(\Omega)$ with compact support.

And for the second item, we consider $1 - \varphi$ as a ΨDO. Then the expression $\psi \mathcal{P} \circ (1 - \varphi)$ is a composition of two properly supported ΨDOs, and all terms in its asymptotic expansion drop out, owing to $\varphi \equiv 1$ on $\operatorname{supp} \psi$. Only the remainder term in that asymptotic expansion survives, and it is a smoothing operator that maps from $\mathcal{E}'(\Omega)$ into $\mathcal{D}(\Omega)$. The result is $\psi \mathcal{P}(\chi \cdot (1-\varphi) u) \in \mathcal{D}(\Omega)$ with support in V. □

We can also **localise** the concept of ellipticity:

Definition 5.8. *We say that an operator $\mathcal{P}$ with symbol in $S^m_{\rm cl}(\Omega \times \mathbb{R}^n)$ is* locally elliptic at $x_0 \in \Omega$ *if there is some constant $c_0 > 0$ such that its principal symbol p_m satisfies*

$$|p_m(x_0, \xi)| \geq c_0 |\xi|^m, \quad \forall \xi \in \mathbb{R}^n \setminus 0.$$

By continuity and compactness of the unit sphere $\{\xi \in \mathbb{R}^n : |\xi| = 1\}$, local ellipticity at a point x_0 implies local ellipticity in a neighbourhood of that point.

Now let us have a look at the construction of the parametrix, and we convince ourselves that this construction can also be **localised**:

Lemma 5.9. *Let $u \in \mathcal{D}'(\Omega)$, and let $\mathcal{P}$ be a classical operator of order m, properly supported, locally elliptic at $x_0 \in \Omega$.*

Then the following holds: if $\mathcal{P}u \in H^{s-m}_{x_0}$ for some $s \in \mathbb{R}$, then $u \in H^s_{x_0}$ for that s.

If we look a bit longer at our previous calculations, we observe that we can localise not only in the space variable x, but also in the co-direction variable ξ.

> **Localising** means to cut-off non-interesting positions x.
> **Micro-localising** means to cut-off non-interesting positions x and to cut-off non-interesting co-directions ξ as well.

We **micro-localise** Sobolev spaces:

Definition 5.10. *Let $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus 0)$, and let $s \in \mathbb{R}$. We say that $u \in \mathcal{D}'(\Omega)$ belongs to $H^s_{x_0,\xi_0}$ if there is a conic neighbourhood $U \times V$ of (x_0, ξ_0) such that*

$$\int_{V_\xi} |(\varphi u)\hat{\ }(\xi)|^2 \langle \xi \rangle^{2s} \, \mathrm{d}\xi < \infty$$

for each $\varphi \in C_0^\infty(U)$. We read this as "$u$ has micro-local H^s regularity at (x_0, ξ_0)".

Exercise: *Take $n = 2$ and consider the function*

$$u(x_1, x_2) = \begin{cases} 0 & : \ x_1 < 0, \\ 1 & : \ x_1 \geq 0. \end{cases}$$

For each $x_0 \in \mathbb{R}^2$, determine those $s \in \mathbb{R}$ for which $u \in H^s_{x_0}$.

For each $(x_0, \xi_0) \in \mathbb{R}^2 \times (\mathbb{R}^2 \setminus 0)$, determine those $s \in \mathbb{R}$ for which $u \in H^s_{x_0,\xi_0}$.

Then the mapping properties can be **micro-localised** as well:

Lemma 5.11. *Let $u \in \mathcal{D}'(\Omega)$ belong to $H^s_{x_0,\xi_0}$ for some $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus 0)$ and some $s \in \mathbb{R}$, and let $\mathcal{P}$ be a classical operator with symbol from $S^m_{\rm cl}(\Omega \times \mathbb{R}^n)$, with proper support.*

Then $\mathcal{P}u \in H^{s-m}_{x_0,\xi}$.

The proof is quite similar to the proof of Lemma 5.7.

Next we **micro-localise** the concept of ellipticity:

Definition 5.12. *We say that an operator $\mathcal{P}$ with symbol in $S^m_{\rm cl}(\Omega \times \mathbb{R}^n)$ is* micro-locally elliptic at $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus 0)$ *if there is some constant $c_0 > 0$ such that its principal symbol satisfies*

$$|p_m(x_0, \xi_0)| \geq c_0 |\xi_0|^m.$$

By continuity and homogeneity of the principal symbol, we then have $|p_m(x, \xi)| \geq \frac{1}{2} c_0 |\xi|^m$ for all (x, ξ) from a conic neighbourhood of (x_0, ξ_0).

And also the construction of the parametrix can be **micro-localised**:

Lemma 5.13. *Let $u \in \mathcal{D}'(\Omega)$, and let $\mathcal{P}$ be a classical operator of order m, properly supported, micro-locally elliptic at $(x_0, \xi_0) \in \Omega \times (\mathbb{R}^n \setminus 0)$.*

Then the following holds: if $\mathcal{P}u \in H^{s-m}_{x_0,\xi_0}$ for some $s \in \mathbb{R}$, then $u \in H^s_{x_0,\xi_0}$ for that s.

Since we have already skipped the proof of Lemma 5.9, we omit this proof as well.

Example: *The wave equation operator $\partial_t^2 - c^2 \triangle$, with a constant $c > 0$, is micro-locally elliptic of order 2 in most directions (but not all). If we know $(\partial_t^2 - c^2 \triangle)u \in L^2_{\mathrm{loc}}(\mathbb{R}^{1+n})$, then we can deduce the micro-local regularity $u \in H^2_{t_0,x_0,\tau_0,\xi_0}$, for all $(t_0, x_0) \in \mathbb{R}^{1+n}$, and most directions $(\tau_0, \xi_0) \in \mathbb{R}^{1+n} \setminus 0$. The details are left to the reader.*

The current situation refers to the elliptic setting, and basically it means: if $\mathcal{P}u$ is microlocally Sobolev-regular at (x_0, ξ_0), then u is microlocally Sobolev-regular at (x_0, ξ_0), and the Sobolev smoothness parameters translate as expected. And since each microlocally elliptic operator has a microlocal parametrix, which is again microlocally elliptic, this description of the regularities is sharp.

Now we present some result that is mostly relevant to hyperbolic situations — a famous theorem of Hörmander about the propagation of singularities. Let $\mathcal{P}$ be a classical ΨDO of order $m \geq 0$, and let p_m be its principal symbol. To this p_m, we consider its HAMILTONian[6] vector field H_p in the phase space $\Omega \times (\mathbb{R}^n \setminus 0)$:

$$H_{p_m} := \sum_{j=1}^n \left(\frac{\partial p_m}{\partial \xi_j} \cdot \frac{\partial}{\partial x_j} - \frac{\partial p_m}{\partial x_j} \cdot \frac{\partial}{\partial \xi_j} \right).$$

Each of its integral curves $(x(t), \xi(t))$ (with t being the parametrisation variable of the curve) is called *bi-characteristic curve*. By definition, this means that the functions $x(t)$ and $\xi(t)$ solve the *Hamilton system*

$$\frac{\mathrm{d}x_j}{\mathrm{d}t} = \frac{\partial p_m(x(t), \xi(t))}{\partial \xi_j}, \qquad j = 1, \dots, n,$$
$$\frac{\mathrm{d}\xi_j}{\mathrm{d}t} = -\frac{\partial p_m(x(t), \xi(t))}{\partial x_j}, \qquad j = 1, \dots, n.$$

Along each such curve, p_m is a constant function of t, because:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} p_m(x(t), \xi(t)) &= \Big(\nabla_x p_m(x(t), \xi(t))\Big) \cdot \dot{x} + \Big(\nabla_\xi p_m(x(t), \xi(t))\Big) \cdot \dot{\xi}(t) \\ &= \Big(\nabla_x p_m(x(t), \xi(t))\Big) \cdot \Big(\nabla_\xi p_m(x(t), \xi(t))\Big) + \Big(\nabla_\xi p_m(x(t), \xi(t))\Big) \cdot \Big(- \nabla_x p_m(x(t), \xi(t))\Big) \\ &= 0.\end{aligned}$$

We are interested in those bi-characteristic curves along which p_m equals zero everywhere (in the elliptic case, they do not exist, because p_m never vanishes except for $\xi = 0$).

Theorem 5.14 (Hörmander). *Let $p \in S^m_{\mathrm{cl}}(\Omega \times \mathbb{R}^n)$, with $m \geq 0$. Suppose that the principal symbol p_m is real-valued everywhere. Let $\gamma \colon [0, T] \to \Omega \times (\mathbb{R}^n \setminus 0)$ be a zero-bi-characteristic curve of p_m, and let Γ be its trace in the phase-space $\Omega \times (\mathbb{R}^n \setminus 0)$. Furthermore, assume $p(x, D)u = f \in H^s_{x,\xi}$ micro-locally, for all $(x, \xi) \in \Gamma$.*

Then the following holds: if $u \in H^{s+m-1}_{\gamma(t_0)}$, micro-locally for one $t_0 \in [0, T]$, then $u \in H^{s+m-1}_{\gamma(t)}$, micro-locally for all *$t \in [0, T]$.*

Let us bring this into context: if (x, ξ) is *not* on any zero-bi-characteristic curve, then $p_m(x, \xi) \neq 0$, hence we have micro-local ellipticity of p_m, and then $u \in H^{s+m}_{x,\xi}$ anyway. This is the boring case.

Now comes the interesting case: (x, ξ) *is* on a certain zero-bi-characteristic curve. Then the singular behaviour of the solution u propagates along this curve.

[6] SIR WILLIAM ROWAN HAMILTON, 1805–1865

An example hopefully brings clarity: consider the wave operator $\partial_t^2 - c^2 \triangle$ in $\mathbb{R}^{1+n}$, and now we call the time variable x_0, and its co-variable becomes ξ_0. Then the operator is

$$\mathcal{P} = \partial_{x_0}^2 - c^2 \sum_{j=1}^n \partial_{x_j}^2, \qquad p(x,\xi) = \sigma(\mathcal{P}) = -\xi_0^2 + c^2(\xi_1^2 + \dots + \xi_n^2), \qquad c > 0.$$

This is a classical operator of order $m = 2$, and indeed its principal symbol is real-valued everywhere. The bi-characteristic curves are $(x_0(t), \dots, x_n(t), \xi_0(t), \dots, \xi_n(t))$, which are required to satisfy

$$\begin{aligned}
\frac{\mathrm{d}x_0(t)}{\mathrm{d}t} &= \frac{\partial p_2(x(t),\xi(t))}{\partial \xi_0} = -2\xi_0(t), \\
\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} &= \frac{\partial p_2(x(t),\xi(t))}{\partial \xi_j} = +2c^2\xi_j(t), \qquad j = 1, \dots, n, \\
\frac{\mathrm{d}\xi_0(t)}{\mathrm{d}t} &= -\frac{\partial p_2(x(t),\xi(t))}{\partial x_0} = 0, \\
\frac{\mathrm{d}\xi_j(t)}{\mathrm{d}t} &= -\frac{\partial p_2(x(t),\xi(t))}{\partial x_j} = 0, \qquad j = 1, \dots, n,
\end{aligned}$$

which is solved by

$$\begin{pmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} 0 \\ x_1(0) \\ \vdots \\ x_n(0) \end{pmatrix} + \begin{pmatrix} -2\xi_0 \\ 2c^2\xi_1 \\ \vdots \\ 2c^2\xi_n \end{pmatrix} t, \qquad \begin{pmatrix} \xi_0(t) \\ \xi_1(t) \\ \vdots \\ \xi_n(t) \end{pmatrix} = \text{const.}$$

Now we are interested in zero-bi-characteristics, which enforces $|\xi_0| = c\sqrt{\xi_1^2 + \dots + \xi_n^2}$. In an attempt of beautification, we introduce a vector Ξ in $\mathbb{R}^n$ of length one,

$$\Xi_j := \frac{\xi_j}{\sqrt{\xi_1^2 + \dots + \xi_n^2}}, \qquad 1 \le j \le n, \qquad c^2\xi_j = c|\xi_0|\Xi_j.$$

Then we get

$$\begin{pmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} 0 \\ x_1(0) \\ \vdots \\ x_n(0) \end{pmatrix} + 2 \begin{pmatrix} -\operatorname{sign}(\xi_0) \\ c\Xi_1 \\ \vdots \\ c\Xi_n \end{pmatrix} |\xi_0| t,$$

and here we see nicely the propagation with speed c: when the time variable $2|\xi_0|t$ changes by one, the space position $2c\Xi|\xi_0|t$ changes by a vector of length c.

Now let us be given a solution $u = u(x_0, x_1, \dots, x_n)$, with some initial values $u_0(x_1, \dots, x_n) = u(0, x_1, \dots, x_n)$, and some initial velocity $\frac{\partial}{\partial x_0} u(0, x_1, \dots, x_n) = u_1(x_1, \dots, x_n)$. Suppose that the initial values u_0 have micro-locally a certain singular behaviour at a position $x_* \in \mathbb{R}^n$ in a certain direction $\xi_* = (\xi_1, \dots, \xi_n)$, in the sense of $u \notin H^{s+1}_{(0,x_*,\xi_0,\xi_*)}$, and ξ_0 being calculated according to the formula $\xi_0 := c|\xi_*|$.

Then this bad behaviour of u propagates through the space-time along the zero-bi-characteristic curve, forward in time and backward in time. We have a very detailed description of the singularities at a micro-local level — not only do we know where the singularities are (described by x), but we also know in which directions the singularities point (described by $\xi \neq 0$).

In linear hyperbolic problems, singularities can not be born or disappear, but they propagate forever along the zero-bi-characteristics of the principal symbol.

Micro-local analysis enables us to describe how regular solutions to some linear PDE are.

5.4 Further Ideas

5.4.1 How About L^p Based Sobolev Spaces ? Other Spaces ?

For $1 < p < \infty$, we define L^p based Sobolev spaces $H^s_p(\mathbb{R}^n)$ as

$$H^s_p(\mathbb{R}^n) := \langle D_x\rangle^{-s} L^p(\mathbb{R}^n).$$

Then ΨDOs with global symbol estimates in the symbol classes $S^m_{1,0}$ map continuously from $H^s_p(\mathbb{R}^n)$ into $H^{s-m}_p(\mathbb{R}^n)$, see [23]. Operators with symbols in $S^m_{\varrho,\delta}$ are harder to handle. A typical result reads: if $1 < p < \infty$ and $a = a(x,\xi) \in S^m_{\varrho,\delta}(\mathbb{R}^n \times \mathbb{R}^n)$ with $0 \le \delta < \varrho \le 1$, global symbol estimates, and

$$m \le -n(1-\varrho)\left|\frac{1}{2} - \frac{1}{p}\right|,$$

then $\mathcal{A}$ maps $L^p(\mathbb{R}^n)$ continuously into itself. Note that we have a *loss of regularity* in case of $\varrho < 1$. The cases $p = 1$ and $p = \infty$ are, in some sense, hopeless. I mean challenging, of course.

How about Hölder[7] spaces $C^{k+\alpha}(\mathbb{R}^n)$? Suppose these are global spaces, in the sense of

$$\forall \beta,\ |\beta| \le k\colon\ \exists C_\beta\colon\ \forall x \in \mathbb{R}^n\colon\ |D^\beta_x u(x)| \le C_\beta;$$
$$\exists C_0\colon\ \forall \beta,\ |\beta| = k\colon\ \forall x, y \in \mathbb{R}^n,\ |x-y| \le 1\colon\ |D^\beta u(x) - D^\beta u(y)| \le C_0|x-y|^\alpha.$$

If $\mathcal{A}$ is a ΨDO with symbol from $S^m_{1,0}$ and global estimates, then $\mathcal{A}$ will map $C^{k+\alpha}(\mathbb{R}^n)$ continuously into $C^{k+\alpha-m}$, **provided that** both $k+\alpha$ and $k+\alpha-m$ are positive and non-integral.

In particular, a first order ΨDO will **not** map $C^3(\mathbb{R}^n)$ into $C^2(\mathbb{R}^n)$, as an example. See [23] and [24]. Remember that the spaces C^k with integer k are ugly anyway (for instance, $C^1([-1,1])$ is not dense in $C^\alpha([-1,1])$ for $0 < \alpha < 1$, see [4]).

A remedy are the Zygmund[8] spaces, defined as follows. Choose some function $\psi_0 = \psi_0(\xi) \in C^\infty_0(\mathbb{R}^n)$ with $\psi_0(\xi) = 1$ for $|\xi| \le 1$, and $\psi_0(\xi) = 0$ for $|\xi| \ge 2$. Define $\Psi_j(\xi) := \psi_0(2^{-j}\xi)$, for $j \in \mathbb{N}_0$. Next set $\psi_j(\xi) := \Psi_j(\xi) - \Psi_{j-1}(\xi)$, for $j \in \mathbb{N}_+$. Then $\sum_{j=0}^\infty \psi_j(\xi) = 1$ on $\mathbb{R}^n_\xi$, and ψ_j is supported in an annulus of approximate radius 2^j.

For positive s, we say that a function u on $\mathbb{R}^n$ belongs to the Zygmund space C^s_* if

$$\sup_{j\in\mathbb{N}_0} 2^{js}\,\|\psi_j(D)u\|_{L^\infty(\mathbb{R}^n)} < \infty.$$

It turns out that, for non-integer s, the Zygmund space C^s_* coincides with the Hölder space C^s. But for integer s, C^s is a subspace of C^s_*. The space C^1_* can be also characterised by the requirement that

$$|u(x+y) + u(x-y) - 2u(x)| \le C_0|y|,$$

with some constant C_0 that does not depend on x or y. There is no space C^0_*.

The good news is that then each ΨDO with symbol from $S^m_{1,0}$ maps C^s_* into C^{s-m}_*, provided that s and $s-m$ are both positive.

5.4.2 What are the Classes $S^m_{\varrho,\delta}$ Good For ?

Suppose we are interested in solving a quasi-linear problem

$$\partial_t U(t,x) - \mathcal{A}(t,x,\{U\},D_x)U(t,x) = F(t,x,U),$$

using methods from microlocal analysis. The solution is expected to live in a certain Sobolev space. Then $\mathcal{A}$ is a PDO whose coefficients depend on the solution U, and therefore $\mathcal{A}$ has only limited regularity with respect to the x-variable. Consequently we need classes of pseudodifferential symbols $p = p(x,\xi)$, for whom the symbol estimates $|\partial^\alpha_x \partial^\beta_\xi p(x,\xi)| \le \dots$ only involve a finite number of α. Now we follow the

[7] Otto Hölder, 1859–1937
[8] Antoni Zygmund, 1900–1992

presentation of [24]. We consider so-called *microlocalizable scales* $\{X^s\}_{s\in\Sigma}$, which are scales of function spaces, such as the Zygmund spaces $\{C^s_*(\mathbb{R}^n)\}_{s\in(0,\infty)}$, or the Sobolev spaces $\{H^s_p(\mathbb{R}^n)\}_{s\in(n/p,\infty)}$. The multiplication of two functions stays inside the scale (which motivates the restriction $s > n/p$ in the Sobolev case).

Then we say that a pseudodifferential symbol $p = p(x,\xi)$ belongs to $X^s S^m_{1,0}$ if and only if

$$\left\|\partial_\xi^\beta p(\cdot,\xi)\right\|_{X^s} \le C_\beta \langle\xi\rangle^{m-|\beta|}, \quad \forall\beta\in\mathbb{N}_0^n.$$

We can also define a subclass $X^s S^m_{\mathrm{cl}}$ of $X^s S^m_{1,0}$ that contains classical symbols.

Then it can be shown that a symbol $p \in X^s S^m_{\mathrm{cl}}$ generates a map $\mathcal{P}(x,D)\colon X^{s+m} \to X^s$ provided that $s \in \Sigma$ and $s+m \in \Sigma$, by means of a tensor product representation.

Now let us be given a symbol $p \in X^s S^m_{1,0}$, and we wish to split it into two parts: $p = p^\sharp + p^b$, with $p^\sharp$ being smooth, and p^b being a symbol of lower order. But there is a price to pay: $p^\sharp$ will only belong to a symbol class $S^m_{1,\delta}$, not $S^m_{1,0}$.

The procedure is as follows. Take a cut-off function $\psi_0 = \psi_0(\xi)$ that is identically equal to 1 for $|\xi| \le 1$, and identically vanishing for $|\xi| \ge 2$. Then consider scaled versions of ψ_0, via $\Psi_j(\xi) := \psi_0(2^{-j}\xi)$. Finally put $\psi_j(\xi) := \Psi_j(\xi) - \Psi_{j-1}(\xi)$. By construction, we have $1 = \sum_{j=0}^\infty \psi_j(\xi)$, for all $\xi \in \mathbb{R}^n$. This is called LITTLEWOOD[9]–PALEY[10] *decomposition of unity.*

Next we build a smoothing operator J_ε, defined as a ΨDO like this:

$$J_\varepsilon f(x) := \psi_0(\varepsilon D) f(x).$$

Pick some $\delta \in (0,1)$, set $\varepsilon_j := 2^{-j\delta}$, and define

$$p^\sharp(x,\xi) := \sum_{j=0}^\infty \Big(J_{\varepsilon_j} p(x,\xi)\Big)\psi_j(\xi),$$

which is an infinitely differentiable function of x, because of the smoothing operator J_{ε_j} acting upon p.

Then it can be shown that $p^\sharp \in S^m_{1,\delta}$, and

$$p^b \in X^{s-t} S^{m-t\delta}_{1,0}, \qquad \text{provided that } s, s-t \in \Sigma, \quad t \ge 0.$$

The nice part is that the leading term $p^\sharp$ now is a smooth symbol, and δ no longer being zero is typically no big deal. The remainder term p^b has even less regularity with respect to x, because of $s - t < s$, but it is a lower order term, because of $m - t\delta < m$.

Further details can be found in [24].

5.4.3 What Remains True for FIOs ?

We did not present formulas about what happens if we compose a ΨDO and a FIO (in whatever order), because this would lead us too far into an integral jungle. But such formulas do exist and can be found in [17], together with their long proofs. Furthermore, we have stayed far away from making any attempt of composing two FIOs with two different phase functions, for two reasons: first the computational complexity, and second — it does not seem to make a lot of sense[11]. Because a FIO is typically related to a hyperbolic equation, and the phase function describes "along which path does the information travel". Why would one then compose two FIOs with different phase functions, and what is this supposed to mean ?

On the other hand, the adjoint operator of a FIO is meaningful, and formulas can be found in [17].

FIOs map between L^2 based Sobolev spaces as expected: if $\mathcal{A}$ is an FIO with an operator phase function, and if its symbol $a \in S^m_{1,0}$ has global symbol estimates, then $\mathcal{A}$ maps from $H^{s+m}(\mathbb{R}^n)$ continuously into

[9] JOHN EDENSOR LITTLEWOOD, 1885–1977

[10] RAYMOND PALEY, 1907–1933

[11] This could be a general principle: if an author's calculations become much more complicated than the author can endure (let alone a reader), then the reason could be that contact to reality has been lost, and the formulas under consideration do not describe a meaningful situation.

$H^s(\mathbb{R}^n)$. However, the situation is much worse in L^p: in [21] you can find mentioning of the result that an operator $\mathcal{A}$ with symbol $a \in S^m_{1,0}$ will map $L^p(\mathbb{R}^n)$ into itself provided that

$$m \leq -(n-1)\left|\frac{1}{2} - \frac{1}{p}\right|.$$

These restrictions on the parameters turn out to be sharp. Similar results for the wave equation have been obtained in the early 1980s by Peral, Beals, Miyachi, and others.

In particular, zero order FIOs will *not* map $L^p(\mathbb{R}^n)$ into itself, except for $p = 2$ and $n = 1$. The special situation $n = 1$ is no surprise, because then we often have solutions in the shape of travelling waves, for which any Lebesgue norm does not change as time goes on. And also the case $p = 2$ is special, because we can prove (without any microlocal methods) that the wave equation can not be well-posed in $L^p(\mathbb{R}^n)$, except if $p = 2$. *Hyperbolic problems don't make sense in* L^p.

And zero order FIOs do map Hölder spaces C^α into $C^{\alpha-(n-1)/2}$, with a loss of $(n-1)/2$ derivatives, which corresponds to a limit $p \to \infty$ in the previous loss formula.

Chapter 6

Boundary Value Problems

6.1 General Principles

This part follows [25].

Let Ω be a bounded domain in $\mathbb{R}^n$ with boundary $\partial\Omega$ in C^∞. We consider the boundary value problem

$$\begin{cases} \mathcal{A}(x, D)u = f(x), & x \in \Omega, \\ \mathcal{B}_j(x, D)u = g_j(x), & x \in \partial\Omega, \quad j = 1, \dots, J, \end{cases} \tag{6.1}$$

with scalar operators $\mathcal{A}$, $\mathcal{B}_1$, ..., $\mathcal{B}_J$ (some J). Suppose $\mathcal{A}$ of order $2m$ is elliptic on $\mathbb{R}^n$, its coefficients are from $C_b^\infty(\mathbb{R}^n)$, and the coefficients of $\mathcal{B}_j$ are from C_b^∞ in a neighbourhood of $\partial\Omega$, and $\operatorname{ord}(\mathcal{B}_j) = m_j \in \mathbb{N}_0$.

Question: given $\mathcal{A}$, what are conditions on $\mathcal{B}_j$ to make (6.1) well-posed ?

Question: what means "well-posed" ?

Being well-posed should imply at least that the map

$$\begin{pmatrix} \mathcal{A} \\ \mathcal{B}_1 \\ \vdots \\ \mathcal{B}_J \end{pmatrix} : H^s(\Omega) \to \begin{pmatrix} H^{s-2m}(\Omega) \\ H^{s-m_1-1/2}(\partial\Omega) \\ \vdots \\ H^{s-m_J-1/2}(\partial\Omega) \end{pmatrix}$$

is an isomorphism, for each integer $s \geq \max(2m, m_1 + 1, \dots, m_J + 1)$.

And a map $\mathcal{T} \in \mathcal{L}(X, Y)$ being an isomorphism between two Hilbert spaces X and Y means (among other things such as boundedness):

- $\mathcal{T}$ is injective
- $\mathcal{T}$ is surjective.

Recall that every bounded map $\mathcal{T}$ from X to Y admits the orthogonal decompositions into closed subspaces

$$X = \ker \mathcal{T} \oplus \overline{(\operatorname{img} \mathcal{T}^*)}, \qquad Y = \ker \mathcal{T}^* \oplus \overline{(\operatorname{img} \mathcal{T})},$$

with the over-line being the topological closure. Now $\mathcal{T}$ is injective iff $\ker \mathcal{T} = \{0\}$, and $\mathcal{T}$ is surjective iff $\operatorname{img} \mathcal{T} = Y$. Now take $\mathcal{T} = (\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_J)^\top$ from above. We wish to show

$$\ker \mathcal{T} = \{0\}, \qquad \ker \mathcal{T}^* = \{0\}, \qquad \operatorname{img} \mathcal{T} \text{ is closed.}$$

To prove this, we quote another lemma from functional analysis: if X and Y are Banach spaces, and $\mathcal{T}$ is a (possibly unbounded) operator acting from a domain $D(\mathcal{T}) \subset X$ into Y, and $\mathcal{T}$ is closed, then the following are equivalent:

- there is some $\varepsilon > 0$ with $\|\mathcal{T}x\|_Y \geq \varepsilon \|x\|_X$ for all $x \in D(\mathcal{T})$;
- $\mathcal{T}$ is injective, and $\operatorname{img} \mathcal{T}$ is closed.

Therefore we wish to prove, with $\mathcal{T} = (\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_J)^\top$, and with some lower regularity bound s_0:

$$\forall s \geq s_0\colon\ \exists C_s\colon\ \forall u \in H^s(\Omega)\colon\ \|u\|_{H^s(\Omega)} \leq C_s \left(\|\mathcal{A}u\|_{H^{s-2m}(\Omega)} + \sum_{j=1}^{J} \|\mathcal{B}_j u\|_{H^{s-m_j-1/2}(\partial\Omega)} \right), \tag{6.2}$$

and if the formal adjoint of $\mathcal{T}$ is $\mathcal{T}^* = (\mathcal{A}^*, \mathcal{B}_1^*, \dots, \mathcal{B}_{J^*}^*)^\top$, with $\operatorname{ord} \mathcal{B}_j^* = m_j^*$, then

$$\forall s \geq s_0\colon\ \exists C_s\colon\ \forall u \in H^s(\Omega)\colon\ \|u\|_{H^s(\Omega)} \leq C_s \left(\|\mathcal{A}^* u\|_{H^{s-2m}(\Omega)} + \sum_{j=1}^{J^*} \left\|\mathcal{B}_j^* u\right\|_{H^{s-m_j^*-1/2}(\partial\Omega)} \right). \tag{6.3}$$

Now what is the formal adjoint operator $\mathcal{T}^*$?

To answer this, we have to dig deeper. Let $\nu(x)$ be the outward unit normal vector at $x \in \partial\Omega$, and let τ denote any of the $n-1$ tangential directions at x, for some suitably chosen coordinate frame at x. Then each $\mathcal{B}_j$ can be written as

$$\mathcal{B}_j(x, D_x) = \gamma_0 \sum_{k+|\beta| \leq m_j} b_{jk\beta}(x) D_\tau^\beta D_\nu^k,$$

with γ_0 being the trace operator at the boundary $\partial\Omega$. We assume that the highest order normal derivative is actually present for each $x \in \partial\Omega$, which means $b_{jm_j0}(x) \neq 0$ for all $x \in \partial\Omega$. The reason for this assumption is twofold: if b_{jm_j0} somewhere equals to zero, and elsewhere not, then we have a mess; but this course is meant as an introduction. And the second reason is that such a situation will turn out to be excluded anyway, by the Shapiro–Lopatinskij[1] condition[2] which we are going to develop in the sequel.

Next we can assume that all the orders m_j are distinct. Because if $m_1 = m_2$ (say), then the boundary conditions expressed by $\mathcal{B}_1$ and $\mathcal{B}_2$ either contradict, and the BVP (6.1) is unsolvable anyway (for non-cooperating g_1, g_2), or we can substitute one boundary condition into the other, and get new boundary conditions for which then $m_1 \neq m_2$ becomes true. By re-ordering we then get $m_1 < m_2 < \ldots < m_J$. We also assume $m_J \leq 2m-1$, without giving a reason. Such a system $(\mathcal{B}_1, \dots, \mathcal{B}_J)$ of boundary operators is called a *normal system*. We remark that it is always possible to find a function u that satisfies all the boundary conditions $\mathcal{B}_j u = g_j$.

Now we have a look at $\mathcal{A}(x, D_x)$, evaluated at a boundary point x_0. By assumption, $\mathcal{A}$ is elliptic, which means that its principal symbol $a_{2m}(x, \xi)$ does not vanish, for each $\xi \neq 0$. Now take $x = x_0$ and $\xi \parallel \nu$: $a_{2m}(x_0, \nu) \neq 0$. Therefore, near x_0, $\mathcal{A}$ has the form

$$\mathcal{A}(x, D_x) = \sum_{\ell=0}^{2m} \tilde{a}_\ell(x, D_\tau) D_\nu^\ell, \qquad \tilde{a}_{2m} \neq 0, \quad \operatorname{ord}(\tilde{a}_\ell) \leq 2m - l.$$

In particular, $\mathcal{A}$ contains an item with $2m$ normal derivatives, everywhere at the boundary. We extend the vector field ν from $\partial\Omega$ to a (two-sided) tubular neighbourhood of $\partial\Omega$. Now we choose a cut-off function χ which has support in a tubular neighbourhood of $\partial\Omega$, and is equal to 1 in a smaller neighbourhood of $\partial\Omega$. Then we write

$$\mathcal{A}(x, D_x) = (1 - \chi(x))\mathcal{A}(x, D_x) + \sum_{\ell=0}^{2m} \chi(x) \tilde{a}_\ell(x, D_\tau) D_\nu^\ell.$$

Consider the integral $\int_\Omega (\mathcal{A}u)\overline{v}\,\mathrm{d}x$, for smooth functions u, v, and try to shift as many derivatives from u towards v, by partial integration.

The first part of $\mathcal{A}$ is easy to handle:

$$\int_\Omega (1-\chi)(\mathcal{A}u) \cdot \overline{v}\,\mathrm{d}x = \int_\Omega u \cdot \overline{\mathcal{A}^*((1-\chi)v)}\,\mathrm{d}x,$$

[1] Yaroslav Borisovich Lopatinskij, 1906–1981

[2] This condition has been discovered 1953 by Z.Ya. Shapiro and Lopatinskij independently, but more abstract and more general results have also been obtained by Mark Iosivovich Vishik (1921–2012) in 1952.

without any boundary integrals. The second part of $\mathcal{A}$ is only relevant near the boundary, and first we handle the tangential derivatives D_τ:

$$\int_\Omega \sum_{\ell=0}^{2m} \chi(x)\Big(\tilde{a}_\ell(x, D_\tau)D_\nu^\ell u\Big) \cdot \overline{v} \,\mathrm{d}x = \int_\Omega \sum_{\ell=0}^{2m} \tilde{\chi}(x)\Big(D_\nu^\ell u\Big) \cdot \overline{\tilde{a}_\ell^t(x, D_\tau)v} \,\mathrm{d}x,$$

for some function $\tilde{\chi}$ supported near the boundary, and some differential operator $\tilde{a}_\ell$ of order $\leq 2m - \ell$ in tangential direction. It remains to bring away the derivatives D_ν^ℓ from u. To this end, we remark that

$$\begin{aligned}
\int_\Omega \chi(x)w(x) \cdot \partial_\nu^\ell u \,\mathrm{d}x &= \int_\Omega \chi(x)w(x)\Big(\vec{\nu} \cdot \nabla\Big)\partial_\nu^{\ell-1} u \,\mathrm{d}x \\
&= \int_\Omega \operatorname{div}\Big(\chi(x)w(x)\vec{\nu}(x)\partial_\nu^{\ell-1}u(x)\Big) - \operatorname{div}\Big(\chi(x)w(x)\vec{\nu}(x)\Big) \cdot \partial_\nu^{\ell-1}u(x) \,\mathrm{d}x \\
&= \int_{\partial\Omega} \chi(x)w(x)\partial_\nu^{\ell-1}u(x) \,\mathrm{d}\sigma \\
&\quad - \int_\Omega \operatorname{div}\Big(\chi(x)\vec{\nu}(x)\Big)w(x) \cdot \partial_\nu^{\ell-1}u(x) \,\mathrm{d}x - \int_\Omega \chi(x)\Big(\partial_\nu w(x)\Big)\partial_\nu^{\ell-1}u(x) \,\mathrm{d}x,
\end{aligned}$$

for a generic function w. Continuing in this style, we end up with

$$\int_\Omega (\mathcal{A}u) \cdot \overline{v} \,\mathrm{d}x = \int_\Omega u \cdot \overline{\mathcal{A}^* v} \,\mathrm{d}x + \sum_{\ell=0}^{2m-1} \int_{\partial\Omega} (D_\nu^\ell u) \cdot \overline{(\mathcal{C}_\ell v)} \,\mathrm{d}\sigma$$

with some differential operators on the boundary,

$$\mathcal{C}_\ell = \sum_{i=0}^{2m-1-\ell} \gamma_0 c_{\ell i}(x, D_\tau)D_\nu^i, \qquad \operatorname{ord}(c_{\ell i}) \leq 2m - 1 - \ell - i, \qquad c_{\ell, 2m-1-\ell} \neq 0.$$

Now we split the set $\{0, 1, \dots, 2m-1\}$ into two disjoint parts,

$$\{0, 1, \dots, 2m-1\} = \mathcal{M} \cup \not\mathcal{M}, \qquad \mathcal{M} := \{m_1, m_2, \dots, m_J\}, \qquad \not\mathcal{M} := \{0, 1, \dots, 2m-1\} \setminus \mathcal{M}.$$

Then we can write

$$\int_\Omega (\mathcal{A}u) \cdot \overline{v} \,\mathrm{d}x = \int_\Omega u \cdot \overline{\mathcal{A}^* v} \,\mathrm{d}x + \sum_{\ell \in \mathcal{M}} \int_{\partial\Omega} (D_\nu^\ell u) \cdot \overline{(\mathcal{C}_\ell v)} \,\mathrm{d}\sigma + \sum_{\ell \in \not\mathcal{M}} \int_{\partial\Omega} (D_\nu^\ell u) \cdot \overline{(\mathcal{C}_\ell v)} \,\mathrm{d}\sigma,$$

and in the first boundary integral, we may plug in the boundary operators $\mathcal{B}_j$, where we take the liberty of re-defining $\mathcal{C}_\ell$ if necessary:

$$\int_\Omega (\mathcal{A}u) \cdot \overline{v} \,\mathrm{d}x = \int_\Omega u \cdot \overline{\mathcal{A}^* v} \,\mathrm{d}x + \sum_{j=1}^{J} \int_{\partial\Omega} (\mathcal{B}_j u) \cdot \overline{(\mathcal{C}_{m_j} v)} \,\mathrm{d}\sigma + \sum_{\ell \in \not\mathcal{M}} \int_{\partial\Omega} (D_\nu^\ell u) \cdot \overline{(\mathcal{C}_\ell v)} \,\mathrm{d}\sigma,$$

from which we can read off what J^* and $\mathcal{B}_j^*$ are: it holds $J^* = 2m - J$, and the adjoint boundary operators $\mathcal{B}_1^*, \dots, \mathcal{B}_{J^*}^*$ are the operators $\mathcal{C}_\ell$ with $\ell \in \not\mathcal{M}$. We remark that the $\mathcal{B}_j^*$ are not uniquely determined (because our choice $D_\nu^\ell u$ of the other factor in the integrand was somewhat arbitrary).

Now we are (finally) in a position to explain what $\mathcal{T}$ and $\mathcal{T}^*$ actually are: since the boundary conditions $\{\mathcal{B}_1, \dots, \mathcal{B}_J\}$ can be satisfied anyway, we may shift u appropriately and assume that all the g_j are zero. Then $\mathcal{T} = \mathcal{A}$, defined on $D(\mathcal{T}) = \{u \in H^{2m}(\Omega) \colon \mathcal{B}_1 u = \dots = \mathcal{B}_J u = 0\}$. And similarly we can set $\mathcal{T} = \mathcal{A}^*$, defined on $D(\mathcal{T}^*) = \{v \in H^{2m}(\Omega) \colon \mathcal{B}_1^* v = \dots = \mathcal{B}_{J^*}^* v = 0\}$. If $u \in D(\mathcal{T})$ and $v \in D(\mathcal{T}^*)$, then we have indeed $\int_\Omega (\mathcal{A}u) \cdot \overline{v} \,\mathrm{d}x = \int_\Omega u \cdot \overline{\mathcal{A}^* v} \,\mathrm{d}x$.

This was just a sketch of how to construct $\mathcal{T}^*$, a thorough presentation can be found in [30].

However, we often have no chance of proving uniqueness of a solution to (6.1), because it is very hard to determine all the eigenvalues precisely; and for given $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_J)^\top$, we have no possibility of knowing whether "we are sitting on an eigenvalue of the problem" or not. Therefore, instead of showing (6.2) and (6.3), we settle for less:

Try to prove that, for each $s \geq s_0$, there is some C_s such that for each $u \in H^s(\Omega)$, we have

$$\|u\|_{H^s(\Omega)} \leq C_s \left(\|\mathcal{A}u\|_{H^{s-2m}(\Omega)} + \sum_{j=1}^{J} \|\mathcal{B}_j u\|_{H^{s-m_j-1/2}(\partial\Omega)} + \|u\|_{H^{s-1}(\Omega)} \right), \tag{6.4}$$

$$\|u\|_{H^s(\Omega)} \leq C_s \left(\|\mathcal{A}^* u\|_{H^{s-2m}(\Omega)} + \sum_{j=1}^{J^*} \left\|\mathcal{B}_j^* u\right\|_{H^{s-m_j^*-1/2}(\partial\Omega)} + \|u\|_{H^{s-1}(\Omega)} \right). \tag{6.5}$$

Then $\ker \mathcal{T}$ and $\ker \mathcal{T}^*$ will be contained in $\cap_{s\geq s_0} H^s(\Omega)$, with higher order derivatives bounded by lower order derivatives. Every bounded and closed subset of $\ker \mathcal{T}$ is then compact, hence $\ker \mathcal{T}$ (and $\ker \mathcal{T}^*$ as well) are finite-dimensional.

We observe that (6.4) and (6.5) are stable under perturbations of $\mathcal{A}$ and $\mathcal{B}_j$, in the following sense: if we change the coefficients of lower order terms of $\mathcal{A}$ and $\mathcal{B}_j$ by whatever smooth function, then we can absorb this perturbation into $\|u\|_{H^{s-1}}$ on the RHS, perhaps making C_s bigger. And if we perturb the highest order coefficients of $\mathcal{A}$ and $\mathcal{B}_j$ by a small amount ("small" measured in appropriate C_b^t norms for some (perhaps large) t), then we can absorb this change into the LHS if the perturbation is only small.

The following two sections will establish:

- if (6.4) holds, then $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_J)$ satisfies the Shapiro-Lopatinskij condition,
- if $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_J)$ satisfies the Shapiro-Lopatinskij condition, then (6.4) holds.

6.2 Necessity

We assume that (6.4) holds for all functions $u \in H^s(\Omega)$, and we look what can be concluded from that.

To keep things easy, we fix the Sobolev order s as $s = 2m$. Pick some point $x_0 \in \partial\Omega$ and some small positive ε. Then (6.4) holds in particular for all $u \in H^s(\Omega)$ with support in $\overline{\Omega} \cap B_\varepsilon(x_0) =: U_\varepsilon$. We introduce a new set of Cartesian coordinates $(y_1, \dots, y_n)$, with origin at x_0, in such a way that the y_n axis points from x_0 along the interior normal direction, and the other axes point tangentially at x_0 along $\partial\Omega$. We rewrite the system and the boundary conditions in the new coordinates, and (6.4) continues to hold, perhaps with bigger C_s, because the Sobolev norms can be an-isotropic (depending on how you have defined them).

Now our intention is to introduce yet another coordinate system, in such a way that $\partial\Omega$ becomes "flat" near x_0. Expressed in the y-system, the set U_ε is

$$U_\varepsilon = \{y \in \mathbb{R}^n : |y| < \varepsilon, y_n \geq h(y_1, y_2, \dots, y_{n-1})\},$$

for some smooth function h with $h(0) = 0$ and $\nabla h(0) = 0$. We also introduce the notation $y' := (y_1, \dots, y_{n-1})$. Then $y = (y', y_n)$.

The new coordinates are $(z_1, \dots, z_n)$ with

$$z_1 := y_1, \dots, z_{n-1} := y_{n-1}, \quad z_n := y_n - h(y_1, \dots, y_{n-1}),$$

and the set U_ε is mapped onto a set V_ε. The boundary $\partial\Omega$ near x_0 is then described by $z_n = 0$. Crucial in this flattening procedure is now $|\nabla h| = \mathcal{O}(\varepsilon)$, for $|y'| \leq \varepsilon$. When we rewrite the PDE $\mathcal{A}u = f$ and the boundary conditions in the z-coordinates, we have to transfer $y \mapsto \Phi(y) = z$, with Φ being a smooth diffeomorphism, and then the principal symbol $p_{\mathrm{pr}}(z, \zeta)$ of a ΨDO $\mathcal{P}$ (in the z world) transfers along the rule

$$p_{\mathrm{pr}}(z, \zeta) \rightsquigarrow p_{\mathrm{pr}}(\Phi(y), (\Phi'(y))^{-\top}\eta)$$

into the y-world, with $(\Phi'(y))^{-\top}$ being the inverse of the transpose of the Jacobi matrix. In our case, $\Phi'(y) = \mathrm{id} + \mathcal{O}(\varepsilon)$ is almost equal to the identity matrix id, which means that the highest order coefficients of $\mathcal{A}$ and of $\mathcal{B}_j$ change only by $\mathcal{O}(\varepsilon)$ (measured in the L^∞ norm), and lower order coefficients can change in a complicated way. We write $\mathcal{A}(z, D_z)$ and $\mathcal{B}_j(z, D_z)$ for the new operators. Similarly, we write $u(x)$ when we refer to the function u in the old x-coordinates, and $u(z)$ when we refer to u in the new z-coordinates.

Moreover, let us freeze the coefficients of $\mathcal{A}$ and $\mathcal{B}_j$: instead of z, we use 0. This brings a change of size $\mathcal{O}(\varepsilon)$ in the coefficients (assuming they are Lipschitz). The result then is: for all $u = u(x) \in H^{2m}(\Omega)$ with $\operatorname{supp} u(x) \subset U_\varepsilon$, we have

$$\|u(z)\|_{H^{2m}(V_\varepsilon)} \le 2C_s \left(\|\mathcal{A}(0, D_z)u\|_{L^2(V_\varepsilon)} + \sum_{j=1}^{J} \|\mathcal{B}_j(0, D_z)u\|_{H^{2m-m_j-1/2}(\partial V_\varepsilon \cap \{z\colon z_n=0\})} + \|u\|_{H^{2m-1}(V_\varepsilon)} \right).$$

Now because of $\operatorname{supp} u(x) \subset U_\varepsilon$, the function $u(z)$ equals zero in a neighbourhood of the "roof boundary" of V_ε. Then we may tacitly extend the function $u(z)$ to the set $\mathbb{R}^n_+ := \{z \in \mathbb{R}^n \colon z_n \ge 0\}$ by zero values, and the result then is

$$\|u(z)\|_{H^{2m}(\mathbb{R}^n_+)} \le 2C_s \left(\|\mathcal{A}(0, D_z)u\|_{L^2(\mathbb{R}^n_+)} + \sum_{j=1}^{J} \|\mathcal{B}_j(0, D_z)u\|_{H^{2m-m_j-1/2}(\mathbb{R}^{n-1})} + \|u\|_{H^{2m-1}(\mathbb{R}^n_+)} \right),$$

valid for all functions $u(x) \in H^{2m}(\Omega)$ with $\operatorname{supp} u(x) \subset U_\varepsilon$.

As a next simplification step, we shift all lower order terms of $\mathcal{A}(0, D_z)$ acting upon $u(z)$ into the remainder $\|u\|_{H^{2m-1}(\mathbb{R}^n_+)}$, at the expense of increasing C_s. We do the same with all lower order terms of $\mathcal{B}_j(0, D_z)$:

$$\|u(z)\|_{H^{2m}(\mathbb{R}^n_+)} \le C_s' \left(\|\mathcal{A}_{\mathrm{pr}}(0, D_z)u\|_{L^2(\mathbb{R}^n_+)} + \sum_{j=1}^{J} \|\mathcal{B}_{j,\mathrm{pr}}(0, D_z)u\|_{H^{2m-m_j-1/2}(\mathbb{R}^{n-1})} + \|u\|_{H^{2m-1}(\mathbb{R}^n_+)} \right),$$

valid for all functions $u(x) \in H^{2m}(\Omega)$ with $\operatorname{supp} u(x) \subset U_\varepsilon$.

Define $f(z) := \mathcal{A}_{\mathrm{pr}}(0, D_z)u(z)$ and $g_j(z') = \mathcal{B}_{j,\mathrm{pr}}(0, D_z)u(z', 0)$. Then we know that

$$\|u(z)\|_{H^{2m}(\mathbb{R}^n_+)} \le C_s' \left(\|f(z)\|_{L^2(\mathbb{R}^n_+)} + \sum_{j=1}^{J} \|g_j(z')\|_{H^{2m-m_j-1/2}(\mathbb{R}^{n-1})} + \|u(z)\|_{H^{2m-1}(\mathbb{R}^n_+)} \right). \tag{6.6}$$

On the other hand, we can calculate u from f and all g_j, as follows: We know

$$\begin{cases} \displaystyle\sum_{|\alpha|=2m} a_\alpha(0) D_z^\alpha u(z) = f(z), & z \in \mathbb{R}^n_+, \\ \displaystyle\sum_{k+|\beta'|=m_j} b_{jk\beta'}(0) D_{z'}^{\beta'} D_n^k u(z', 0) = g_j(z'), & z' \in \mathbb{R}^{n-1}, \quad j = 1, \dots, J, \end{cases}$$

and now it becomes possible to perform a partial Fourier transform in tangential direction, replacing z' by its co-variable ζ'. Observe that $\mathcal{A}_{\mathrm{pr}}(0, D_z)u(z) = f(z)$ can also be expressed as

$$\sum_{k=0}^{2m} \sum_{|\alpha'|=2m-k} a_{\alpha',k}(0) D_{z'}^{\alpha'} D_n^k u(z', z_n) = f(z', z_n).$$

We introduce the pseudodifferential symbols

$$a_k(\zeta') := \sum_{|\alpha'|=2m-k} a_{\alpha',k}(0)(\zeta')^{\alpha'}, \qquad b_{jk}(\zeta') := \sum_{|\beta'|=m_j-k} b_{jk\beta'}(0)(\zeta')^{\beta'},$$

and the partial Fourier transforms

$$\hat{u}(\zeta', z_n) := \mathcal{F}_{z'\to\zeta'}\{u(z', z_n)\}, \qquad \hat{f}(\zeta', z_n) := \mathcal{F}_{z'\to\zeta'}\{f(z', z_n)\}, \qquad \hat{g}_j(\zeta') := \mathcal{F}_{z'\to\zeta'}\{g(z')\},$$

and then we know that

$$\begin{cases} \displaystyle\sum_{k=0}^{2m} a_k(\zeta') D_n^k \hat{u}(\zeta', z_n) = \hat{f}(\zeta', z_n), & \zeta' \in \mathbb{R}^{n-1}, \quad 0 \le z_n < \infty, \\ \displaystyle\sum_{k=0}^{m_j} b_{jk}(\zeta') D_n^k \hat{u}(\zeta', 0) = \hat{g}_j(\zeta'), & \zeta' \in \mathbb{R}^{n-1}, \quad j = 1, \dots, J. \end{cases} \tag{6.7}$$

And moreover, from $u(z) \in H^{2m}(\mathbb{R}^n_+)$ we in particular get $u \in L^2(\mathbb{R}^n_+)$, which implies (by Plancherel)

$$\int_{z_n=0}^{\infty} \int_{\mathbb{R}^{n-1}_{\zeta'}} |\hat{u}(\zeta', z_n)|^2 \, \mathrm{d}\zeta' \, \mathrm{d}z_n < \infty.$$

Similarly, $f(z) \in L^2(\mathbb{R}^n_+)$ gives us

$$\int_{z_n=0}^{\infty} \int_{\mathbb{R}^{n-1}_{\zeta'}} |\hat{f}(\zeta', z_n)|^2 \, \mathrm{d}\zeta' \, \mathrm{d}z_n < \infty.$$

Now (6.7) is an ODE in the variable z_n, with constant coefficients, and with J boundary conditions at $z_n = 0$. Therefore, the solution $\hat{u}(\zeta', \cdot)$ can be constructed by methods which we know from the ODE courses (it is just tedious).

Now we present the famous Shapiro-Lopatinskij condition first, and afterwards we explain why it follows from the validity of (6.4).

Condition 1 (Shapiro–Lopatinskij). *For all $\zeta' \in \mathbb{R}^{n-1} \setminus 0$, the initial-boundary value problem*

$$\begin{cases} \sum_{k=0}^{2m} a_k(\zeta') D_n^k v(z_n) = 0, & 0 \le z_n < \infty, \\ \lim_{z_n \to +\infty} v(z_n) = 0, & \\ \sum_{k=0}^{m_j} b_{jk}(\zeta') D_n^k v(0) = 0, & j = 1, \dots, J \end{cases} \tag{6.8}$$

possesses only the solution $v \equiv 0$.

Our strategy will be: assume that Condition 1 is wrong. Then we will carefully construct a function u for which (6.6) is wrong. But (6.6) has been found as a conclusion from (6.4).

Let us be given some $\zeta' \in \mathbb{R}^{n-1} \setminus 0$. We calculate all functions $v = v(z_n)$ that satisfy the first two equations of (6.8). We make the ansatz

$$v(z_n) = e^{\mathrm{i}\tau z_n}$$

with unknown $\tau \in \mathbb{C}$, and substituting into the first equation then gives

$$\left(\sum_{k=0}^{2m} a_k(\zeta') \tau^k \right) e^{\mathrm{i}\tau z_n} \overset{!}{=} 0,$$

which implies $a_{\mathrm{pr}}(0, \zeta', \tau) \overset{!}{=} 0$, and $a_{\mathrm{pr}}(0, \zeta)$ is the pseudodifferential symbol of $\mathcal{A}_{\mathrm{pr}}(0, D_z)$. The equation $a_{\mathrm{pr}}(0, \zeta', \tau) = 0$ is a polynomial in the variable τ of degree $2m$. Because $\mathcal{A}$ is elliptic, the solution τ cannot be a real number. We assume for simplicity that the polynomial $\tau \mapsto a_{\mathrm{pr}}(0, \zeta', \tau)$ has no multiple roots. Then there are m_+ roots $\tau_1, \tau_2, \dots, \tau_{m_+}$ with strictly positive real part, and there are $2m - m_+$ roots with strictly negative real part (which do not matter because of the second equation in (6.8)).[3] Hence we have shown that all functions $v = v(z_n)$ which decay for $z_n \to +\infty$ and solve $\sum_{k=0}^{2m} a_k(\zeta') D_n^k v(z_n) = 0$ are given as

$$v(z_n) = \sum_{\ell=1}^{m_+} c_\ell \exp(\mathrm{i}\tau_\ell z_n), \qquad c_\ell \in \mathbb{C}. \tag{6.9}$$

If $\zeta' \in \mathbb{R}^{n-1}$ is being replaced by $t\zeta'$ (with $t \in \mathbb{R}_+$), then $\tau_\ell \in \mathbb{C} \setminus \mathbb{R}$ is being replaced by $t\tau_\ell$. Now we understand the second equation in (6.8): if some item $\exp(\mathrm{i}\tau z_n)$ with $\Im\tau < 0$ were present in v, then this item would explode for fixed positive z_n and $|\zeta'| \to \infty$, contradicting the requirement that the function $\zeta' \mapsto \hat{u}(\zeta', z_n)$ belong to $L^2(\mathbb{R}^{n-1}_{\zeta'})$.

[3] We remark that it can be shown $m_+ = m$ if $n \geq 3$.

Now we assume Condition 1 to be wrong. Then there is some $\mathfrak{z}' \in \mathbb{R}^{n-1} \setminus 0$ and $c_1, \ldots, c_{m_+} \in \mathbb{C}$ (at least one of them non-zero) such that the function

$$v_t(z_n) := \sum_{\ell=1}^{m_+} c_\ell \exp(\mathrm{i} t \tau_\ell z_n), \qquad a_{\mathrm{pr}}(0, t\mathfrak{z}', t\tau_\ell) = 0,$$

satisfies all equations of (6.8), for all $t \in \mathbb{R}_+$. Let us assume that the roots $\tau_1, \ldots, \tau_{m_+}$ are numbered in such a way that $c_1 \neq 0$ and τ_1 has the smallest (positive) imaginary part of the $\tau_1, \ldots, \tau_{m_+}$. We choose a smooth function $\varphi = \varphi(z)$ with small support near $z = 0$, and then we set

$$u(z) := \varphi(z) \exp(\mathrm{i} t \mathfrak{z}' \cdot z') \cdot v_t(z_n).$$

We claim that this function violates (6.6) for large t. The only purpose of the function φ is to make sure that the function $u(x)$ (expressed in the x-language) has support in U_ε.

Let $\alpha \in \mathbb{N}_0^n$ be a multi-index. What is the largest contribution to $|D_z^\alpha u(z)|$, for large t ? If one derivative lands on $\varphi(z)$, then we waste one factor t. Hence we find that the largest contribution is

$$|\varphi(z)| \cdot t^{|\alpha|} \cdot \left| (\mathfrak{z}')^{\alpha'} \right| \cdot |c_1| \cdot |\tau_1|^{\alpha_n} \cdot |\exp(-\Im \tau_1 t z_n)| \,.$$

Therefore, the LHS of (6.6) has growth order $t^{2m} |\exp(-\Im \tau_1 t z_n)|$, which still needs to be integrated over z_n, resulting in $t^{2m-1/2}$.

The first term on the RHS has growth order $t^{2m-1} |\exp(-\Im \tau_1 t z_n)|$, because now at least one derivative lands on $\varphi(z)$, by the very construction of the function v_t. The same applies to the second term on the RHS of (6.6), which has a growth order $t^{2m-3/2}$. And the third item $\|u(z)\|_{H^{2m-1}(\mathbb{R}^n_+)}$ on the RHS is obviously weaker than the LHS.

This reveals to us that for large t the RHS of (6.6) becomes smaller than the LHS, which is absurd.

Therefore, Condition 1 is necessary for (6.4) to hold. Now the third line of (6.8) must be strong enough to enforce that all the coefficients $c_1, \ldots, c_{m_+}$ from (6.9) are zero. This makes $J \geq m_+$ necessary.

Now we can repeat this reasoning with the other inequality (6.5) that refers to the adjoint problem. The numbers τ_ℓ do not change. And in that case, $2m - J \geq m_+$ becomes necessary.

Together with $m_+ = m$ (which can be proved for $n \geq 3$ and is an additional requirement (but a very realistic one) for $n = 2$) we then find

$$J = m,$$

as necessary condition for the well-posedness of (6.1).

6.3 Sufficiency

Now we assume the following: the operator $\mathcal{A}$ is elliptic, the Condition 1 holds at every point x_0 of the boundary $\partial\Omega$, and $m_+ = J = m$. We claim that then (6.4) holds.

For clarity, let us retrace the steps we have gone in the necessity part:

- pick a point $x_0 \in \partial\Omega$ and a small one-sided neighbourhood $U_\varepsilon \subset \overline{\Omega}$ of x_0,
- translate and rotate the coordinate system, obtain $y_1, \ldots, y_n$ as new coordinates near x_0,
- flatten the boundary, obtain a new neighbourhood, and new coordinates $z_1, \ldots, z_n$, with x_0 now corresponding to $z = 0$,
- throw away lower order parts of $\mathcal{A}$ and $\mathcal{B}_j$, obtain $\mathcal{A}_{\mathrm{pr}}$ and $\mathcal{B}_{j,\mathrm{pr}}$,
- freeze the coefficients of $\mathcal{A}_{\mathrm{pr}}$ and $\mathcal{B}_{j,\mathrm{pr}}$,
- perform partial Fourier transform in tangential direction, obtain an ODE in z_n direction with constant coefficients that depend in a polynomial way on the Fourier variable ζ', which is the frequency variable associated to z'.

The solution to an initial-value problem to an ODE of order $2m$ with constant coefficients is built from $2m$ linearly independent solutions of the form $z_n \mapsto \exp(\mathrm{i}\tau z_n)$, with the usual modifications in case of higher multiplicities. By the assumption $m_+ = m$, m of these $2m$ linearly independent solutions have exponential growth for $z_n \to +\infty$, which makes them inadmissible (because these functions have also exponential growth for $|\zeta'| \to \infty$, thus they are not members of L^2). Hence only m linearly independent functions of the type $z_n \mapsto \exp(\mathrm{i}\tau_\ell z_n)$ remain. Furthermore, the condition $J = m$ means that we have m initial conditions prescribed at $z_n = 0$. After we have done all the necessary substitutions, we end up with a system of m linear equations for m unknown complex numbers $c_1, \dots, c_m$, which is a problem known from linear algebra. The Shapiro–Lopatinskij condition means that the resulting matrix of size $m \times m$ has only a trivial null-space. But this is equivalent to the invertibility of this $m \times m$ matrix.

The final conclusion then is that the above mentioned inhomogeneous ODE of order $2m$ is uniquely solvable; and it can be shown that its solution enjoys an estimate that is compatible to (6.4).

What do we have now: we can split the function u into many little pieces, and each piece satisfies an estimate which looks like (6.4). What remains to be done is to glue the pieces together.

Let $\overline{\Omega}$ be covered by open sets U_k, with $k = 1, \dots, K$. Note that (in deviation to the notation in the previous section) some sets U_k will contain portions of the complement set $\mathbb{R}^n \setminus \Omega$. To each U_k, let us be given a cut-off function φ_k with support inside U_k, values between 0 and 1, and suppose that $\sum_{k=1}^K \varphi_k(x) = 1$, for all $x \in \overline{\Omega}$. For later use, we prepare functions ψ_k with bigger support than φ_k, with $\psi_k(x) = 1$ for $x \in \operatorname{supp}\varphi_k$. The result then is $\varphi_k = \varphi_k\psi_k$. We can arrange the supports in such a way that nowhere do more than N_0 supports of the functions ψ_k overlap, with some natural number N_0 that only depends on the space dimension n. This would turn out to be relevant should we later find it necessary to make the sets U_k smaller — then the number N_0 will not grow.

Then we can split the given function $u \in H^{2m}(\Omega)$ according to $u = \sum_{k=1}^K (\varphi_k u)$. Now fix some k.

The inner case: the set U_k does not intersect $\partial\Omega$

The elliptic operator $\mathcal{A}$ has a parametrix $\mathcal{A}^\sharp$, which is a ΨDO of order $-2m$, such that $\mathcal{A}^\sharp \circ \mathcal{A} = \mathrm{id} - \mathcal{R}$ with a smoothing operator $\mathcal{R}$. It follows (with $[\mathcal{A}, \varphi_k]$ being the commutator $\mathcal{A} \circ \varphi_k - \varphi_k \mathcal{A}$)

$$
\begin{aligned}
\varphi_k u &= \mathcal{A}^\sharp \circ \mathcal{A}\varphi_k u + \mathcal{R}\varphi_k u, \\
\|\varphi_k u\|_{H^{2m}(\Omega)} &\le C \|\mathcal{A}(\varphi_k u)\|_{L^2(\mathbb{R}^n)} + C \|\varphi_k u\|_{L^2(\mathbb{R}^n)} \\
&\le C \|\varphi_k \mathcal{A} u\|_{L^2(\mathbb{R}^n)} + C \left\| \Big[\mathcal{A}, \varphi_k\Big] u \right\|_{L^2(\mathbb{R}^n)} + C \|\varphi_k u\|_{L^2(\mathbb{R}^n)} \\
&= C \|\varphi_k \mathcal{A} u\|_{L^2(\Omega)} + C \left\| \Big[\mathcal{A}, \varphi_k\Big] \psi_k u \right\|_{L^2(\mathbb{R}^n)} + C \|\varphi_k u\|_{L^2(\Omega)} \\
&\le C \|\varphi_k \mathcal{A} u\|_{L^2(\Omega)} + C \|\psi_k u\|_{H^{2m-1}(\Omega)} + C \|\varphi_k u\|_{L^2(\Omega)} \\
&\le C \|\varphi_k \mathcal{A} u\|_{L^2(\Omega)} + C \|\psi_k u\|_{H^{2m-1}(\Omega)} .
\end{aligned}
$$

The boundary case: U_k intersects $\partial\Omega$

We recall the various coordinate systems, in the variables x and y and z. We also need names for the various operators:

- $\mathcal{A}$ and $\mathcal{B}_j$ are the operators we are actually interested in,
- $\tilde{\mathcal{A}}_{\mathrm{pr}}$ and $\tilde{\mathcal{B}}_{j,\mathrm{pr}}$ mean the following operators: in the z-language, we have thrown away the lower order terms, and we have frozen the coefficients of the highest order terms. If we then translate back into the x–language, we obtain $\tilde{\mathcal{A}}_{\mathrm{pr}}$ and $\tilde{\mathcal{B}}_{j,\mathrm{pr}}$.
- $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{B}}_j$: as before, but now we add back the lower order terms.

The operators $\mathcal{A}$ and $\tilde{\mathcal{A}}$ do not differ in the lower order terms; and their coefficients of the highest order terms differ in size $\mathcal{O}(\varepsilon)$. Similarly for $\mathcal{B}_j$ and $\tilde{\mathcal{B}}_j$.

The operators $(\tilde{\mathcal{A}}_{\mathrm{pr}}, \tilde{\mathcal{B}}_{1,\mathrm{pr}}, \dots \tilde{\mathcal{B}}_{m,\mathrm{pr}})$ constitute a uniquely solvable BVP, hence we know

$$
\|\varphi_k u\|_{H^{2m}(\Omega)} \le C \left(\left\| \tilde{\mathcal{A}}_{\mathrm{pr}}(\varphi_k u) \right\|_{L^2(\Omega)} + \sum_{j=1}^m \left\| \tilde{\mathcal{B}}_{j,\mathrm{pr}}(\varphi_k u) \right\|_{H^{2m-m_j-1/2}(\partial\Omega)} \right).
$$

Now we add back the lower order terms, which are operators of order $2m-1$ in the $\mathcal{A}$ case, and of order m_j-1 in the $\mathcal{B}_j$ case. Consequently,

$$\|\varphi_k u\|_{H^{2m}(\Omega)} \le C\left(\left\|\tilde{\mathcal{A}}(\varphi_k u)\right\|_{L^2(\Omega)} + \sum_{j=1}^m \left\|\tilde{\mathcal{B}}_j(\varphi_k u)\right\|_{H^{2m-m_j-1/2}(\partial\Omega)} + \|\varphi_k u\|_{H^{2m-1}(\Omega)}\right).$$

Now we repair the highest order terms in the operators:

$$\begin{aligned}&\|\varphi_k u\|_{H^{2m}(\Omega)}\\ &\quad\le C\left(\|\mathcal{A}(\varphi_k u)\|_{L^2(\Omega)} + \sum_{j=1}^m \|\mathcal{B}_j(\varphi_k u)\|_{H^{2m-m_j-1/2}(\partial\Omega)} + \varepsilon\,\|\varphi_k u\|_{H^{2m}(\Omega)} + \|\varphi_k u\|_{H^{2m-1}(\Omega)}\right).\end{aligned}$$

Note that the many constants C can change from place to place, but they do not depend on ε. Hence we can arrange that $0 < C\varepsilon \le \frac{1}{2}$, and then the corresponding term from the RHS can be brought to the LHS.

We can now unify both cases:

$$\|\varphi_k u\|_{H^{2m}(\Omega)} \le C\left(\|\varphi_k \mathcal{A} u\|_{L^2(\Omega)} + \sum_{j=1}^m \|\varphi_k \mathcal{B}_j u\|_{H^{2m-m_j-1/2}(\partial\Omega)} + \|\psi_k u\|_{H^{2m-1}(\Omega)}\right).$$

We square it, and note that the RHS contains a sum of $m+2$ items $t_1, \ldots, t_{m+2}$, for which we recall the classic inequality between the arithmetic mean and the quadratic mean:

$$\frac{t_1+\cdots+t_{m+2}}{m+2} \le \sqrt{\frac{t_1^2+\cdots+t_{m+2}^2}{m+2}}, \quad \text{hence} \quad (t_1+\cdots+t_{m+2})^2 \le (m+2)(t_1^2+\cdots+t_{m+2}),$$

from which we then deduce that (changing C by a factor depending only on m)

$$\|\varphi_k u\|^2_{H^{2m}(\Omega)} \le C^2\left(\|\varphi_k \mathcal{A} u\|^2_{L^2(\mathbb{R}^n)} + \sum_{j=1}^m \|\varphi_k \mathcal{B}_j u\|^2_{H^{2m-m_j-1/2}(\partial\Omega)} + \|\psi_k u\|^2_{H^{2m-1}(\Omega)}\right).$$

Now we sum over k,

$$\sum_{k=1}^K \|\varphi_k u\|^2_{H^{2m}(\Omega)} \le C^2 \sum_{k=1}^K \left(\|\varphi_k \mathcal{A} u\|^2_{L^2(\Omega)} + \sum_{j=1}^m \|\varphi_k \mathcal{B}_j u\|^2_{H^{2m-m_j-1/2}(\partial\Omega)} + \|\psi_k u\|^2_{H^{2m-1}(\Omega)}\right),$$

and the first term on the RHS can be handled by $\varphi_k^2 \le \varphi_k$, because of $0 \le \varphi_k \le 1$:

$$\sum_{k=1}^K \|\varphi_k \mathcal{A} u\|^2_{L^2(\Omega)} = \int_{\Omega_x} \sum_{k=1}^K \varphi_k^2(x)\left|(\mathcal{A}u)(x)\right|^2 \mathrm{d}x \le \int_{\Omega_x} \sum_{k=1}^K \varphi_k(x)\left|(\mathcal{A}u)(x)\right|^2 \mathrm{d}x = \|\mathcal{A}u\|^2_{L^2(\Omega)}.$$

The remaining two items on the RHS are to be treated in a similar style, but now it will happen that some derivatives will act upon φ_k and ψ_k, leading to large factors if these cut-off functions are steep.

Now it remains to handle the LHS:

$$\begin{aligned}\sum_{k=1}^K \|\varphi_k u\|^2_{H^{2m}(\Omega)} &= \sum_{k=1}^K \sum_{|\alpha|\le 2m} \int_\Omega \left|\partial_x^\alpha(\varphi_k u)\right|^2 \mathrm{d}x\\ &\ge \frac{1}{2}\sum_{k=1}^K \sum_{|\alpha|\le 2m} \int_\Omega \left|\varphi_k \partial_x^\alpha u\right|^2 \mathrm{d}x - C\sum_{k=1}^K \|\psi_k u\|^2_{H^{2m-1}}\\ &= \frac{1}{2}\sum_{|\alpha|\le 2m} \int_\Omega \left(\sum_{k=1}^K \varphi_k^2(x)\right)\cdot\left|\partial_x^\alpha u\right|^2 \mathrm{d}x - C\sum_{k=1}^K \|\psi_k u\|^2_{H^{2m-1}}\\ &\ge \frac{1}{2N_0}\sum_{|\alpha|\le 2m} \int_\Omega \left(\sum_{k=1}^K \varphi_k(x)\right)^2\cdot\left|\partial_x^\alpha u\right|^2 \mathrm{d}x - C\sum_{k=1}^K \|\psi_k u\|^2_{H^{2m-1}}.\end{aligned}$$

This then will conclude the proof of (6.4).

6.4 Indications on a General Microlocal Theory of BVPs

Let us have a look at

- the operators that constitute an (elliptic) BVP,
- the operators that solve it,

and try to get a unified perspective, in the language of pseudodifferential calculus.

We have a domain $\Omega \subset \mathbb{R}^n$ with smooth (C^∞) boundary $\partial\Omega$. The easiest case is $\Omega = \mathbb{R}^n_+ := \{(x_1, \dots, x_n) \colon x_n > 0\}$. Further, we have a (scalar or matrix) operator $\mathcal{A}(x, D_x)$ in Ω with smooth coefficients, and a boundary operator $\mathcal{B}(x, D_x)$. We wish to solve

$$\begin{cases} \mathcal{A}u = f, & \text{in } \Omega, \\ \mathcal{B}u = g, & \text{on } \partial\Omega, \end{cases}$$

for given f, g. The functions u, f, g can be scalar or vectorial, and $\mathcal{A}$, $\mathcal{B}$ can be scalars or matrices of appropriate shape.

The best we can hope for is: for each $f \in \overline{C_b^\infty(\Omega)}$, $g \in C_b^\infty(\partial\Omega)$, there is a unique solution $u \in \overline{C_b^\infty(\Omega)}$. Here $\overline{C_b^\infty(\Omega)}$ consists of all functions from $C^\infty(\Omega)$ that are bounded together with all their derivatives, and all derivatives can be extended continuously from Ω to the closure $\overline{\Omega}$.

If $\mathcal{A}$ is a PDO, then $\mathcal{A}$ maps $\overline{C_b^\infty(\Omega)} \to \overline{C_b^\infty(\Omega)}$, and $\mathcal{B}$ maps $\overline{C_b^\infty(\Omega)} \to C_b^\infty(\partial\Omega)$. We write this as

$$\begin{pmatrix} \mathcal{A} \\ \mathcal{B} \end{pmatrix} : \quad \overline{C_b^\infty(\Omega)} \quad \to \quad \begin{matrix} \overline{C_b^\infty(\Omega)} \\ \oplus \\ C_b^\infty(\partial\Omega), \end{matrix}$$

and $\binom{\mathcal{A}}{\mathcal{B}}$ is to be read as a matrix with one column and two rows, which maps from a scalar object to a two-component vectorial object.

Can we make $\mathcal{A}$ a ΨDO ? Then $\mathcal{A}$ is a nonlocal operator, and care is needed. Suppose that $\sigma(\mathcal{A})$ is known on $\mathbb{R}^n_x \times \mathbb{R}^n_\xi$. To define the action of $\mathcal{A}$ upon a function $u \in \overline{C_b^\infty(\Omega)}$, we define the extension e_+u from Ω to $\mathbb{R}^n$,

$$(e_+u)(x) := \begin{cases} u(x) & : x \in \Omega, \\ 0 & : x \notin \Omega, \end{cases}$$

which is an expression to which we can apply the ΨDO $\mathcal{A}$. The resulting function $\mathcal{A}(e_+u)$ is then restricted from $\mathbb{R}^n$ to Ω, using a restriction operator r_+. Instead of $\mathcal{A}$ (to be applied upon functions defined on $\mathbb{R}^n$) we therefore consider the operator

$$\mathcal{A}_+ := r_+ \circ \mathcal{A} \circ e_+,$$

to be applied upon functions from $\overline{C_b^\infty(\Omega)}$.

Does each ΨDO $\mathcal{A}$ generate a restricted operator $\mathcal{A}_+$ that maps $\overline{C_b^\infty(\Omega)}$ into itself ? No, already $\langle D_x \rangle$ is a counter-example. The reason is (take $\Omega = \mathbb{R}^n_+$ for sake of clarity): each function $u \in \overline{C_b^\infty(\Omega)}$ allows for a Taylor expansion (in the normal variable x_n) at the boundary:

$$\forall L \in \mathbb{N}\colon \quad u(x_1, \dots, x_n) = \sum_{\ell=0}^{L} c_\ell(x_1, \dots, x_{n-1}) \cdot x_n^\ell + \mathcal{O}(x_n^{L+1}), \qquad \text{for } x_n \to 0_{+0},$$

with some coefficients c_ℓ that depend on the tangential variables $(x_1, \dots, x_{n-1})$.

Now if $\mathcal{A}$ is a general ΨDO (even if it is a classical operator), it may happen that $\mathcal{A}_+u$ has an asymptotic expansion in x_n that not only contains the powers x_n^0, x_n^1, x_n^2, $\dots$, but also additional terms such as $x_n^3 \log x_n$, $x_n^3(\log x_n)^2$, and other stuff (which we do not want).

We wish that $\mathcal{A}_+$ "preserves Taylor asymptotics", and for this it is necessary and sufficient that $\sigma(\mathcal{A})$ fulfils a so-called "transmission condition". Then we indeed find that

$$\begin{pmatrix} \mathcal{A}_+ \\ \mathcal{B} \end{pmatrix} : \quad \overline{C_b^\infty(\Omega)} \quad \to \quad \begin{matrix} \overline{C_b^\infty(\Omega)} \\ \oplus \\ C_b^\infty(\partial\Omega). \end{matrix} \tag{6.10}$$

The operator $\mathcal{A}_+$ here is a ΨDO that satisfies the transmission condition, and $\mathcal{B}$ is called a *trace operator*.

Now we look at how to solve BVPs. Our example is

$$\begin{cases} \triangle u = f, & \text{in } \Omega, \\ \quad u = g, & \text{on } \partial\Omega. \end{cases}$$

Solving this BVP is done in two steps: we ignore the boundary condition, extend f to $\mathbb{R}^n$ by zero, then apply (1.2) (to the extended version $e_+ f$ of f) and get $u_1(x) = (r_+ \mathcal{P} e_+ f)(x)$, with $\mathcal{P}$ being a ΨDO on $\mathbb{R}^n$ with the (non-smooth) symbol $\frac{1}{-|\xi|^2}$. For $f \in \overline{C_b^\infty(\Omega)}$, we have $u_1 \in C^\infty(\Omega)$, and since $\frac{1}{-|\xi|^2}$ satisfies the transmission condition, we even have $u_1 \in \overline{C_b^\infty(\Omega)}$.

By construction $\triangle u_1(x) = f(x)$ for $x \in \Omega$. But u_1 does not have Dirichlet values g at $\partial\Omega$, so we need a correction u_2 that solves

$$\begin{cases} \triangle u_2 = 0, & \text{in } \Omega, \\ \quad u_2 = g - u_1, & \text{on } \partial\Omega. \end{cases}$$

If Ω is a ball, then Proposition 1.1 tells us how to calculate u_2 from $g - u_1$: we apply a so-called *potential operator* $\mathcal{K}$ which maps functions defined on $\partial\Omega$ to functions defined in Ω:

$$u_2 = \mathcal{K}(g - u_1).$$

Such operators $\mathcal{K}$ exist for any smooth Ω. Then we set $u = u_1 + u_2$, and we write this as

$$\begin{aligned} u &= r_+ \mathcal{P} e_+ f + \mathcal{K}(g - r_+ \mathcal{P} e_+ f) \\ &= \Big(r_+ \circ \mathcal{P} \circ e_+ - \mathcal{K} \circ r_+ \circ \mathcal{P} \circ e_+\Big) f + \mathcal{K} g. \end{aligned}$$

In matrix notation, this then becomes

$$u = \begin{pmatrix} [r_+ \circ \mathcal{P} \circ e_+ - \mathcal{K} \circ r_+ \circ \mathcal{P} \circ e_+] & \mathcal{K} \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}.$$

The mapping matrix has one row and two columns, and it is being applied to a column vector with two entries.

The item $-\mathcal{K} \circ r_+ \circ \mathcal{P} \circ e_+$ is called *singular Green operator*[4], and it takes a function defined in Ω, from which it produces again a function defined in Ω.

This solving procedure (first determine u_1 solving the interior problem, then finding a correction term u_2 to take care of the boundary conditions) is applicable to *any* elliptic BVP, and we get a solution operator in matrix form

$$\begin{pmatrix} [\mathcal{P}_+ + \mathcal{G}] & \mathcal{K} \end{pmatrix} : \begin{matrix} \overline{C_b^\infty(\Omega)} \\ \oplus \\ C_b^\infty(\partial\Omega) \end{matrix} \to \overline{C_b^\infty(\Omega)}, \tag{6.11}$$

with $\mathcal{P}$ as a ΨDO, $\mathcal{G}$ as a singular Green operator, and $\mathcal{K}$ as a potential operator.

Now comes the ingenious idea of Boutet de Monvel[5]: to bring the column matrix $\binom{\mathcal{A}_+}{\mathcal{B}}$ from (6.10) and the row matrix $\begin{pmatrix} [\mathcal{P}_+ + \mathcal{G}] & \mathcal{K} \end{pmatrix}$ from (6.11) into a 2×2 matrix:

$$\begin{pmatrix} [\mathcal{A}_+ + \mathcal{G}] & \mathcal{K} \\ \mathcal{B} & \boxed{?} \end{pmatrix} : \begin{matrix} \overline{C_b^\infty(\Omega)} \\ \oplus \\ C_b^\infty(\partial\Omega) \end{matrix} \to \begin{matrix} \overline{C_b^\infty(\Omega)} \\ \oplus \\ C_b^\infty(\partial\Omega). \end{matrix}$$

[4] George Green, 1793–1841
[5] Louis Boutet de Monvel, 1941–2014

The entries map like this:

$$\begin{aligned} \mathcal{A}_+ + \mathcal{G} &\colon \overline{C_b^\infty(\Omega)} \to \overline{C_b^\infty(\Omega)}, \\ \mathcal{K} &\colon C_b^\infty(\partial\Omega) \to \overline{C_b^\infty(\Omega)}, \\ \mathcal{B} &\colon \overline{C_b^\infty(\Omega)} \to C_b^\infty(\partial\Omega). \end{aligned}$$

The entry $\boxed{?}$ is not yet needed, and the only choice that makes sense is to keep it available for some mapping from $C_b^\infty(\partial\Omega)$ into itself, which will turn out to be a ΨDO defined on $\partial\Omega$.

Operators of the shape

$$\mathfrak{A} = \begin{pmatrix} [\mathcal{A}_+ + \mathcal{G}] & \mathcal{K} \\ \mathcal{B} & \mathcal{S} \end{pmatrix}$$

are called *Operators in the Boutet de Monvel calculus.*

The main results are:

- those operators $\mathfrak{A}$ form an algebra. This means that they form a vector space, and we have an operator multiplication which is compatible with the vector space operations. The operator multiplication is obtained by composing two such operators $\mathfrak{A}$ and $\tilde{\mathfrak{A}}$ in the sense of a 2×2 matrix-matrix product, and the resulting product is again an operator from the Boutet de Monvel calculus.
- taking adjoint operators stays in the calculus, under extra conditions.
- we have asymptotic expansions in descending orders of certain symbols.
- we have a concept of ellipticity, as invertibility of the principal symbol of $\mathfrak{A}$. This incorporates the Shapiro–Lopatinskij condition naturally.
- elliptic operators have parametrices, and these are again operators in the Boutet de Monvel calculus.

We have omitted all the details and technicalities. A nice introduction is [20], and the canonical reference is [6].

Chapter 7

Applications

7.1 General Principles

We recall what it means when we say that a collection of **scalar** operators $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$ constitute an elliptic BVP on a domain $\Omega \subset \mathbb{R}^n$: we have a PDO $\mathcal{A}$ of order $2m$ and PDOs $\mathcal{B}_j$ of order $m_j \leq 2m-1$, all having smooth coefficients (and all their derivatives are bounded), such that:

$\mathcal{A}$ is uniformly elliptic on $\overline{\Omega}$:

there is a positive constant C such that the principal symbol a_{pr} of $\mathcal{A}$ satisfies

$$|a_{\mathrm{pr}}(x,\xi)| \geq C, \quad \forall x \in \overline{\Omega}, \quad \forall |\xi| = 1;$$

$(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$ satisfy the Shapiro–Lopatinskij condition everywhere on $\partial\Omega$:

For every $x_0 \in \partial\Omega$, the following holds: at x_0, introduce a new coordinate frame $(y_1, \dots, y_n)$, with y' pointing tangentially, and y_n pointing in normal direction. Freeze the coefficients of $\mathcal{A}$ and $\mathcal{B}_j$ at $y = 0$, and consider the principal symbols $a_{\mathrm{pr}}(0,\eta)$ and $b_{j,\mathrm{pr}}(0,\eta)$. Then (for each $\eta' \in \mathbb{R}^{n-1} \setminus 0$) the ODE

$$\begin{cases} a_{\mathrm{pr}}(0,\eta',D_{y_n})v(y_n) = 0, & 0 \leq y_n < \infty, \\ \lim\limits_{y_n \to +\infty} v(y_n) = 0, & \\ b_{j,\mathrm{pr}}(0,\eta',D_{y_n})v(0) = 0, & j = 1,\dots,m, \end{cases}$$

possesses only the trivial solution $v(y_n) \equiv 0$.

We have seen in the previous section that the null space of this BVP then is a finite-dimensional space of smooth functions; and the same holds for the null space of the adjoint BVP; and we have the estimates (6.4) and (6.5).

Now let us consider the **matrix** case: we have operators $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$, with orders as above, but now each of them is a matrix differential operator of size $N \times N$. Again, all the coefficients are smooth functions, with all the derivatives being bounded. We say that this collection of matrix differential operators constitute an elliptic BVP on Ω if the following holds:

$\mathcal{A}$ is uniformly elliptic on $\overline{\Omega}$:

there is a positive constant C such that the $N \times N$ matrix-valued principal symbol a_{pr} of $\mathcal{A}$ satisfies

$$|\det a_{\mathrm{pr}}(x,\xi)| \geq C, \quad \forall x \in \overline{\Omega}, \quad \forall |\xi| = 1;$$

$(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$ satisfy the Shapiro–Lopatinskij condition everywhere on $\partial\Omega$:

For every $x_0 \in \partial\Omega$, the following holds: at x_0, introduce a new coordinate frame $(y_1, \dots, y_n)$, with y' pointing tangentially, and y_n pointing in normal direction. Freeze the coefficients of $\mathcal{A}$ and $\mathcal{B}_j$ at

$y = 0$, and consider the principal symbols $a_{\rm pr}(0,\eta)$ and $b_{j,{\rm pr}}(0,\eta)$. Then (for each $\eta' \in \mathbb{R}^{n-1}\setminus 0$) the system of N ODEs

$$\begin{cases} a_{\rm pr}(0,\eta',D_{y_n})v(y_n) = 0 \in \mathbb{C}^N, \quad 0 \le y_n < \infty, \\ \lim\limits_{y_n\to+\infty} v(y_n) = 0, \\ b_{j,{\rm pr}}(0,\eta',D_{y_n})v(0) = 0 \in \mathbb{C}^N, \quad j = 1,\dots,m, \end{cases}$$

possesses only the trivial solution $v(y_n) \equiv 0 \in \mathbb{C}^N$.

And again we are able to show that the null spaces of this BVP and its adjoint BVP are finite-dimensional spaces of smooth functions, and we have estimates very similar to (6.4) and (6.5). The proofs are basically the same. The only issue which requires some attention is that the matrix product is non-commutative.

Ellipticity means invertibility of the principal symbol.

In the scalar case, the principal symbol of $\mathcal{A}$ is a scalar function that must be separated away from zero; and in the matrix case, the principal symbol of $\mathcal{A}$ is an $N\times N$ matrix that must be uniformly invertible. The principal symbol of the parametrix $\mathcal{A}^\sharp$ of $\mathcal{A}$ then is the inverse matrix to $a_{\rm pr}(x,\xi)$, which then has again smooth and bounded matrix entries (which follows from the well-known formula of an inverse matrix via determinant and cofactor matrix).

We now have a look at a sub-class of elliptic systems. The condition $|\det a_{\rm pr}(x,\xi)| \ge C$ in particular means that 0 is no eigenvalue of the matrix principal symbol. In many realistic situations however, much more is true: all the eigenvalues of the principal symbol avoid a certain sector $\mathcal{L}$ of the complex plane.

Definition 7.1 (Parameter-elliptic BVP). *Let $\mathcal{L}$ be an open sector in $\mathbb{C}$, centred at the origin, augmented by 0:*

$$\mathcal{L} := \{z \in \mathbb{C} \colon \theta_1 < \arg z < \theta_2\} \cup \{0\}, \quad \textit{with some } \theta_1, \theta_2.$$

We say that $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$ is a parameter-elliptic BVP *in the sector $\mathcal{L}$ if $\mathcal{A}$ and $\mathcal{B}_j$ are matrix-valued PDOs with smooth coefficients, and orders $2m$ (for $\mathcal{A}$) and $m_j \le 2m-1$ (for $\mathcal{B}_j$) such that the following conditions hold:*

$\mathcal{A}$ is uniformly parameter-elliptic on $\overline{\Omega}$:

there is a positive constant C such that the $N\times N$ matrix-valued principal symbol $a_{\rm pr}$ of $\mathcal{A}$ satisfies

$$|\det(a_{\rm pr}(x,\xi) - \lambda\,{\rm id}_N)| \ge C, \quad \forall x \in \overline{\Omega}, \quad \forall(\xi,\lambda) \in \mathbb{R}^n \times \mathcal{L} \colon \ |\lambda| + |\xi| = 1;$$

$(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$ satisfy the Shapiro–Lopatinskij condition everywhere on $\partial\Omega$:

For every $x_0 \in \partial\Omega$, the following holds: at x_0, introduce a new coordinate frame $(y_1,\dots,y_n)$, with y' pointing tangentially, and y_n pointing in normal direction. Freeze the coefficients of $\mathcal{A}$ and $\mathcal{B}_j$ at $y = 0$, and consider the principal symbols $a_{\rm pr}(0,\eta)$ and $b_{j,{\rm pr}}(0,\eta)$. Then (for each $(\eta',\lambda) \in (\mathbb{R}^{n-1}\times\mathcal{L})\setminus(0,0)$) the system of N ODEs

$$\begin{cases} \Big(a_{\rm pr}(0,\eta',D_{y_n}) - \lambda\,{\rm id}_N\Big)v(y_n) = 0 \in \mathbb{C}^N, \quad 0 \le y_n < \infty, \\ \lim\limits_{y_n\to+\infty} v(y_n) = 0, \\ b_{j,{\rm pr}}(0,\eta',D_{y_n})v(0) = 0 \in \mathbb{C}^N, \quad j = 1,\dots,m, \end{cases}$$

possesses only the trivial solution $v(y_n) \equiv 0 \in \mathbb{C}^N$.

The probability is very high that any elliptic BVP which you have ever seen[1] in your career is actually parameter-elliptic in a certain sector.

[1] with the exception of the Cauchy–Riemann equations from complex analysis

Take for instance $\mathcal{A}(x, D_x) = \triangle + \sum_{k=1}^n b_k(x) D_{x_k} + c(x)$ with boundary condition $\mathcal{B}_1(x, D_x) = \frac{\partial}{\partial \nu} + \gamma(x)$. The principal parts are the Laplacian and the Neumann boundary condition. To check the Shapiro–Lopatinskij condition, we have to consider

$$a_{\mathrm{pr}}(0, \eta', D_{y_n}) = -|\eta'|^2 + \partial_{y_n}^2, \qquad b_{\mathrm{pr}}(0, \eta', D_{y_n}) = \partial_{y_n}.$$

With λ coming from a certain sector that will be determined later, we then have a look at

$$\begin{cases} \left(\partial_{y_n}^2 - |\eta'|^2 - \lambda\right) v(y_n) = 0, & 0 \le y_n < \infty, \\ \lim_{y_n \to +\infty} v(y_n) = 0, & \\ \partial_{y_n} v(0) = 0. & \end{cases}$$

We multiply the differential equation by $\overline{v(y_n)}$ and integrate over $(0, \infty)$:

$$0 = \int_0^\infty v'' \overline{v} - \left(|\eta'|^2 + \lambda\right) |v|^2 \,\mathrm{d}y_n = -\int_0^\infty |v'(y_n)|^2 + |\eta'|^2 |v(y_n)|^2 \,\mathrm{d}y_n - \lambda \int_0^\infty |v(y_n)|^2 \,\mathrm{d}y_n,$$

and from this we read directly: whenever $\Im \lambda \neq 0$, then $v \equiv 0$. And whenever λ is a positive real number, then also $v \equiv 0$.

The only numbers λ, for which the Shapiro–Lopatinskij condition could be violated, are the negative reals. Hence we have found that this BVP is parameter-elliptic in every sector

$$\mathcal{L} = \Sigma_{\pi - \varepsilon} := \{ z \in \mathbb{C} \colon |\arg(z)| < \pi - \varepsilon \}, \qquad 0 < \varepsilon < \pi.$$

Now comes the first result in this course which guarantees us the **existence** of a solution to some equation:

Theorem 7.2. *Consider a scalar operator $\mathcal{A}$ of order $2m$, and boundary operators $\mathcal{B}_1, \dots, \mathcal{B}_m$, of orders $m_1 < m_2 < \dots < m_m$. Fix some Lebesgue exponent $p \in (1, \infty)$. Suppose that the BVP $(\mathcal{A}, \mathcal{B}_1, \dots, \mathcal{B}_m)$ is parameter-elliptic in some sector $\mathcal{L}$.*

Then there is some (large) number $\lambda_0 \in \mathbb{R}_+$ such that the following holds: if $\lambda \in \mathcal{L}$ and $|\lambda| \ge \lambda_0$, then the BVP

$$\begin{cases} \mathcal{A}u - \lambda u = f & \text{in } \Omega, \\ \mathcal{B}_j u = g_j & \text{on } \partial\Omega, \quad j = 1, \dots, m, \end{cases}$$

is uniquely solvable for all $f \in L^p(\Omega)$, $g_j \in W_p^{2m - m_j - 1/p}(\partial\Omega)$; the solution u is in $W_p^{2m}(\Omega)$, and it enjoys the estimate

$$\|u\|_{W_p^{2m}(\Omega)} + |\lambda| \cdot \|u\|_{L^p(\Omega)} \le C_0 \left(\|f\|_{L^p(\Omega)} + \sum_{j=1}^m \|g_j\|_{W_p^{2m - m_j - 1/p}(\partial\Omega)} \right).$$

Here $W_p^{2m - m_j - 1/p}(\partial\Omega)$ is a fractional order Sobolev space on the boundary that occurs naturally as the function space of traces of functions from the integer order Sobolev space $W_p^{2m - m_j}(\Omega)$.

The proof is very long, but its main ideas are not so hard: we are already able to construct a local parametrix inside the domain Ω, and we know how to solve approximate problems in small neighbourhoods of a point at the boundary. We only have to glue the pieces together using many cut-off functions. Then additional terms will appear (which we have pushed into terms $\|u\|_{H^{2m-1}(\Omega)}$ when we studied the Shapiro–Lopatinskij condition), but they are no problem because "λ is in a good sector and large enough". Recall that we can interpolate

$$\|v\|_{W_p^{2m-1}(\Omega)} \le C \|v\|_{W_p^{2m}(\Omega)}^{\frac{2m-1}{2m}} \cdot \|v\|_{L^p(\Omega)}^{\frac{1}{2m}} \le \varepsilon \|v\|_{W_p^{2m}(\Omega)} + \frac{C}{\varepsilon^{\text{huge}}} \|v\|_{L^p(\Omega)},$$

for any function $v \in W_p^{2m}(\Omega)$. We can bring $\varepsilon \|v\|_{W_p^{2m}(\Omega)}$ to the LHS for small ε, and we can eliminate the other item using $|\lambda| \cdot \|u\|_{L^p}$ for large $|\lambda|$.

We come to some important conclusion.

Theorem 7.3. *Let X be a Banach space, and let $\mathcal{A}\colon D(\mathcal{A}) \to X$ be a closed operator, with dense domain $D(\mathcal{A}) \subset X$. Assume that there is a sector $\mathcal{L} = \Sigma_\theta$ with $\theta > \pi/2$, such that for each $\lambda \in \mathcal{L}$, the resolvent $(\mathcal{A} - \lambda\,\mathrm{id})^{-1}$ exists as a bounded map of X into $D(\mathcal{A})$, with an operator norm estimate*

$$\left\|(\mathcal{A} - \lambda\,\mathrm{id})^{-1}\right\|_{X\to X} \le \frac{C}{|\lambda|}, \qquad \forall\lambda \in \mathcal{L},$$

and C is independent of λ. Then the operator $\mathcal{A}$ generates an analytic semigroup on X.

This implies then in particular that initial-value problems

$$\begin{cases} \partial_t u(t) - \mathcal{A}u(t) = f(t), & 0 < t < T, \\ \qquad\qquad u(0) = u_0, \end{cases}$$

with given $u_0 \in X$ and given f become meaningfully solvable, assuming some conditions on u_0 and f, such as $u_0 \in D(\mathcal{A})$ and $f \in C^\alpha([0,T], X)$, for some Hölder exponent $\alpha \in (0,1)$.

In case of the above Neumann Laplacian, we have $X = L^p(\Omega)$ and $D(\mathcal{A}) = \{u \in W^2_p(\Omega)\colon \mathcal{B}u = 0\}$. Some shifted operator $\mathcal{A}$ + const. then generates an analytic semigroup on X, hence also $\mathcal{A}$.

7.2 Population Models

There seem to be some diverging concepts of ellipticity in the mathematics community. In this course, ellipticity always means invertibility of the principal symbol. For a scalar operator $\mathcal{A}$ of second order in $\Omega \subset \mathbb{R}^n$, the principal symbol is $a_{\mathrm{pr}}(x,\xi) = \sum_{j,k=1}^n a_{jk}(x)\xi_j\xi_k$, and then ellipticity implies

$$|a_{\mathrm{pr}}(x,\xi)| \ge c(x)|\xi|^2, \qquad x \in \Omega, \quad \xi \in \mathbb{R}^n,$$

for some $c(x) > 0$. Now this keeps the possibility open of $c(x)$ approaching zero for x approaching $\partial\Omega$, and typically this is an unwanted degeneracy. To exclude this behaviour, the concept of *uniform ellipticity* has been invented, which requires:

$$\exists c > 0\colon\ \forall (x,\xi) \in \Omega \times \mathbb{R}^n\colon\ |a_{\mathrm{pr}}(x,\xi)| \ge c|\xi|^2.$$

Ellipticity does not always imply that some quadratic form is positive definite.

In case of scalar second order operators with principal symbol $a_{\mathrm{pr}}(x,\xi) = \sum_{j,k=1}^n a_{jk}(x)\xi_j\xi_k$, we can naturally assume $a_{jk} = a_{kj}$ because of $\partial_{x_j}\partial_{x_k}u = \partial_{x_k}\partial_{x_j}u$ (valid for classical derivatives and weak derivatives and distributional derivatives, see Remark 2.22), and in most applied situations, the coefficients a_{jk} are real-valued anyway, and then the matrix with entries a_{jk} is either uniformly positive definite or uniformly negative definite.

The situation changes completely for second order **matrix** differential operators. Let us consider a population model with two population densities u_1 and u_2, and their fluxes J_1, J_2, in $\mathbb{R}^n$:

$$\begin{aligned} \partial_t u_1 - \operatorname{div} J_1 &= 0, & J_1 &= \nabla\Big((\delta_1 + \delta_{11}u_1 + \delta_{12}u_2)u_1\Big), \\ \partial_t u_2 - \operatorname{div} J_2 &= 0, & J_2 &= \nabla\Big((\delta_2 + \delta_{21}u_1 + \delta_{22}u_2)u_2\Big), \end{aligned}$$

where we have neglected many lower-order terms. We may write this as

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \operatorname{div}\left(A \cdot \begin{pmatrix} \nabla u_1 \\ \nabla u_2 \end{pmatrix}\right) = 0, \qquad A = \begin{pmatrix} \delta_1 + 2\delta_{11}u_1 + \delta_{12}u_2 & \delta_{12}u_1 \\ \delta_{21}u_2 & \delta_2 + \delta_{21}u_1 + 2\delta_{22}u_2 \end{pmatrix}.$$

The principal symbol of the matrix differential operator is

$$\sigma_{\mathrm{pr}}(\mathcal{A}) = -\begin{pmatrix} \delta_1 + 2\delta_{11}u_1 + \delta_{12}u_2 & \delta_{12}u_1 \\ \delta_{21}u_2 & \delta_2 + \delta_{21}u_1 + 2\delta_{22}u_2 \end{pmatrix} |\xi|^2,$$

and this matrix is indeed invertible for δ_1, $\delta_2 > 0$, all other $\delta_{jk} \geq 0$, and u_1, $u_2 \geq 0$. It is easy to check that its eigenvalues are in a sector that is strictly smaller than the left half-plane of $\mathbb{C}$, which makes this problem parabolic. And it is obviously quasi-linear. The methods of [5] show us how to solve it: we choose the functional-analytical base space $X = L^p(\mathbb{R}^n)$ with $p > n$, in which $\mathcal{A}$ generates an analytic semi-group (for fixed u_1, $u_2 > 0$), and then the existence and uniqueness of a solution (u_1, u_2) can be proven for short times. The assumption $p > n$ gives us the embedding $W_p^1 \subset L^\infty$, which enables us to handle the nonlinearities. And this solution persists as long as it does not explode in certain norms and stays positive. All this follows from the general machinery of [5].

But in the case of strong cross-diffusion, the operator $\mathcal{A}$ does not generate a positive definite form. Take for instance $\delta_1 = \delta_2 = 1$, the cross-diffusion coefficients $\delta_{12} = \delta_{21} = 1$, and the self-diffusion coefficients $\delta_{11} = \delta_{22} = 0$. Take $u_1 = 20$ and $u_2 = 80$ in some region of $\mathbb{R}^n$. Then

$$A = \begin{pmatrix} 81 & 20 \\ 80 & 21 \end{pmatrix},$$

which is invertible, but it generates the quadratic form

$$\begin{pmatrix} \nabla u_1 \\ \nabla u_2 \end{pmatrix} \mapsto 81|\nabla u_1|^2 + 20(\nabla u_1)(\nabla u_2) + 80(\nabla u_2)(\nabla u_1) + 21|\nabla u_2|^2,$$

which is not positive definite, because its matrix

$$\begin{pmatrix} 81 & 50 \\ 50 & 21 \end{pmatrix}$$

has one negative eigenvalue (its determinant is negative).

On a more general level: a second order matrix differential operator $\mathcal{A}(x, D_x)$, with coefficients $a_{jk}(x)$ (each a_{jk} being an $N \times N$ matrix with complex entries a_{jk}^{rs}), is called *uniformly strongly elliptic* if

$$\Re\Big(\sum_{r,s=1}^{N} \sum_{j,k=1}^{n} a_{jk}^{rs}(x)\xi_j\xi_k\lambda_r\overline{\lambda_s} \Big) \geq c|\xi|^2 \cdot |\lambda|^2,$$

for some $c > 0$ and all $x \in \Omega$, all $\xi \in \mathbb{R}^n$, all $\lambda \in \mathbb{C}^N$. The population model with strong cross-diffusion has shown that ellipticity (which means invertibility of the pseudodifferential principal symbol) does not imply strong ellipticity (which means that the usual quadratic form obtained by partial integration is positive definite).

However, there is some general procedure which is sometimes helpful. Consider the problem

$$\partial_t U - \mathcal{A}(x, D_x)U = 0,$$

where U is some vector function $U\colon \mathbb{R}_t \times \mathbb{R}_x^n \to \mathbb{C}^N$, and $\mathcal{A}$ is some matrix ΨDO of size $N \times N$. Nothing has been said about the order of $\mathcal{A}$, or whether this problem is hyperbolic or parabolic or hybrid structure (think thermo-elasticity). If $\langle \cdot, \cdot \rangle$ denotes the usual $L^2(\mathbb{R}^n \to \mathbb{C}^N)$ scalar product, and U is a solution, then

$$\begin{aligned} \partial_t \|U\|^2 &= \langle \partial_t U, U\rangle + \langle U, \partial_t U\rangle = 2\Re\langle \partial_t U, U\rangle = 2\Re\langle \mathcal{A}U, U\rangle = \langle \mathcal{A}U, U\rangle + \overline{\langle \mathcal{A}U, U\rangle} \\ &= \langle \mathcal{A}U, U\rangle + \langle U, \mathcal{A}U\rangle = \langle \mathcal{A}U, U\rangle + \langle \mathcal{A}^*U, U\rangle, \end{aligned}$$

with $\mathcal{A}^*$ being the adjoint operator to the ΨDO $\mathcal{A}$.

This is the usual approach, which is not very helpful if $\mathcal{A}$ is elliptic, but not strongly elliptic. But consider now an arbitrary matrix ΨDO $\mathcal{M}(x, D_x)$, perhaps with symbol from S_{cl}^0, and define $V(x) = \mathcal{M}(x, D_x)U(t, x)$. Suppose that an inverse operator $\mathcal{M}^{-1}$ to $\mathcal{M}$ exists. Then we find

$$\partial_t V = \mathcal{M}\partial_t U = \mathcal{M}\mathcal{A}U = \Big(\mathcal{M}\mathcal{A}\mathcal{M}^{-1}\Big)V =: \tilde{\mathcal{A}}V,$$

and we can also calculate (as above)

$$\partial_t \|V\|^2 = \left\langle \left(\tilde{\mathcal{A}} + \tilde{\mathcal{A}}^* \right) V, V \right\rangle .$$

Note that $\tilde{\mathcal{A}}$ is a matrix ΨDO with principal matrix symbol

$$\sigma_{\rm pr}(\tilde{\mathcal{A}}) = \sigma_{\rm pr}(\mathcal{M}) \cdot \sigma_{\rm pr}(\mathcal{A}) \cdot (\sigma_{\rm pr}(\mathcal{M}))^{-1},$$

and therefore $\sigma_{\rm pr}(\tilde{\mathcal{A}})$ has the same eigenvalues as $\sigma_{\rm pr}(\mathcal{A})$. If you choose $(\sigma_{\rm pr}(\mathcal{M}))^{-1}$ as the collection of the eigenvectors of $\sigma_{\rm pr}(\mathcal{A})$, then $\tilde{\mathcal{A}}$ has a diagonal principal part, and then there is hope that estimating $\|V\|$ allows for much better estimates of $\|U\|$ because the system for V becomes de-coupled in the principal part.

Of course, there are challenges:

- What to do if there not enough eigenvectors of $\sigma_{\rm pr}(\mathcal{A})$? This happens for so-called weakly hyperbolic systems.
- What to do if $\mathcal{A}$ depends on time ? Then also $\mathcal{M}$ will depend on time, and more terms appear.
- Are the lower order terms really harmless ?
- What to do if $\mathcal{A}$ depends itself on U, as in the population model ?

In particular in the nonlinear cases, various entropies have been applied with great success; but the key idea behind their choice has often been to diagonalise some matrix differential operator. And if you cannot diagonalise that operator, at least make it symmetric.

7.3 Quantum Hydrodynamics

This part is an edited version of [9].

The (viscous) model of quantum hydrodynamics describes the transport of electrons in various semiconductor devices, and it contains the unknowns n (scalar electron density), J (vectorial current density), V (scalar electric potential). The system reads

$$\begin{cases} \partial_t n - \operatorname{div} J = \nu_1 \triangle n, \\ \partial_t J - \operatorname{div}\left(\dfrac{J \otimes J}{n}\right) - \nabla p(n) + n\nabla V + \dfrac{\varepsilon^2}{2} n \nabla \left(\dfrac{\triangle \sqrt{n}}{\sqrt{n}}\right) = \nu_2 \triangle J - \dfrac{J}{\tau}, \\ \lambda_D^2 \triangle V = n - C(x), \end{cases} \tag{7.1}$$

for $(t,x) \in (0,T_0) \times \Omega$, where the spatial domain Ω is an open subset of $\mathbb{R}^d$, $d = 1,2,3$, with smooth boundary. The three equations can be understood as conservation of mass, conservation of momentum, and an elliptic equation connecting the electric potential V to the density of the electric charges.

In the system $p = p(n)$ is a pressure term, typically having the form $p(n) = \text{const.}\, n^\gamma$, for some $\gamma \geq 1$. The viscosity constants $\nu_{1,2}$ are typically equal, and they describe the quantum mechanical interaction of the electrons with oscillations of the semiconductor crystal. The third order derivative term is called *Bohm potential*[2],

$$B(n) = n\nabla \frac{\triangle \sqrt{n}}{\sqrt{n}} = \frac{1}{2}\nabla \triangle n - \frac{1}{2} \operatorname{div}\left(\frac{(\nabla n) \otimes (\nabla n)}{n}\right).$$

The parameter τ is the relaxation time of the current variable J, and λ_D is called *Debye length*[3]. The function C is known and models the density of positive ions in the crystal.

We prescribe certain boundary conditions, and initial values

$$n(0,x) = n_0(x), \qquad J(0,x) = J_0(x), \qquad x \in \Omega, \tag{7.2}$$

[2] David Bohm, 1917–1992
[3] Peter Joseph William Debye, 1884–1966, Nobel prize 1936

The linear principal part of the steady state problem of quantum hydrodynamics is

$$\mathcal{A}\begin{pmatrix} n(x) \\ J(x) \end{pmatrix} = \begin{pmatrix} f_0(x) \\ f'(x) \end{pmatrix}, \qquad x \in \Omega,$$

where A is given as

$$\mathcal{A} = \begin{pmatrix} \nu_1 \triangle & \operatorname{div} \\ -\frac{\varepsilon^2}{4}\nabla\triangle & \nu_2 \triangle \operatorname{id}_d \end{pmatrix}. \tag{7.3}$$

This is a mixed order matrix differential operator, and the *mixed order* refers to the third order derivative in the lower left corner, which is “in a certain sense compensated” by the first order derivative in the upper right corner. Below we will define what an elliptic BVP for such an operator is, in the sense of a Shapiro–Lopatinskij criterion.

We assume that the viscosity parameters satisfy

$$\begin{cases} \nu_1, \nu_2 \geq 0, \quad \nu_1 + \nu_2 > 0, & : d = 1, \\ \nu_1 \geq 0, \quad \nu_2 > 0, & : d \geq 2. \end{cases} \tag{7.4}$$

Theorem 7.4. *Suppose (7.4) and $\varepsilon \neq 0$. Then for $d \geq 1$, each of the following boundary conditions on n and J satisfies the Shapiro–Lopatinskij criterion:*

$$\begin{aligned} n_{|\partial\Omega} &= n_\Gamma, & J_{|\partial\Omega} &= J_\Gamma, \\ \partial_\nu n_{|\partial\Omega} &= n_\Gamma, & J_{|\partial\Omega} &= J_\Gamma, \\ n_{|\partial\Omega} &= n_\Gamma, & (J_\|)_{|\partial\Omega} &= J_{\|,\Gamma}, \quad (\partial_\nu J_\perp)_{|\partial\Omega} = J_{\perp,\Gamma}, \quad (\nu_1 > 0), \end{aligned}$$

where $J_\|$ and $J_\perp$ are the components of J tangential and perpendicular to $\partial\Omega$.

There is a sector Σ_ϑ with

$$\Sigma_\vartheta = \{z \in \mathbb{C} \setminus \{0\} : |\arg z| < \vartheta\}, \qquad \frac{\pi}{2} < \vartheta < \pi, \tag{7.5}$$

and for each $p \in (1, \infty)$ there is a positive $\lambda_0(p)$ such that: for zero boundary values and $\lambda \in \Sigma_\vartheta$ with $|\lambda| \geq \lambda_0(p)$, the solution (n, J) to the problem

$$\mathcal{A}\begin{pmatrix} n \\ J \end{pmatrix} - \lambda \begin{pmatrix} n \\ J \end{pmatrix} = \begin{pmatrix} f_0 \\ f' \end{pmatrix} \in W_p^1(\Omega) \times (L^p(\Omega))^d,$$

with homogeneous boundary conditions, exists in $W_p^3(\Omega) \times (W_p^2(\Omega))^d$ and enjoys the a priori *estimate*

$$\|n\|_{W_p^3(\Omega)} + |\lambda|^{3/2}\|n\|_{L^p(\Omega)} + \|J\|_{W_p^2(\Omega)} + |\lambda|\,\|J\|_{L^p(\Omega)} \leq C\left(\|f_0\|_{W_p^1(\Omega)} + |\lambda|^{1/2}\|f_0\|_{L^p(\Omega)} + \|f'\|_{L^p(\Omega)}\right). \tag{7.6}$$

This *a priori* estimate then will enable us to verify the operator $\mathcal{A}$ as the generator of a semigroup. Our approach is as follows: Let $\mathcal{B}_n$ and $\mathcal{B}_J$ be boundary condition operators, with either $\mathcal{B}_n = 1$ or $\mathcal{B}_n = \partial_\nu$, and either $\mathcal{B}_J = I_d$, $\mathcal{B}_J = I_d\partial_\nu$, or $\mathcal{B}_J = (P_\|, \partial_\nu P_\perp)$, with $P_\|$ and $P_\perp$ being the projectors onto the tangential and normal parts of a vector field.

For $1 < p < \infty$, we consider the operator $\mathcal{A}$ from (7.3) with domain

$$D(\mathcal{A}) = \left\{(n, J) \in W_p^3(\Omega) \times (W_p^2(\Omega))^d : \ (\mathcal{B}_n n)_{|\partial\Omega} = 0, \ (\mathcal{B}_J J)_{|\partial\Omega} = 0\right\}.$$

Theorem 7.5. *Assume (7.4) and $\varepsilon \neq 0$. Then $\mathcal{A}$ generates an analytic semigroup on the space*

$$X = \left\{(f_0, f') \in W_p^1(\Omega) \times (L^p(\Omega))^d : \ (\mathcal{B}_n f_0)_{|\partial\Omega} = 0\right\}.$$

This brings us in a position to study local well–posedness to the system (7.1) with initial conditions (7.2) and boundary conditions

$$\begin{cases} (\mathcal{B}_n n)(t, x) = n_\Gamma(x), \\ (\mathcal{B}_J J)(t, x) = J_\Gamma(x), \qquad (t, x) \in (0, T_0) \times \partial\Omega. \\ V(t, x) = V_\Gamma(x), \end{cases} \tag{7.7}$$

Theorem 7.6. *We suppose that the initial data possess the regularity* $n_0 \in W_p^3(\Omega)$*,* $J_0 \in W_p^2(\Omega)$*; and for the boundary data we assume* $n_\Gamma \in W_p^{3-\operatorname{ord}\mathcal{B}_n-1/p}(\partial\Omega)$*,* $J_\Gamma \in W_p^{2-\operatorname{ord}\mathcal{B}_J-1/p}(\partial\Omega)$ *as well as* $V_\Gamma \in W_p^{2-1/p}(\partial\Omega)$*, where* $p > d$*. The doping profile* C *is assumed to be an* $L^p(\Omega)$ *function. Moreover, suppose* $\inf_{x\in\Omega} n_0(x) > 0$ *and the compatibility conditions*

$$(\mathcal{B}_n n_0)(x) = n_\Gamma(x), \qquad (\mathcal{B}_J J_0)(x) = J_\Gamma(x), \qquad x \in \partial\Omega.$$

Then the problem (7.1), (7.2), (7.7) has a unique time–local classical solution (n, J, V) *with*

$$\begin{aligned} n &\in C([0,T_0], W_p^3(\Omega)), & \partial_t n &\in C([0,T_0], W_p^1(\Omega)),\\ J &\in C([0,T_0], W_p^2(\Omega)), & \partial_t J &\in C([0,T_0], L^p(\Omega)),\\ V &\in C([0,T_0], W_p^2(\Omega)), & \partial_t V &\in C([0,T_0], W_p^3(\Omega)),\end{aligned}$$

for some positive T_0*.*

Now we explain our concept of ellipticity. We consider an $N \times N$ matrix differential operator $\mathcal{A}(x, D_x)$ consisting of entries $a_{jk}(x, D_x)$ with $D = \frac{1}{\mathrm{i}}\nabla$ as usual. We suppose that there are integers $s_1, \dots, s_N$ and $m_1, \dots, m_N$ such that $s_j + m_j =: m \in \mathbb{N}_+$ is independent of j, with the property that the order of a_{jk} is no more than $s_j + m_k$ for all $j, k = 1, \dots, N$. We do not lose generality if we suppose that $m_1 \geq m_2 \geq \ldots \geq m_N = 0$. Additionally, we assume that $a_{jk} \equiv 0$ in case of $s_j + m_k < 0$.

We wish to solve the system of partial differential equations

$$(\mathcal{A}(x, D_x) - \lambda\,\mathrm{id}_N)u(x) = f(x), \qquad x \in \Omega, \tag{7.8}$$

for all λ in a certain sector $\mathcal{L}$ of $\mathbb{C}$. This interior problem is complemented with boundary conditions

$$\mathcal{B}_j(x, D_x)u(x) = g_j(x), \qquad x \in \partial\Omega, \quad j = 1, \dots, mN/2 \in \mathbb{N}, \tag{7.9}$$

where $\mathcal{B}_j$ is a $1 \times N$ matrix differential operator with entries b_{jk}, $k = 1, \dots, N$, whose order does not exceed $r_j + m_k$. Here we assume that such numbers $r_1, \dots, r_{mN/2} \in \mathbb{Z}$ exist with $r_j < m$ and that $b_{jk} \equiv 0$ in case of $r_j + m_k < 0$.

Through this section, we assume that the coefficients of $\mathcal{A}$ and $\mathcal{B}_j$ belong to $C^\infty(\overline{\Omega})$ and $\Omega \subset \mathbb{R}^n$, $\partial\Omega \in C^{\max m_j + m}$.

Definition 7.7. *Let* $\mathcal{L}$ *be a closed sector in the complex plane with vertex at the origin. Write* $a_{\mathrm{pr},jk}(x, D_x)$*,* $b_{\mathrm{pr},jk}(x, D_x)$ *for the principal parts of* a_{jk} *and* b_{jk}*, with* $\operatorname{ord} a_{\mathrm{pr},jk} = s_j + m_k$ *and* $\operatorname{ord} b_{\mathrm{pr},jk} = r_j + m_k$*. Let* $\mathcal{A}_{\mathrm{pr}}$ *and* $\mathcal{B}_{\mathrm{pr},j}$ *be the* $N \times N$ *and* $1 \times N$ *matrices with entries* $a_{\mathrm{pr},jk}$ *and* $b_{\mathrm{pr},jk}$*. Their pseudodifferential symbols are* a_{pr} *and* $b_{\mathrm{pr},j}$*.*

The boundary value problem (7.8), (7.9) is called elliptic with parameter in the sector $\mathcal{L}$ *if the following conditions hold:*

interior ellipticity condition: $\det(a_{\mathrm{pr}}(x,\xi) - \lambda\,\mathrm{id}_N) \neq 0$ *for all* $(x, \xi, \lambda) \in \overline{\Omega} \times \mathbb{R}^n \times \mathcal{L}$ *with* $|\xi| + |\lambda| > 0$*.*

Shapiro–Lopatinskij condition: *Let* $x^0 \in \partial\Omega$ *and the system (7.8), (7.9) be rewritten in local coordinates near* x^0 *(using a translation and a rotation), in such a way that the boundary at* x^0 *corresponds to* $x_n = 0$*, and the interior normal vector corresponds to the half–axis with* $x_n > 0$*. Then the boundary value problem on the half–line*

$$\begin{cases} a_{\mathrm{pr}}(0, \xi', D_{x_n})v(t) - \lambda v(t) = 0, & 0 < t = x_n < \infty,\\ \lim_{t\to+\infty} v(t) = 0, & \\ b_{\mathrm{pr},j}(0, \xi', D_{x_n})v(t) = 0, & t = 0, \quad j = 1, \dots, mN/2, \end{cases} \tag{7.10}$$

has only the trivial solution $v \equiv 0$*, for all* $(x^0, \xi', \lambda) \in \partial\Omega \times \mathbb{R}^{n-1} \times \mathcal{L}$ *with* $|\xi'| + |\lambda| > 0$*.*

In [3], it has been shown that the interior ellipticity condition implies $mN \in 2\mathbb{N}$.

For a vector–valued function u on Ω of regularity $W_p^2(\Omega)$, where $1 < p < \infty$ and $s \in \mathbb{N}_0$, we define a parameter–dependent norm,

$$\|u\|_{s,p,\Omega,\lambda} = \|u\|_{W_p^s(\Omega)} + |\lambda|^{s/m}\, \|u\|_{L^p(\Omega)}\,, \quad \lambda \in \mathbb{C} \setminus \{0\}.$$

Similarly, for a function v living on the boundary $\partial\Omega$ with regularity $W_p^{s-1/p}(\partial\Omega)$, where $1 < p < \infty$ and $s \in \{1, 2, \dots, m\}$, we define a norm

$$\|v\|_{s-1/p,p,\partial\Omega,\lambda} = \|v\|_{W_p^{s-1/p}(\partial\Omega)} + |\lambda|^{\frac{s-1/p}{m}} \|v\|_{L^p(\partial\Omega)}, \qquad \lambda \in \mathbb{C} \setminus \{0\}.$$

We quote a well–posedness result from [12], see also [2, Theorem 6.4.1].

Theorem 7.8. *Suppose that the boundary value problem (7.8), (7.9) is elliptic with parameter in the sector $\mathcal{L}$. Then there exists a $\lambda_0 = \lambda_0(p)$ such that for $\lambda \in \mathcal{L}$ with $|\lambda| \geq \lambda_0$, the boundary value problem has a unique solution $(u_1, \dots, u_N) \in \prod_{j=1}^N W_p^{m_j+m}(\Omega)$ for any right-hand side $f = (f_1, \dots, f_N) \in \prod_{j=1}^N W_p^{m_j}(\Omega)$ and all boundary values $g = (g_1, \dots, g_{mN/2}) \in \prod_{j=1}^{mN/2} W_p^{m-r_j-1/p}(\partial\Omega)$, and the* a priori *estimate*

$$\sum_{j=1}^N \|u_j\|_{m_j+m,p,\Omega,\lambda} \leq C \left(\sum_{j=1}^N \|f_j\|_{m_j,p,\Omega,\lambda} + \sum_{j=1}^{mN/2} \|g_j\|_{m-r_j-1/p,p,\partial\Omega,\lambda} \right)$$

holds, where the constant C does not depend upon f, g and λ.

The stationary vQHD system can be written as

$$\begin{cases} \nu_1 \triangle n + \operatorname{div} J = 0, \\ -\dfrac{\varepsilon^2}{4} \nabla \triangle n + \nu_2 \triangle J = -\operatorname{div}\left(\dfrac{J \otimes J + \frac{\varepsilon^2}{4} (\nabla n) \otimes (\nabla n)}{n} \right) - \nabla p(n) + n \nabla V + \dfrac{1}{\tau} J, \\ \triangle V = \dfrac{1}{\lambda_D^2}(n - C). \end{cases}$$

We put $u = \binom{n}{J}$ and define a $(1+d) \times (1+d)$ matrix differential operator $\mathcal{A}$ as in (7.3) with pseudo–differential symbol

$$a(x, \xi) = \begin{pmatrix} -\nu_1 |\xi|^2 & \mathrm{i}\xi_1 & \mathrm{i}\xi_2 & \dots & \mathrm{i}\xi_d \\ \mathrm{i}\frac{\varepsilon^2}{4}\xi_1|\xi|^2 & -\nu_2|\xi|^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathrm{i}\frac{\varepsilon^2}{4}\xi_d|\xi|^2 & 0 & 0 & \dots & -\nu_2|\xi|^2 \end{pmatrix}. \tag{7.11}$$

Now we are in a position to prove Theorem 7.4.

Proof. The operator $\mathcal{A}$ has families of orders $(s_1, s_2, \dots, s_{d+1}) = (1, 2, \dots, 2)$ and $(m_1, m_2, \dots, m_{d+1}) = (1, 0, \dots, 0)$, and we have $N = d + 1$, $m = 2$ as well as $\mathcal{A} = \mathcal{A}_{\mathrm{pr}}$. Then the eigenvalues of a_{pr} are the solutions λ to

$$(\nu_1|\xi|^2 + \lambda)(\nu_2|\xi|^2 + \lambda)^d - (\nu_2|\xi|^2 + \lambda)^{d-1}\frac{\varepsilon^2}{4}|\xi|^4 = 0,$$

hence

$$\lambda_{1,\dots,d-1} = -\nu_2|\xi|^2,$$
$$\lambda_{d,d+1} = -\frac{1}{2}(\nu_1 + \nu_2)|\xi|^2 \pm \frac{1}{2}\sqrt{(\nu_1 - \nu_2)^2 - \varepsilon^2}|\xi|^2.$$

Recalling that the parameters satisfy (7.4), we then can find an angle ϑ (even if $\nu_1 = 0$) with $\pi/2 < \vartheta < \pi$ in such a way that the closure of the sector Σ_ϑ as in (7.5) contains none of the values $\lambda_1, \dots, \lambda_{d+1}$, provided $|\xi| > 0$. This will be a first condition on the choice of the sector $\mathcal{L}$ as in Definition 7.7.

In order to discuss the Shapiro–Lopatinskij condition, we pick a point $x^0 \in \partial\Omega$, and then we rotate and shift the coordinates in such a way that the interior normal direction at x^0 is given by $(0, \dots, 0, 1) \in \mathbb{R}^d$. We consider the boundary value problem on the half–line

$$\begin{cases} (a_{\mathrm{pr}}(\xi', D_{x_d}) - \lambda\, \mathrm{id}_{d+1})v(x_d) = 0, & 0 < x_d < \infty, \\ \lim_{x_d \to +\infty} v(x_d) = 0, & \\ b_{\mathrm{pr},j}(\xi', D_{x_d})v(x_d) = 0, & x_d = 0, \quad j = 1, \dots, d+1, \end{cases} \tag{7.12}$$

where $\xi' \in \mathbb{R}^{d-1}$ and $a_{\text{pr}} = a_{\text{pr}}(\xi', D_{x_d})$ is given as follows:

$$a_{\text{pr}} = \begin{pmatrix} -\nu_1(|\xi'|^2 + D_{x_d}^2) & \mathrm{i}\xi_1 & \dots & \mathrm{i}D_{x_d} \\ \mathrm{i}\frac{\varepsilon^2}{4}\xi_1(|\xi'|^2 + D_{x_d}^2) & -\nu_2(|\xi'|^2 + D_{x_d}^2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{i}\frac{\varepsilon^2}{4}D_{x_d}(|\xi'|^2 + D_{x_d}^2) & 0 & \dots & -\nu_2(|\xi'|^2 + D_{x_d}^2) \end{pmatrix}.$$

Our intention is to show that the choice of boundary condition operators $\mathcal{B}_{\text{pr},j}(\xi', D_{x_d})$ listed in Theorem 7.4 implies $v(x_d) \equiv 0$ for all $\xi' = (\xi_1, \dots, \xi_{d-1})$ and all $\lambda \in \mathcal{L}$, but $|\xi'| + |\lambda| > 0$.

The first line of (7.12) is a system of ODEs with constant coefficients and of mixed order, and it is clear that any decaying solution $v = v(x_d)$ of this system must decay exponentially for $x_d \to \infty$, as well as all derivatives of v. A thorough description of the structure of the solutions to mixed order ODE systems can be found in [1]. We have

$$\begin{aligned} -\nu_1(|\xi'|^2 + D_{x_d}^2)n + \mathrm{i}\xi_1 J_1 + \dots + \mathrm{i}\xi_{d-1}J_{d-1} + \mathrm{i}D_{x_d}J_d &= \lambda n, \\ \mathrm{i}\frac{\varepsilon^2}{4}\xi_k(|\xi'|^2 + D_{x_d}^2)n - \nu_2(|\xi'|^2 + D_{x_d}^2)J_k &= \lambda J_k, \qquad k \le d-1, \\ \mathrm{i}\frac{\varepsilon^2}{4}D_{x_d}(|\xi'|^2 + D_{x_d}^2)n - \nu_2(|\xi'|^2 + D_{x_d}^2)J_d &= \lambda J_d. \end{aligned}$$

Write $\langle \cdot, \cdot \rangle$ for the usual scalar product on $L^2(\mathbb{R}_+)$: $\langle u, v \rangle := \int_0^\infty u\overline{v}\,\mathrm{d}x_d$ and $\|u\|^2 := \langle u, u \rangle$. We take this scalar product of the equations for J_k with J_k and perform appropriate integrations by parts (which produce no boundary terms due to the choice of $\mathcal{B}_n$ and $\mathcal{B}_J$):

$$\begin{aligned} -\frac{\varepsilon^2}{4}\left\langle (|\xi'|^2 + D_{x_d}^2)n, \mathrm{i}\xi_k J_k \right\rangle - \nu_2|\xi'|^2\|J_k\|^2 - \nu_2\|D_{x_d}J_k\|^2 &= \lambda\|J_k\|^2, \\ -\frac{\varepsilon^2}{4}\left\langle (|\xi'|^2 + D_{x_d}^2)n, \mathrm{i}D_{x_d}J_d \right\rangle - \nu_2|\xi'|^2\|J_d\|^2 - \nu_2\|D_{x_d}J_d\|^2 &= \lambda\|J_d\|^2. \end{aligned}$$

Summing up and plugging in the equation for n then give

$$-\frac{\varepsilon^2}{4}\nu_1\left\|(|\xi'|^2 + D_{x_d}^2)n\right\|^2 - \nu_2\sum_{k=1}^{d}\left(|\xi'|^2\|J_k\|^2 + \|D_{x_d}J_k\|^2\right) = \lambda\sum_{k=1}^{d}\|J_k\|^2 + \overline{\lambda}\frac{\varepsilon^2}{4}\left(|\xi'|^2\|n\|^2 + \|D_{x_d}n\|^2\right).$$

The LHS is a non-positive real number, which enforces $n \equiv 0$ and $J \equiv 0$ in the case when $\Re\lambda \ge 0$. There is even a closed sector $\mathcal{L}$, strictly larger than the right complex half-plane, such that $\lambda \in \mathcal{L}$ implies $n \equiv 0$ and $J \equiv 0$. This can be seen as follows. By scaling arguments, we can assume $|\xi'| = 1$, or $\xi' = 0$ and $|\lambda| = 1$. Keep ξ' fixed. The Shapiro–Lopatinskij criterion is violated exactly for those λ, for which the Lopatinskij determinant vanishes. These values of λ form a discrete set in $\mathbb{C}$, which continuously depends on $\xi' \in S^{d-1}$, the unit sphere. But S^{d-1} is compact, which ensures the existence of a sector $\mathcal{L}$ with the desired properties.

This completes the proof of Theorem 7.4. □

Now we give proofs to the Theorems 7.5 and 7.6.

Proof of Theorem 7.5. First we derive a resolvent estimate for $\mathcal{A}$, improving the *a priori* estimates of (7.6). Let $\lambda \in \Sigma_\vartheta$ with $|\lambda| \ge \lambda_0(p)$ as in Theorem 7.4, and consider the problem

$$(\mathcal{A} - \lambda\,\mathrm{id}_{d+1})\begin{pmatrix} n \\ J \end{pmatrix} = \begin{pmatrix} f_0 \\ f' \end{pmatrix} \in X. \tag{7.13}$$

We define an operator $\mathcal{P} = \triangle$ with domain $D(\mathcal{P}) = \{v \in W_p^2(\Omega) \colon (\mathcal{B}_n v)_{|\partial\Omega} = 0\}$, and then we set $n^* := (\mathcal{P} - \lambda\,\mathrm{id})^{-1}f_0$. For this function we have, by classical results, the resolvent estimate

$$|\lambda|\,\|n^*\|_{L^p(\Omega)} + \|n^*\|_{W_p^2(\Omega)} \le C\,\|f_0\|_{L^p(\Omega)}, \quad \text{if } f_0 \in L^p(\Omega), \tag{7.14}$$

which we are now going to "lift" by two Sobolev orders. We choose and fix a complex number μ for which $\mathcal{P} - \mu\,\mathrm{id}$ is an isomorphism from $D(\mathcal{P})$ onto $L^p(\Omega)$. For some constant C_1 (which depends only on μ, but neigher on λ nor on n^*), we then have

$$C_1^{-1}\|n^*\|_{W_p^4(\Omega)} \le \|(\mathcal{P} - \mu\,\mathrm{id})n^*\|_{W_p^2(\Omega)} \le C_1\|n^*\|_{W_p^4(\Omega)}, \qquad \forall\, n^* \in W_p^4(\Omega) \cap D(\mathcal{P}). \tag{7.15}$$

Now we assume $f_0 \in D(\mathcal{P})$. Then also $\mathcal{P}n^* = f_0 + \lambda n^* \in D(\mathcal{P})$ (which implies $\mathcal{B}_n(\mathcal{P}n^*) = 0$), hence $n^* \in D(\mathcal{P}^2)$. This permits us to write

$$(\mathcal{P} - \lambda\,\mathrm{id})(\mathcal{P} - \mu\,\mathrm{id})n^* = (\mathcal{P} - \mu\,\mathrm{id})f_0, \qquad \mathcal{B}_n(\mathcal{P} - \mu\,\mathrm{id})n^* = 0.$$

We apply now (7.14):

$$|\lambda|\,\|(\mathcal{P} - \mu\,\mathrm{id})n^*\|_{L^p(\Omega)} + \|(\mathcal{P} - \mu\,\mathrm{id})n^*\|_{W^2_p(\Omega)} \le C\,\|(\mathcal{P} - \mu\,\mathrm{id})f_0\|_{L^p(\Omega)}\,.$$

And this can be further estimated using (7.15):

$$|\lambda|\,\|n^*\|_{W^2_p(\Omega)} + \|n^*\|_{W^4_p(\Omega)} \le C\,\|f_0\|_{D(\mathcal{P})}\,, \qquad \forall\, f_0 \in D(\mathcal{P}). \tag{7.16}$$

Interpolating between the estimates (7.14) and (7.16) we then find

$$|\lambda|\,\|n^*\|_{W^1_p(\Omega)} + |\lambda|^{1/2}\,\|n^*\|_{W^2_p(\Omega)} + \|n^*\|_{W^3_p(\Omega)} \le C\,\|f_0\|_{W^1_p(\Omega)}\,, \qquad \forall\, f_0 \in D(\mathcal{P}).$$

By density, this estimate holds for all $f_0 \in W^1_p(\Omega)$ with $(\mathcal{B}_n f_0)_{|\partial\Omega} = 0$.

Now we put $n =: n^* + m$ and apply the inequality (7.6) to the problem

$$\mathcal{A}\begin{pmatrix} m \\ J \end{pmatrix} - \lambda\begin{pmatrix} m \\ J \end{pmatrix} = \begin{pmatrix} f_0 - \nu_1 \triangle n^* + \lambda n^* \\ f' + \frac{\varepsilon^2}{4}\nabla\triangle n^* \end{pmatrix},$$

and recalling that $f_0 = \triangle n^* - \lambda n^*$ (by the very definition of n^*) we thusly deduce that

$$\begin{aligned}
&\|m\|_{W^3_p(\Omega)} + |\lambda|^{3/2}\,\|m\|_{L^p(\Omega)} + \|J\|_{W^2_p(\Omega)} + |\lambda|\,\|J\|_{L^p(\Omega)} \\
&\quad \le C\left(\|\triangle n^*\|_{W^1_p(\Omega)} + |\lambda|^{1/2}\,\|\triangle n^*\|_{L^p(\Omega)} + \|f'\|_{L^p(\Omega)} + \|n^*\|_{W^3_p(\Omega)}\right) \\
&\quad \le C\left(\|f_0\|_{W^1_p(\Omega)} + \|f'\|_{L^p(\Omega)}\right).
\end{aligned}$$

Summing up ($n = n^* + m$) and using $|\lambda|\,\|m\|_{W^1_p(\Omega)} \le C(\|m\|_{W^3_p(\Omega)} + |\lambda|^{3/2}\,\|m\|_{L^p(\Omega)})$, we then find

$$\|n\|_{W^3_p(\Omega)} + |\lambda|\,\|n\|_{W^1_p(\Omega)} + \|J\|_{W^2_p(\Omega)} + |\lambda|\,\|J\|_{L^p(\Omega)} \le C\left(\|f_0\|_{W^1_p(\Omega)} + \|f'\|_{L^p(\Omega)}\right), \tag{7.17}$$

which can be expressed, for $|\lambda| \ge \max(1, \lambda_0(p))$, as

$$|\lambda| \cdot \left\|(\mathcal{A} - \lambda\,\mathrm{id}_{d+1})^{-1}\right\|_{\mathcal{L}(X,X)} + \left\|(\mathcal{A} - \lambda\,\mathrm{id}_{d+1})^{-1}\right\|_{\mathcal{L}(X,D(\mathcal{A}))} \le C.$$

Put $\lambda_1 = \lambda_0(p) + 1$. Then we have

$$\sup_{\lambda \in \Sigma_\vartheta} \left\|\lambda(\mathcal{A} - (\lambda + \lambda_1)\,\mathrm{id}_{d+1})^{-1}\right\|_{\mathcal{L}(X,X)} < \infty.$$

Since $D(\mathcal{A})$ is dense in X, the operator $\mathcal{A} - \lambda_1\,\mathrm{id}_{d+1}$ then is a sectorial operator with spectral angle greater than $\pi/2$. Consequently, the operator $\mathcal{A} - \lambda_1\,\mathrm{id}_{d+1}$ (and then also $\mathcal{A}$) generates an analytic semigroup on X. This completes the proof of Theorem 7.5. □

Now we demonstrate Theorem 7.6.

Proof of Theorem 7.6. We write the system as $\partial_t u = \mathcal{A}u + F(u)$, $u(0) = u_0$ with $u = \binom{n}{J}$, $u_0 = \binom{n_0}{J_0}$, $\mathcal{A}$ as in (7.3), and

$$F(u) = \begin{pmatrix} 0 \\ \operatorname{div}\left(\frac{1}{n}\left(J \otimes J + \frac{\varepsilon^2}{4}(\nabla n) \otimes (\nabla n)\right)\right) + \nabla p(n) - n\nabla V - \frac{1}{\tau}J \end{pmatrix}.$$

With $u = u^* + u_0$ we then wish to solve

$$u^*(t) = \int_{s=0}^{t} \exp(\mathcal{A}(t-s))\,(\mathcal{A}u_0 + F(u_0 + u^*)(s))\,\mathrm{d}s, \tag{7.18}$$

by means of the iteration scheme

$$
\begin{aligned}
u_0^*(t) &= 0,\\
u_{k+1}^*(t) &= \int_{s=0}^{t} \exp(\mathcal{A}(t-s))\,(\mathcal{A}u_0 + F(u_0+u_k^*)(s))\,\mathrm{d}s.
\end{aligned}
$$

The analytic semigroup $(\exp(\mathcal{A}t))_{t\geq 0}$ on the space X enjoys the estimate

$$
\|\exp(\mathcal{A}t)v\|_{D(\mathcal{A})} \leq \frac{C(T_0)}{t}\,\|v\|_X\,, \qquad 0 < t \leq T_0, \tag{7.19}
$$

for all $v \in X$. Define by complex interpolation

$$
Y = [D(\mathcal{A}), X]_{1/2} = \left\{(n,J) \in W_p^2(\Omega) \times (W_p^1(\Omega))^d \colon (\mathcal{B}_n n)_{|\partial\Omega} = (\mathcal{B}_J J)_{|\partial\Omega} = 0\right\}.
$$

Then the representation formula of u_{k+1}^* gives us, since $W_p^1(\Omega) \subset L^\infty(\Omega)$ because of $p > d$,

$$
\left\|u_{k+1}^*(t)\right\|_Y \leq Ct^{1/2} \sup_{[0,t]} \|\mathcal{A}u_0 + F(u_0+u_k^*)(s)\|_X \leq Ct^{1/2}\left(1 + \sup_{[0,t]} \|u_k^*(s)\|_Y^3\right),
$$

and the convergence $u_{k+1}^* \to u^*$ in the space $C([0,T_0],Y)$ can be shown by the contraction mapping principle for small T_0. This function $u^* \in C([0,T_0],Y)$ then is a mild solution to the problem

$$
\begin{cases}
\partial_t u^* = \mathcal{A}u^* + \mathcal{A}u_0 + F(u_0+u^*) =: \mathcal{A}u^* + f(t),\\
u^*(0) = 0.
\end{cases}
$$

Now we bring the standard regularity theory into play: since $\mathcal{A}$ is the infinitesimal generator of an analytic semigroup on X, and since $f \in L^q((0,T_0),X)$ for any $q \in (1,\infty)$ (we have even $f \in C([0,T_0],X)$), it follows that

$$
u^* \in C^\theta([0,T_0],X), \qquad \theta = (q-1)/q.
$$

From (7.19) and the representation of u^* we also get $u^* \in L^\infty((0,T_0),[D(\mathcal{A}),X]_\gamma)$ for any $\gamma \in (0,1)$, and interpolating once more we then find

$$
u^* \in C^{1/3}([0,T_0],Y).
$$

This implies $f \in C^{1/3}([0,T_0],X)$. Then, by standard theory, u^* is a classical solution with regularity

$$
\begin{aligned}
&\mathcal{A}u^*,\ \partial_t u^* \in C^{1/3}([\gamma,T_0],X), \qquad \forall\gamma > 0,\\
&\mathcal{A}u^*,\ \partial_t u^* \in C([0,T_0],X).
\end{aligned}
$$

The proof of Theorem 7.6 is complete. □

Appendix A

Facts from Topology and Functional Analysis

We collect some facts for reference, mostly without proofs.

Let X be an arbitrary set (immagine a subset of $\mathbb{R}^n$ or a function space). To avoid nonsense statements, assume $X \neq \emptyset$.

Definition A.1 (**Topology**). *A* topology *is a set τ of subsets of X with the following properties*

- $\emptyset \in \tau$,
- $X \in \tau$,
- *the intersection of finitely many elements of τ is always in τ,*
- *the union of arbitrarily many (even uncountably many) elements of τ is always in τ.*

The pair (X, τ) is called topological space. *If $A \in \tau$ then A is called an* open set, *and the complement $X \setminus A$ is called* closed set.

Definition A.2 (Hausdorff[1] **space**). *A topological space (X, τ) is called a* Hausdorff space *if the following holds: for each distinct $x_1, x_2 \in X$ there are always open sets $A_1, A_2 \in \tau$ with $x_1 \in A_1$, $x_2 \in A_2$ and $A_1 \cap A_2 = \emptyset$.*

Topological spaces which are not Hausdorff turned out to be very unpopular. For this reason, all spaces are now assumed Hausdorff.

Definition A.3 (**Continuous map**). *A map $f\colon X \to Y$ between two zwei topological spaces (X, τ) and (Y, σ) is called* continuous *if the pre-image of each in Y open set is open in X.*

Definition A.4. *A set $N \subset X$ is called* neighbourhood of a point $x_0 \in X$ *if there is an open set U with $x_0 \in U \subset N$.*

We note that neighbourhoods need not be open, and they need not be small.

Definition A.5 (**Basis of a topology**). *Let (X, τ) be a topological space. A family σ of subsets of X is called* basis of the topology τ *if each element of σ is open, and if each open set of X can be written as (possibly uncountable) union of elements of σ.*

A typical example for $X = \mathbb{R}^n$ is this: let ε run through the set $\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$, and take all the open balls with radius ε and with centre point of rational coordinates. This is then a countable basis of the usual metric topology of $\mathbb{R}^n$.

Definition A.6 (**Sub-basis of a topology**). *Let (X, τ) be a topological space. A family σ of subsets of X is called* sub-basis of the topology τ *if the set of finite intersections of members of σ is a basis of the topology τ.*

[1] Felix Hausdorff, 1868–1942

Proposition A.7 (Interior of a set). *Let (X, τ) be a topological space, and let A be a subset of X. Then there is exactly one greatest open set $U \in \tau$ with $U \subset A$. This set U is called* interior of A.

For a proof, it suffices to take all open sets that are contained in A, and then their union.

Proposition A.8 (Closure of a set). *Let (X, τ) be a topological space, and let A be a subset of X. Then there is exactly one smallest closed set U with $A \subset U$. This set U is called* closure of A.

For a proof: just take all closed sets that contain A, and then their intersection.

Definition A.9 (Induced topology). *Let (X, τ) be a topologicaler space, and let $\emptyset \neq A \subset X$. Then all the sets $A \cap U$, where $U \in \tau$, form the* induced topology *of A.*

Attention. Take $X = \mathbb{R}^1$ with the usual metric topology, and choose $A = [-1, 1]$. Then $(0.5, 1]$ is an open set in the induced topology of A, but it is not an open set in X.

We define the convergence of a sequence as expected:

Definition A.10 (Convergence of a sequence). *We say that a sequence $(x_1, x_2, \dots) \subset X$* converges *to an element $x^* \in X$ if for each neighbourhood U of x^* some natural number $N_0 = N_0(U)$ exists such that $x_n \in U$ for each $n \geq N_0$.*

Warning: *For complete metric spaces X, Y, the continuity of a map $f : X \to Y$ is equivalent to the* sequential continuity, *which is defined as $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$, for each converging sequence $(x_1, x_2, \dots) \subset X$. This equivalence does not nold in general topological spaces: sequential continuity does not imply continuity.*

Warning: *In general topological spaces, also the concepts* dense *and* sequentially dicht *do not coincide. We say that a set A is **dense** in a topological space X if the closure of A equals X. And A is called **sequentially dense** in X if each element of X is the limit of a sequence in A.*

Definition A.11 (Compactness). *A subset $A \subset X$ is called* compact *if each covering of A by open sets contains a finite sub-covering.*

Warning: *In complete metric spaces we have the following: a set A is compact if and only if each sequence in A contains a converging sub-sequence with limit in A. This equivalence will **not** hold any longer in general topological spaces. But this equivalence can be restored if the words "sequence" and "sub-sequence" are replaced by "net" and "sub-net" (we make no attempt at defining that).*

Definition A.12 (Product topology). *Let (X, τ) and (Y, σ) be topological spaces. The* produkt topology *of $X \times Y$ is that topology which has the sub-basis consisting of all $U \times V$ where U runs through all open sets in X, and V runs through all open sets in Y.*

Definition A.13 (Topological vector space). *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The standard norm turns $\mathbb{K}$ into a topological space. Let X be a vector space over the field $\mathbb{K}$. We say that X is a* topological vector space (tvs) *if (X, τ) is a topological Hausdorff space, with its topolgy τ designed in such a way that each set with exactly one element is always closed, and that the addition of vectors and the multiplication by scalars are continuous maps from $X \times X$ into X and from $\mathbb{K} \times X$ into X, respectively. Here, the sets $X \times X$ and $\mathbb{K} \times X$ are to be equipped with the product topologies.*

Proposition A.14. *A linear map between two topolical vector spaces is continuous if and only if it is continuous at the origin.*

Definition A.15 (Bounded subset of a tvs). *A subset A of a tvs X is called* bounded *if for each neighbourhood U of the origin, there is a positive λ_0 with $A \subset \lambda U$ for each $\lambda \geq \lambda_0$.*

Intuitively, this means that the set A will be swallowed by each neighbourhood of the origin if you pump up the neighbourhood big enough.

Proposition A.16. *Let X be a vector space over the field $\mathbb{K}$, and let $\{p_i : i \in I\}$ be a family of seminorms (here the index set I may be uncountable). We assume that this family* separates points*: if $p_i(x_0) = 0$ for each $i \in I$, then necessarily $x_0 = 0$.*

To each seminorm p_i and to each $\varepsilon > 0$, we consider sets

$$U_{i,\varepsilon} := \{x \in X : p_i(x) < \varepsilon\}.$$

Then the family

$$x + \{U_{i,\varepsilon} : i \in I,\ \varepsilon > 0\},$$

where x runs through X, is the sub-basis of a topology τ of X.

This topology has the following properties:

- *every open set of X can be written as union of sets $U_{x_0,J,\varepsilon}$ defined as*

  $$U_{x_0,J,\varepsilon} := \{x \in X : p_j(x - x_0) < \varepsilon,\ j \in J\},$$

  where ε runs through $\mathbb{R}_+$, J runs through all finite subsets of I, and x runs through X,
- *a linear set f of X into Y is continuous if and only if for each neighbourhood $V_{0,J,\varepsilon}$ of $0 \in Y$ there is a neighbourhood $U_{0,\tilde{J},\delta}$ of $0 \in X$ with $f(U_{0,\tilde{J},\delta}) \subset V_{0,J,\varepsilon}$,*
- *a set $A \subset X$ is bounded if and only if for each $i \in I$ there is a constant C_i with $p_i(x) \leq C_i$ for each $x \in A$,*
- *a sequencee $(x_1, x_2, \dots) \subset X$ converges to a limit $x^* \in X$ if and only if $\lim_{m\to\infty} p_i(x_m - x^*) = 0$ for each seminorm p_i.*

Such a vector space X, equipped with this topology, is called *locally convex space.*

The situation becomes nicer if there are only countably many seminorms $(p_1, p_2, \dots)$, in which case we can define a function $d\colon X \times X \to \mathbb{R}$ like this:

$$d(x_1, x_2) = \sum_{j=1}^{\infty} 2^{-j} \frac{p_j(x_1 - x_2)}{1 + p_j(x_1 - x_2)}, \qquad x_1, x_2 \in X.$$

We quickly show that d has all the properties of a metric. We say that the tvs X is *metrizable.*

If we now suppose additionally that X is a complete space (which means that each metric Cauchy sequence has a limit in X), then X is called a Fréchet space. The advantage is now that continuity is again equivalent to sequential continuity, density equivalent to sequential density, compactness equivalent to sequential compactness.

We give a hint why $\mathcal{D}(\mathbb{R}^n)$ cannot be a Fréchet space, by means of constructing a similar space (but easier) that is not metrizable.

Lemma A.17. *Choose the (algebraic) space*

$$C^0_{\text{comp}}(\mathbb{R}^n) := \{u\colon \mathbb{R}^n \to \mathbb{R} : u \text{ is continous and has compact support}\}.$$

Consider its sub-spaces

$$C^0_M(\mathbb{R}^n) := \left\{u \in C^0_{\text{comp}}(\mathbb{R}^n) : \operatorname{supp} u \subset B_M(0)\right\}, \qquad M > 0,$$

with $B_M(0)$ being the ball about 0 with radius M.

Then the following holds: any topology of $C^0_{\text{comp}}(\mathbb{R}^n)$ that induces on all its sub-spaces $C^0_M(\mathbb{R}^n)$ their natural Banach space topology, cannot be metrizable.

Proof. The natural Banach space topology of the $C^0_N(\mathbb{R}^n)$ comes from the supremum norm. Pick $0 < M < N$. Then $C^0_M(\mathbb{R}^n)$ is a closed sub-space of $C^0_N(\mathbb{R}^n)$, and $C^0_M(\mathbb{R}^n)$ is *nowhere-dense* in $C^0_N(\mathbb{R}^n)$, because every open ball in $C^0_N(\mathbb{R}^n)$ with arbitrary center and with radius $\varepsilon > 0$ contains functions that do not vanish in the annulus, and these functions are impossible to approximate by functions from $C^0_M(\mathbb{R}^n)$.

We have obviously $C^0_{\text{comp}}(\mathbb{R}^n) = \cup_{M=1}^{\infty} C^0_M(\mathbb{R}^n)$, hence $C^0_{\text{comp}}(\mathbb{R}^n)$ is the countable union of nowhere-dense closed sets.

Now Baires[2] theorem says: a non-empty complete metric space cannot be the countable collection of nowhere-dense closed sets. □

Definition A.18 (Topological dual space). *Let X be a tvs over the field $\mathbb{K}$. The set of all linear and continuous maps from X into $\mathbb{K}$ is denoted by X', and it is called the* topological dual space[3].

[2] René–Louis Baire, 1874–1932

[3] We quickly check that X' is an algebraic vector space over the field $\mathbb{K}$

Each $x \in X$ generates a seminorm p_x on X' via

$$T \mapsto p_x(T) := |\langle T, x\rangle_{X'\times X}|, \qquad T \in X'.$$

The topology of X', generated by these seminorms, is called the weak–∗–topology.

Definition A.19 (**Transposed operator**). *Let X and Y be locally convex spaces, and let $A\colon X \to Y$ be linear and continuous. Then A generates a linear operator operator $A^t\colon Y' \to X'$ via*

$$\left\langle A^t y', x\right\rangle_{X'\times X} := \langle y', Ax\rangle_{Y'\times Y}, \qquad y' \in Y', \quad x \in X,$$

which we call transposed operator.

Proposition A.20. *With the above notations, A^t is continuous. If A is a topological isomorphism (which means that also A^{-1} is continuous, then also A^t is a topological isomorphism, and we have $(A^t)^{-1} = (A^{-1})^t$.*

Proof. For each seminorm p of X', we need a finite collection of seminorms $q_1, \dots, q_k$ of Y' and a constant C, such that

$$p(A^t y') \le C \sum_{j=1}^{k} q_j(y'), \qquad \forall y' \in Y'.$$

Only those p are interesting that are generated by an $x \in X$ via

$$p(x') = |\langle x', x\rangle_{X'\times X}|, \qquad \forall x' \in X',$$

hence we have

$$p(A^t y') = |\left\langle A^t y', x\right\rangle| = |\langle y', Ax\rangle| =: q(y'),$$

where q is a seminorm on Y' that is being generated by $Ax \in Y$. This proves the first claim.

Assume now A as bijective, and A^{-1} continuous. Then there is $(A^{-1})^t\colon X' \to Y'$ defined by

$$\left\langle (A^{-1})^t x', y\right\rangle_{Y'\times Y} := \left\langle x', A^{-1}y\right\rangle_{X'\times X},$$

and it is continuous.

The operator A^t is injective, because: suppose $A^t y' = 0 \in X'$, then we have, for all $x \in X$,

$$0 = \left\langle A^t y', x\right\rangle_{X'\times X} = \langle y', Ax\rangle_{Y'\times Y},$$

hence $0 = \langle y', y\rangle_{Y'\times Y}$ for all $y \in Y$, because A is surjective as a map from X to Y. This means $y' = 0$.

The operator A^t is surjective, because: let us be given $x' \in X'$, and we wish to find $y' \in Y'$ with $A^t y' = x'$. For all $x \in X$ we then have

$$\langle x', x\rangle = \left\langle x', A^{-1}Ax\right\rangle = \left\langle (A^{-1})^t x', Ax\right\rangle = \left\langle A^t (A^{-1})^t x', x\right\rangle,$$

which enforces $x' = A^t(A^{-1})^t x'$. It suffices to choose $y' = (A^{-1})^t x'$. And we also get $(A^t)^{-1} = (A^{-1})^t$. □

Appendix B

Exercises

Microlocal Analysis and Boundary Value Problems — Homework Sheet 1

1. Prove that

$$T(\varphi) = \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x + \int_{\varepsilon}^{\infty} \frac{\varphi(x)}{x}\,\mathrm{d}x\right), \qquad \varphi \in C_0^\infty(\mathbb{R}^1)$$

is a linear map from $C_0^\infty(\mathbb{R}^1)$ into $\mathbb{C}$. Determine whether T is a distribution.

2. Determine a function $E\colon \mathbb{R}^n \to \mathbb{R}$ such that

$$(-1)^n \int_{\mathbb{R}^n} E(x)\partial_{x_1}\dots\partial_{x_n}\varphi(x)\,dx = \varphi(0)$$

holds for all $\varphi \in C_0^\infty(\mathbb{R}^n)$.

3. Let $f_n(x) = \sin(nx)$ for $x \in \mathbb{R}$ and $n \in \mathbb{N}$. The associated regular distributions shall be denoted by T_n, which means $\langle T_n, \varphi\rangle = \int_{x\in\mathbb{R}} f_n(x)\varphi(x)\,\mathrm{d}x$. Check whether

   (a) the sequence of functions $(f_n)_{n\in\mathbb{N}}$ enjoys pointwise convergence almost everwhere,

   (b) the sequence of distributions $(T_n)_{n\in\mathbb{N}}$ converges in $\mathcal{D}'(\mathbb{R})$.

4. Show the continuity of the embedding $\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$.

   Prove that the Fourier transform maps from $\mathcal{E}'(\mathbb{R}^n)$ into $C^\infty(\mathbb{R}^n)$, and it enjoys the formula

$$(\mathcal{F}u)(\xi) = \langle u, \exp(-\mathrm{i}x\cdot\xi)\rangle_{\mathcal{E}'(\mathbb{R}^n_x)\times\mathcal{E}(\mathbb{R}^n_x)}, \qquad \xi \in \mathbb{R}^n, \quad u \in \mathcal{E}'(\mathbb{R}^n_x).$$

   Hint: tensor products of distributions from $\mathcal{S}'$, perhaps.

   Moreover, show that $\mathcal{F}u$ has slow growth, which means

$$|\mathcal{F}u(\xi)| \le C_u(1+|\xi|)^{N_u}, \qquad \forall \xi \in \mathbb{R}^n.$$

   If possible, show that $\mathcal{F}u$ is an analytic function of ξ.

5. Let $\chi = \chi(\xi) \in C_0^\infty(\mathbb{R}^n)$ be a cut-off function which is identically equal to 1 in a neighbourhood of $\xi = 0$. Show that the family of functions $\{\chi_\varepsilon = \chi(\varepsilon\xi)\}_{0<\varepsilon<1}$ belongs to $S^0_{1,0}$ with uniform in ε symbol estimates.

Microlocal Analysis and Boundary Value Problems — Homework Sheet 2

1. To show: if $u \in \mathcal{S}'(\mathbb{R}^n)$ is a distributional solution to $\triangle u = 0$, then u is a polynomial.

   Hint: Fourier transform

   Are there distributional solutions to $\triangle u = 0$, which are no polynomials ?

2. Let $u \in \mathcal{E}'(\mathbb{R}^n)$. Prove that then some $k \in \mathbb{N}_+$ exists and some $f \in L^2(\mathbb{R}^n)$ with $u = (1 - \triangle)^k f$ (as an identity of distributions).

3. Prove that the intersection $\bigcap_{m\in\mathbb{R}} S^m_{\varrho,\delta}(\Omega \times \mathbb{R}^N)$ indeed does not depend on $\varrho, \delta \in [0, 1]$.

   Determine a pseudodifferential symbol $p \in \bigcap_{0<\varepsilon<1/3} S^0_{1-\varepsilon,\varepsilon}(\mathbb{R}^n \times \mathbb{R}^n)$, for which however $p \notin S^0_{1,0}(\mathbb{R}^n \times \mathbb{R}^n)$.

4. Let $a = a(x, \xi) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$, and let $\mathcal{A}$ be the associated ΨDO. Prove that

   $$(\mathcal{A}(x, D_x)u)(x) = \int_{\mathbb{R}^{2n}_{p,q}} \hat{a}(q, p) e^{\mathrm{i}q\cdot X} e^{\mathrm{i}p\cdot D} u(x) \,\mathrm{d}q \,\mathrm{d}p \qquad \forall u \in \mathcal{S}(\mathbb{R}^n),$$

   where the operators $e^{\mathrm{i}q\cdot X}$ (“phase factor”) and $e^{\mathrm{i}p\cdot D}$ (“Taylor expansion”) are defined via

   $$(e^{\mathrm{i}q\cdot X}u)(x) := e^{\mathrm{i}q\cdot x}u(x), \qquad (e^{\mathrm{i}p\cdot D}u)(x) := u(x + p).$$

5. Let $a, b \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$, and let $\mathcal{A}$, $\mathcal{B}$ be the associated ΨDOs. To show: then $\mathcal{C} := \mathcal{A} \circ \mathcal{B}$ is again a ΨDO. Find a representation formula for its symbol $c = c(x, \xi)$ as an oscillating integral.

   Hint: previous question

Microlocal Analysis and Boundary Value Problems — Homework Sheet 3

1. Let $\chi = \chi(\xi) \in C_0^\infty(\mathbb{R}^n)$ be a cut-off function, hence (e.g.) $\chi(\xi) = 1$ for $|\xi| < 1$ and $\chi(\xi) = 0$ for $|\xi| > 2$. For some pseudodifferential symbol $p = p(x, \xi) \in S^m_{\varrho,\delta}$ with global symbol estimates, its operator $\mathcal{P}$, and $\varepsilon > 0$, define $p_\varepsilon(x, \xi) := p(x, \xi)\chi(\varepsilon\xi)$, and let the associated operator be denoted by $\mathcal{P}_\varepsilon(x, D_x)$.

   To show: if $\varrho, \delta \in [0, 1]$ and $u \in \mathcal{S}(\mathbb{R}^n)$, then $\lim_{\varepsilon\to+0} \mathcal{P}_\varepsilon u = \mathcal{P}u$, with convergence in the topology of $\mathcal{S}(\mathbb{R}^n)$.

   To show: if $\varrho, \delta \in [0, 1]$, where $\delta < 1$, and if $u \in \mathcal{S}'(\mathbb{R}^n)$, then $\lim_{\varepsilon\to+0} \mathcal{P}_\varepsilon u = \mathcal{P}u$, with convergence in the weak-$*$ topology of $\mathcal{S}'(\mathbb{R}^n)$.

2. Let $a(x, \xi) \sim \sum_{k=0}^\infty a_k(x, \xi)$ be an asymptotically convergent series with $a \in S^m_{1,0}(\mathbb{R}^n_x \times \mathbb{R}^n_\xi)$ and $a_k \in S^{m-k}_{1,0}(\mathbb{R}^n_x \times \mathbb{R}^n_\xi)$. Equip the statement "each asymptotically convergent series can be differentiated termwise" with a meaning and with a proof.

3. Let $a \in S^m_{1,0}(\mathbb{R}^n_x \times \mathbb{R}^n_\xi)$ with global symbol estimates, and let $\mathcal{A}$ be the associated ΨDO. To show that $\mathcal{A}$ maps $\mathcal{S}(\mathbb{R}^n)$ continuously into itself, and $\mathcal{S}'(\mathbb{R}^n)$ continuously into itself as well.

4. Consider in $\mathbb{R}^2$ the function

$$u(x_1, x_2) = \begin{cases} 0 & : \ x_1 < 0, \\ 1 & : \ x_1 \geq 0. \end{cases}$$

   For each $x_0 \in \mathbb{R}^2$, determine those $s \in \mathbb{R}$ for which $u \in H^s_{x_0}$.

   For each $(x_0, \xi_0) \in \mathbb{R}^2 \times (\mathbb{R}^2 \setminus 0)$, determine those $s \in \mathbb{R}$ for which $u \in H^s_{x_0,\xi_0}$.

Bibliography

[1] Shmuel Agmon, Avron Douglis, and Louis Nirenberg. Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. II. *Commun. Pure Appl. Math.*, 17:35–92, 1964.

[2] Mikhail Semenovich Agranovich. *Partial Differential Equations IX*, volume 79 of *Encyclopaedia of Mathematical Sciences*, chapter Elliptic boundary problems, pages 1–145. Springer, 1997.

[3] Mikhail Semenovich Agranovich and Mark Iosifovich Vishik. Elliptic problems with a parameter and parabolic systems of general form. *Russ. Math. Surv.*, 19:53–157, 1964.

[4] Hans Wilhelm Alt. *Lineare Funktionalanalysis*. Springer, 2002.

[5] Herbert Amann. *Linear and quasilinear parabolic problems. Vol. I*, volume 89 of *Monographs in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1995. Abstract linear theory.

[6] Louis Boutet de Monvel. Boundary problems for pseudo-differential operators. *Acta Math.*, 126(1-2):11–51, 1971.

[7] Klemens Burg, Herbert Haf, and Friedrich Wille. *Höhere Mathematik für Ingenieure. Band 3. Gewöhnliche Differentialgleichungen, Distributionen, Integraltransformationen. 3., durchges. Aufl.* Stuttgart: Teubner, 1993.

[8] Alberto P. Calderon and Remi Vaillancourt. On the boundedness of pseudo–differential operators. *J. Math. Soc. Japan*, 23(2):374–378, 1971.

[9] Li Chen and Michael Dreher. Viscous quantum hydrodynamics and parameter-elliptic systems. *Math. Methods Appl. Sci.*, 34(5):520–531, 2011.

[10] Arthur Erdélyi, Wilhelm Magnus, Fritz Oberhettinger, and Francesco Tricomi. *Higher Transcendental Functions. Bateman Manuscript Project*. McGraw–Hill, New York, 1953.

[11] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[12] Melvin Faierman. Eigenvalue asymptotics for a boundary problem involving an elliptic system. *Math. Nachr.*, 279(11):1159–1184, 2006.

[13] Gerd Grubb. *Distributions and Operators*, volume 252 of *Graduate Texts in Mathematics*. Springer, 2009.

[14] Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 1-4. Springer, 1983.

[15] Hans Jarchow. *Locally Convex Spaces*. Mathematische Leitfäden. B. G. Teubner, 1981.

[16] Jürgen Jost. *Partielle Differentialgleichungen*. Springer, 1998.

[17] Hitoshi Kumano-go. *Pseudo-Differential Operators*. MIT Press, Cambridge, 1982.

[18] Jeffrey Rauch. *Partial Differential Equations*. Springer, 1991.

[19] Walter Rudin. A failure of the associative law. *Delta (Waukesha)*, 5:28–31, 1975.

[20] Elmar Schrohe. A short introduction to Boutet de Monvel's calculus. In *Approaches to singular analysis (Berlin, 1999)*, volume 125 of *Oper. Theory Adv. Appl.*, pages 85–116. Birkhäuser, Basel, 2001.

[21] Andreas Seeger, Christopher Sogge, and Elias Stein. Regularity properties of Fourier integral operators. *Annals of Mathematics*, 134(2):231–251, 1991.

[22] Mikhail Shubin. *Pseudodifferential Operators and Spectral Theory.* Springer, 2001.

[23] Michael Eugene Taylor. *Pseudodifferential Operators.* Princeton University Press, 1981.

[24] Michael Eugene Taylor. *Pseudodifferential Operators and Nonlinear PDE.* Birkhäuser Boston, 1991.

[25] Francois Treves. *Basic Linear Partial Differential Equations.* Academic Press, 1975.

[26] Hans Triebel. *Höhere Analysis.* Thun, Frankfurt am Main: Verlag Harri Deutsch, 1980.

[27] Vasilij Sergeevich Vladimirov. *Gleichungen der mathematischen Physik.* Deutscher Verlag der Wissenschaften, 1972.

[28] Wolfgang Walter. *Einführung in die Theorie der Distributionen.* B.I. Wissenschaftsverlag, 1994.

[29] Wolfgang Walter. *Ordinary Differential Equations.* Graduate Texts in Mathematics. 182. Springer, 1998.

[30] Joseph Wloka. *Partielle Differentialgleichungen.* B. G. Teubner, Stuttgart, 1982. Sobolevräume und Randwertaufgaben. [Sobolev spaces and boundary value problems], Mathematische Leitfäden. [Mathematical Textbooks].