Alma Mater Studiorum · Università di Bologna

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica

On solvability of PDEs

Tesi di Laurea in Analisi Matematica

Relatore:
Chiar.mo Prof.
Alberto Parmeggiani

Presentata da:
Luca Pallucchini

II Sessione
Anno Accademico 2011/2012

alla mia famiglia e a Giulia …

Introduction

The most natural question one can ask about PDEs is whether there exists a solution of it. For ODEs we have satisfactory theorems about the existence of solutions (at least locally). Malgrange and Ehrenpreis have proved that all constant coefficient linear partial differential equations have local solutions, and that, by Cauchy-Kovelevsky Theorem, all analytic partial differential equation have local analytic solutions. Therefore, it came as a complete surprise when in 1957 Hans Lewy discovered the first non-solvable operator,

$$L = D_{x_1} + iD_{x_2} - 2i(x_1 + ix_2)D_{x_3}.$$

Note that L is the tangential Cauchy-Riemann operator on the boundary of the strictly pseudoconvex domain

$$\{(z, w) \in \mathbb{C}^2; |z|^2 - \mathsf{Im}(w) < 0\}. \tag{1}$$

In fact, consider $\mathsf{Im}(w) = |z|^2$ and the vector field $\alpha\partial/\partial\bar{z} + \beta\partial/\partial\bar{w}$,

$$\Big(\alpha\frac{\partial}{\partial\bar{z}} + \beta\frac{\partial}{\partial\bar{w}}\Big)\Big(\underbrace{|z|^2}_{=z\bar{z}} - \underbrace{\mathsf{Im}(w)}_{=(w-\bar{w})/(2i)}\Big) = 0$$

gives $\beta = -2iz\alpha$. Now consider the map

$$\Phi : (x, y, u) \longmapsto (x, y, u, x^2 + y^2),$$

where $z = x + iy$ and $w = u + iv$. We get :

$$\Phi_*(\partial_x) = \partial_x + 2x\partial_v, \quad \Phi_*(\partial_y) = \partial_y + 2y\partial_v, \quad \Phi_*(\partial_u) = \partial_u$$

and

$$\Phi_*(D_x + iD_y - 2izD_u) = D_x + iD_y - 2iz(D_u + iD_v).$$

The discovery opened up a new research area, that of solvability of partial differential operators, with the aim of understanding necessary and/or sufficient conditions for solvability.
Solvable differential operators include operators with constant coefficients and elliptic operators.
In fact by Malgrange-Ehrenpreis theorem [Chapter 3] every constant coefficient linear partial differential equation has a fundamental solution E, i.e. there exists a distribution E s.t. $P(D)E = \delta$ and so there exists a solution of the equation $P(D)u = f$ with $f \in \mathscr{E}'$, $u = E * f$ is a solution of it (for the general case one can see [21],[10] about P-convexity, but we are not going to treat these topics).
For an elliptic operator there exists a *parametrix* and this gives solvability [Chapter 4].
The simplest class of non-elliptic solvable operators is that of operator of *real principal type* [8].

The conditions for solvability, as we will see, lie in the geometry of the operator. Elaborating on the Lewy operator, Hörmander [8] found the first general necessary condition for solvability, (H)[chapter 2]. One can see that for the Lewy operator (H) is violated. This was a remarkable advance because it explained a phenomenon, that had appeared as an isolated example, in terms of very general geometric properties of the principal symbol, an invariantly defined object.
In spite of this success, it turned out that condition (H) was not accurate enough to discriminate the solvable operator from the non-solvable ones. In fact, if one considers the Mizohata operator,

$$M_k = \frac{\partial}{\partial y} - iy^k \frac{\partial}{\partial x},$$

(H) is satisfied but M_k is locally solvable at the origin iff k is even.
Nirenberg and Treves [22] identified a property that turned out to be the

right condition for local solvability of *differential* operator of principal type (condition $\mathscr{P}$), see [Chapter 2]. The sufficiency of condition ($\mathscr{P}$) was proved by Nirenberg and Treves [23] for analytic differential operators and by Beals and Fefferman [4] for *pseudodifferential operators* of principal type.
A related and more general condition, condition (Ψ), also introduce by Nirenberg and Treves [22], relevant for solvability of *pseudodifferential operators*, was shown to be necessary in dimension 2 by Moyer (unpublished) and in several dimensions by Hörmander [12]. The sufficiency of the condition was next shown by Lerner [17] in 2 dimensions and finally in the general case by Dencker [5].
Condition (Ψ) is more general then condition ($\mathscr{P}$) for *pseudodifferential operators*, whereas it coincides with ($\mathscr{P}$) for *differential operators.*

This dissertation aims at being a rapid introduction (although, far to be complete) to solvability of PDEs. In this work, I have followed several papers quoted in the Bibliography and some notes by G.Mendoza and A.Parmeggiani. It is divided into six chapters:

- The first chapter is a review [6] of notions of symplectic geometry that will be used throughout.

- The second chapter introduce the conditions for solvability, with examples.

- The third chapter gives two proofs of the Malgrange-Ehrenpreis theorem about the existence of the fundamental solution of PDEs with constant coefficients. In this chapter one can find two different way to prove the theorem: the first one is the description of Atiyah's proof about the division of distributions, the second one is Hörmander's proof.
  In this case, as is well known, the solvability of $P(D)u = f \in \mathscr{E}'(\mathbb{R}^n)$ is solved. In the last part of the chapter we give an elementary proof due to D. Jerison [14], of the L^2 local solvability, in which use is made of the SAK principle by C. Fefferman and D. H. Phong (see [7]).

- In the fourth chapter we describe the process of constructing a parametrix for an elliptic differential operator.
- In the fifth chapter we describe the link between solvability and hypoellipticity.
- In the sixth chapter we give Hörmander's proof of the invariance of condition (Ψ) and an elementary proof of the sufficiency of condition (Ψ) in two dimensions due to H. Smith [30]. (Originally existence in 2 dimensions due to N. Lerner [17]).

Introduzione in italiano

La domanda più naturale che uno potrebbe porsi riguardo alle PDEs è se esista una soluzione. Per le ODEs abbiamo soddisfacenti teoremi riguardo l'esistenza di soluzioni (almeno localmente). Malgrange ed Ehrenpreis hanno provato che tutte le equazioni differenziali alle derivate parziali con coefficenti costanti hanno soluzioni locali e, grazie al teorema di Cauchy-Kovelevsky, tutte le equazioni differenziali alle derivate parziali analitiche hanno soluzioni analitiche. Per questo fu una grande sorpresa quando nel 1957 Hans Lewy scoprì il primo operatore non risolubile,

$$L = D_{x_1} + iD_{x_2} - 2i(x_1 + ix_2)D_{x_3}.$$

Notiamo che L è l'operatore di Cauchy-Riemann tangenziale sul bordo del dominio strettamente pseudoconvesso

$$\{(z, w) \in \mathbb{C}^2; |z|^2 - \mathsf{Im}(w) < 0\}.$$

La scoperta aprì una nuova area di ricerca riguardante la risolubilità di equazioni differenziali alle derivate parziali, con lo scopo di capire necessarie e/o sufficienti condizioni per la risolubilità.
Tra gli operatori differenziali alle derivate parziali risolubili troviamo quelli a coefficenti costanti e gli operatori ellittici.

Per i primi, infatti, grazie al teorema di Malgrange-Ehrenpreis [Chapter 3], sappiamo esistere una soluzione fondamentale, cioè una distribuzione E per cui $P(D)E = \delta$, da cui poi possiamo trovare una soluzione di $P(D)u = f$ convolvendo E con f, questo a patto di prendere $f \in \mathscr{E}'(\mathbb{R}^n)$ (per il caso generale si rimanda ai lavori [21],[10] riguardo la P-convessità, questi non verranno trattati nella tesi).
Riguardo agli operatori ellittici sappiamo esistere una parametrice e questa dà risolubilità [Chapter 4]. La classe più semplici di operatori non ellittici risolubili è quella degli operatori di *tipo principale* [8].
Le condizioni per la risolubilità, come vedremo, riguardano la geometria dell'operatore.
La prima condizione per la risolubilità è stata trovata da Hörmander [8], la condizione (H) [Chapter 2]. Si vede che l'operatore L di Lewy non soddisfa questa condizione. Questo fu un risultanto importante in quanto ci spiega un fenomeno apparso come uno esempio isolato, in termini di proprietà geometriche del simbolo principale.
Nonostante questo successo, la condizione (H) si scoprì non abbastanza accurata per discriminare gli operatori risolubili da quelli non risolubili, infatti l'operatore di Mizohata

$$M_k = \frac{\partial}{\partial y} - iy^k \frac{\partial}{\partial x},$$

soddisfa (H) ma è localmente risolubile nell'origine se e solo se k è pari.
Nirenberg e Treves [22] identificarono una proprietà che si rivelò essere la condizione sufficiente per la risolubilità di operatori differenziali di *tipo principale*, la condizione $(\mathscr{P})$ [Chapter 2].
La sufficienza di $(\mathscr{P})$ fu dimostrata da Nirenberg e Treves [23] per operatori differenziali analitici e da Beals e Fefferman [4] per *operatori pseudodifferenziali di tipo principale.*
Una condizione più generale di $(\mathscr{P})$, la condizione (Ψ), anch'essa introdotta da Nirenberg e Treves [22], fu mostrata essere necessaria in due dimensioni da Moyer (non pubblicata) e in più dimensioni da Hörmander [12]. La sufficienza fu poi provata da Lerner [17] in due dimensioni e da Dencker [5] in

generale per più dimensioni.
La condizione (Ψ) è più generale della condizione $(\mathscr{P})$ per *operatori pseudodifferenziali*, anche se esse coincidono per *operatori differenziali*.
Questa tesi vuole essere una rapida introduzione alla risolubilità di PDEs. In questo lavoro ho fatto affidamento su vari articoli citati in Bibliografia e su appunti di G. Mendoza e A. Parmeggiani. Per un piccolo riassunto riguardo i temi svolti nei successivi capitoli si veda l'introduzione in inglese.

Contents

Chapter 1

Local symplectic geometry

We recall briefly some notions such as: manifold, tangent and cotagent vectors, differential form and vector bundle.

1.1 Tangent and cotangent vectors

Let X be a smooth manifold of dimension n.
Let $x_0 \in X$.

If $\gamma, \tilde{\gamma} : (-1, 1) \to X$, we say that $\gamma, \tilde{\gamma}$ are equivalent if $\|\gamma(t) - \tilde{\gamma}(t)\| = o(t), t \to 0$, (this is well defined through local coordinates).

A tangent vector is by definition an equivalence class. If γ is a curve as above we denote by $\dot{\gamma}(0)$ or $(\frac{d}{dt})_{t=0}\gamma(t)$ the corresponding tangent vector. The set of tangent vectors at a point $x_0 \in X$ is denoted by $T_{x_0}X$.

If $f, \tilde{f} : X \to \mathbb{R}$ are two C^1-functions, we say that f and $\tilde{f}$ are equivalent if $\big(f(x) - f(x_0)\big) - \big(\tilde{f}(x) - \tilde{f}(x_0)\big) = o(\|x - x_0\|), x \to x_0$.
We let $df(x_0)$ denote the equivalence class of f. It is called a differential 1-form at x_0 or a cotangent vector at $x_0 \in X$; also it is the differential of f at x_0.
The sets $T^*_{x_0}X$ and $T_{x_0}X$ are n-dimensional (real) vector spaces dual to each other, the duality being given by

$$\langle df(x_0), \dot{\gamma}(0)\rangle = \Big(\frac{d}{dt}\Big)_{t=0} f(\gamma(t)). \tag{1.1}$$

If $x_1, ..., x_n$ are local coordinates, defined in a neighborhood of x_0, then $dx_1(x_0), ..., dx_n(x_0)$ or $(dx_1, ...dx_n$ for short) is a basis in $T^*_{x_0}X$ and the corresdonding dual basis is (by definition) $\frac{\partial}{\partial x_1}(x_0), ..., \frac{\partial}{\partial x_n}(x_0)$ (or $\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n}$ for short).

The sets $TX = \bigsqcup_{x_0 \in X} T_{x_0}X$ (disjoint union) and $T^*X = \bigsqcup_{x_0 \in X} T^*_{x_0}X$ are vector bundles and in particular C^∞-manifolds. If $x_1, ..., x_n$ are local coordinates on X then we get corresponding local coordinates $(x, t) = (x_1, ..., x_n, t_1, ..., t_n)$ on TX and $(x, \xi) = (x_1, ..., x_n, \xi_1, ..., \xi_n)$ on T^*X by representing $\nu \in TX$ and $\rho \in T^*X$ by their base point x so that $\nu \in T_xX$, $\rho \in T^*_xX$ and then writing $\nu = \sum t_j \frac{\partial}{\partial x_j}(x)$, $\rho = \sum \xi_j dx_j$. The local coordinates $(x_1, ..., x_n, \xi_1, ..., \xi_n)$ are called canonical (local) coordinates on T^*X. If $y_1, ..., y_n$ is a second systems of local coordinates, then in the intersection of the two open sets in X parametrized by the two systems of coordinates we have the relations $t = \frac{\partial x}{\partial y}s$, $\eta = {}^t\big(\frac{\partial x}{\partial y}\big)\xi$ for the corresponding local coordinates (x, t), (y, η) on T^*X. Here $\frac{\partial x}{\partial y} = \big(\frac{\partial x_j}{\partial y_k}\big)$ is the standard Jacobian matrix.

If $\rho \in T^*X$, we let $\pi(\rho) \in X$ be the corresponding base point. A section in T^*X is a map $\omega : X \to T^*X$ with $\pi \circ \omega(x) = x$, $\forall x \in X$. (The same definition can be given for TX or for any given vector bundle). Sections in T^*X are called differential 1-form, and sections in TX are called vector fields. A vector field can be written in local coordinates as $\nu = \sum_1^n t_j(x)\frac{\partial}{\partial x_j}$ and a differential 1-form as $\omega = \sum_1^n \xi_j(x)dx_j$.

If Y is a second manifold and $f : Y \to X$ is a map of class C^1, $y_0 \in Y$, $x_0 = f(y_0) \in X$, then we have a natural map $f_* = df : T_{y_0}Y \to T_{x_0}X$ which in local coordinates is given by the ordinary Jacobian matrix. The adjoint is $f^* : T^*_{x_0}X \to T^*_{y_0}Y$ and we notice that $d(u \circ f)(y_0) = f^*(du(x_0))$ if u is a C^1 function on X. If Z is a third manifold, $g : Z \to Y$ in C^1 and $z_0 \in Z$, $g(z_0) = y_0$, then $(f \circ g)_* = f_* \circ g_*$, $(f \circ g)^* = g^* \circ f^*$. When passing to sections we see that if ω is a 1-form on Y then $f^*\omega$ is a well defined 1-form on Y (this the pull-back of ω by means of f). Notice that the corresponding push-forward $f_*\nu$ of a vector field ν on Y can be defined if f is a C^1 diffeo-

morphism but not in general.
If $\gamma : (a,b) \to Y$ is a C^1 curve and $t_0 \in (a,b)$ we defined its tangent at $\gamma(t_0)$ as $\gamma_*\left(\frac{\partial}{\partial t}\right) := \frac{\partial \gamma}{\partial t}(t_0) = \dot{\gamma}(t_0)$. (This definition coincides with the earlier one.) If ν is a C^∞ vector field on X then for every $x_0 \in X$ we can find $T_+(x_0), T_-(x_0) \in (0,\infty]$ such that we have a unique smooth curve

$$(-T_-(x_0), T_+(x_0)) \ni t \longmapsto \gamma(t) = \exp(t\nu)(x_0) \in X \tag{1.2}$$

with $\gamma(0) = x_0$, $\dot{\gamma}(t) = \nu(\gamma(t))$.
Choosing $T_+(x_0), T_-(x_0)$ maximal, we get a smooth map

$$\begin{aligned} \Phi : \{(t,x) \in \mathbb{R} \times X; -T_-(x) < t < T_+(x)\} \to X \\ \Phi(t,x) = \exp(t\nu)(x), \end{aligned} \tag{1.3}$$

where $\Phi(0,x) = x$, $\partial\Phi(t,x)/\partial t = \nu(\Phi(t,x))$ and $T_+(x), T_-(x)$ are lower semi-continuous. We have

$$\exp(t\nu) \circ \exp(s\nu)(x) = \exp((t+s)\nu)(x), \tag{1.4}$$

for t,s such that both sides are defined. For details see [6], [26]

1.2 The canonical 1- and 2- forms

Let $\pi : T^*X \to X$ be the natural projection. For $\rho \in T^*X$ we consider $\pi^* : T^*_{\pi(\rho)}X \to T^*_\rho(T^*X)$ and since $\rho \in T^*X$ we can define the *canonical 1-form* $\omega_\rho \in T^*_\rho(T^*X)$ by $\omega_\rho = \pi^*(\rho)$. Varying ρ we get a smooth 1-form on T^*X. In canonical coordinates we get $\omega = \sum_1^n \xi_j dx_j$.

We next recall a few facts about forms of higher degree. If L is a finite-dimensional real vector space and L^* is the dual space, then we have a natural duality between the k-fold exterior product spaces $\bigwedge^k L$ and $\bigwedge^k L^*$, given by

$$\langle u_1 \wedge ... \wedge u_k, v_1 \wedge ... \wedge v_k \rangle = \det(\langle u_j, v_k \rangle), \quad u_j \in L, v_k \in L^*. \tag{1.5}$$

If M is a C^∞ manifold of dimension m then a differential k-form is a section v of the vector bundle $\bigwedge^k T^*M$. In local coordinates $x_1, ..., x_m$,

$$v = \sum_{|I|=k} v_I(x) dx^I, \tag{1.6}$$

where in general $I = (i_1, ..., i_l) \in \{1, ..., m\}^l, |I| = l, dx^I = dx_{i_1} \wedge ... \wedge dx_{i_l}$. (The representation (1.6) becomes unique if we restrict to those I's with $i_1 < i_2 < ... < i_k$.) If v is a k-form of class C^1 locally given by (1.6), we define the $(k+1)$-form

$$dv = \sum_{|I|=k} dv_I \wedge dx^I \quad \text{(the exterior differential of } v). \tag{1.7}$$

This definition does not depend on the choice of local coordinates or on how we choose the representation (1.6). We have the following facts:

(i) Twice the exterior differential is 0, that is

$$d^2 = 0. \tag{1.8}$$

(ii) If ω is a C^∞ $(k+1)$-form which is closed in the sense that $d\omega = 0$, then in every open set in M diffeomorphic to a ball, we can find a smooth k-form v such that $dv = \omega$. (Poincaré's lemma)

(iii) If $f : Y \to X$ is a smooth map between two smooth manifolds then there is a unique way of extending the pull-back f^* of 1-forms to k-forms by multilinearity. Moreover if v is a smooth k-form on X, then $d(f^*v) = f^*(dv)$.

We now return to the canonical 1-form ω on T^*X and define the *canonical 2-form* on T^*X as $\sigma = d\omega$. In canonical coordinates,

$$\sigma = \sum_1^n d\xi_j \wedge dx_j. \tag{1.9}$$

This 2-form is also called *the canonical symplectic form.*
For $\rho \in T^*X$, σ_ρ can be viewed as a linear form on $\bigwedge^2 T_\rho(T^*X)$ or equivalently as an skewsymmetric bilinear form on $T_\rho(T^*X) \times T_\rho(T^*X)$ given by

$$\sigma_\rho(t, s) = \langle \sigma_\rho, t \wedge s \rangle, \quad t, s \in T_\rho(T^*X). \tag{1.10}$$

In canonical coordinates we write $t = (t_x, t_\xi)\big(\text{where } t = \sum t_{x_j} \frac{\partial}{\partial x_j} + t_{\xi_j} \frac{\partial}{\partial \xi_j}\big)$, $s = (s_x, s_\xi)$ and get

$$\sigma_\rho(t, s) = \langle t_\xi, s_x \rangle - \langle s_\xi, t_x \rangle = \sum \big(t_{\xi_j} s_{x_j} - s_{\xi_j} t_{x_j}\big). \tag{1.11}$$

From this it is clear that σ_ρ is a non-degenerate bilinear form and we therefore have a bijection $H : T^*_\rho(T^*X) \to T_\rho(T^*X)$ defined by

$$\sigma(s, Hu) = \langle s, u\rangle, \quad s \in T_\rho(T^*X), \quad u \in T^*_\rho(T^*X). \tag{1.12}$$

In canonical coordinates, if $u = u_x dx + u_\xi d\xi = \sum(u_{x_j} dx_j + u_{\xi_j} d\xi_j)$ we get $Hu = u_\xi \frac{\partial}{\partial x} - u_x \frac{\partial}{\partial \xi}$.

If $f(x, \xi)$ is of class C^1 on (some open set in) T^*X, we define its Hamilton (vector) field by $H_f = H(df)$. In canonical coordinates,

$$H_f = \sum_1^n \Big(\frac{\partial f}{\partial \xi_j}\frac{\partial}{\partial xi} - \frac{\partial f}{\partial x_j}\frac{\partial}{\partial \xi_j}\Big). \tag{1.13}$$

In a more sophisticated way, let M be a manifold, $\rho \in M$, $t \in T_\rho M$ and define $t\lrcorner : \bigwedge^k T^*_\rho M \to \bigwedge^{k-1} T^*_\rho M$ as the adjoint of the left exterior multiplication $t\wedge : \bigwedge^{k-1} T_\rho M \to \bigwedge^k T_\rho M$. Then with $M = T^*X$, the Hamilton field is defined by the pointwise relation

$$H_f \lrcorner \sigma = -df. \tag{1.14}$$

Eqivalently, $\sigma(v, H_f) = df(v)\ \forall v \in T_\rho T^*X$. If f, g are two C^1 functions defined on the same open set in T^*X, we define their *Poisson bracket* as the continuous function on T^*X given by

$$\{f, g\} = H_f(g) = \langle H_f, dg\rangle = \sigma(H_f, H_g), \tag{1.15}$$

where in the second expression we view H_f as a first-order differential operator. In canonical coordinates,

$$\{f, g\} = \sum \Big(\frac{\partial f}{\partial \xi_j}\frac{\partial g}{\partial x_j} - \frac{\partial f}{\partial x_j}\frac{\partial g}{\partial \xi_j}\Big). \tag{1.16}$$

Notice that $\{f, g\} = -\{g, f\}$ and, in particular, $\{f, f\} = 0$.

1.3 Lie derivatives

Let v be a C^∞ vector field on a manifold M and let ω be a $C^\infty k$-form on M. Then the Lie derivative of ω along v is poitwise defined by

$$\mathcal{L}_v\omega = \Big(\frac{d}{dt}\Big)_{t=0}((\exp tv)^*\omega). \tag{1.17}$$

If ω is a second smooth vector field on M we also define

$$\mathcal{L}_v u = \big(\frac{d}{dt}\big)_{t=0}((\exp -tv)_* u). \tag{1.18}$$

In the latter definition we observe that in this case the push-forward of a vector field is made through a local diffeomorphism. We have the following identities :

(i) When ω is a 0-form and hence a function, then $\mathcal{L}_v\omega = v(\omega)$

(ii) $\mathcal{L}_v u = [v, u] = vu - uv$, where u, v are viewed as first-order differential operators in the last two expressions.

(iii) $\mathcal{L}_v(d\omega) = d(\mathcal{L}_v\omega)$

(iv) $\mathcal{L}_v(\omega_1 \wedge \omega_2) = (\mathcal{L}_v\omega_1) \wedge \omega_2 + \omega_1 \wedge (\mathcal{L}_v\omega_2)$

(v) $\mathcal{L}_v(u \lrcorner \omega) = (\mathcal{L}_v u) \lrcorner \omega + u \lrcorner (\mathcal{L}_v\omega)$

(vi) $\mathcal{L}_v\omega = v \lrcorner d\omega + d(v \lrcorner \omega)$

(vii) $\mathcal{L}_{v_1+v_2} = \mathcal{L}_{v_1} + \mathcal{L}_{v_2}$

Lemma 1.3.1. *If f is a C^∞ function on some open set in T^*X, then $\mathcal{L}_{H_f}\sigma =$ 0*

Proof. One computes

$$\mathcal{L}_{H_f}\sigma = H_f \lrcorner d\sigma + d(H_f \lrcorner \sigma) = H_f \lrcorner d^2\omega - d^2 f = 0. \tag{1.19}$$

□

The maps $T^*X \ni \rho \mapsto \Phi_t(\rho) = \exp(tH_f)(\rho)$ is a local diffeomorphism when $|t|$ is sufficiently small. They are also local symplectimorphisms, that is $\Phi_t^*\sigma = \sigma$. This immediately follows from

$$\frac{d}{dt}\Phi_t^*\sigma = \big(\frac{d}{ds}\big)_{s=0}\Phi_t^*\Phi_s^*\sigma = \Phi_t^*\mathcal{L}_{H_f}\sigma = 0 \tag{1.20}$$

1.4 Lagrangian manifolds

A submanifold $\Lambda \subset T^*X$ is called a *Lagrangian manifold* if $\dim \Lambda = \dim X$ and $\sigma|_\Lambda = 0$. In general we define the restriction of a differential k-form to a submanifold as the pull-back of this form through the natural inclusion map. Viewing $(\sigma|_\Lambda)_\rho$, $\rho \in \Lambda$, as a bilinear form on $T_\rho\Lambda \times T_\rho\Lambda$ we simply have $(\sigma|_\Lambda)_\rho(t,s) = \sigma_\rho(t,s)$, $t,s \in T_\rho\Lambda$, where $T_\rho\Lambda$ is identified with a subspace of $T_\rho T^*X$ (namely the image of $T_\rho\Lambda$ by the differential of the natural inclusion map). If $T_\rho\Lambda^\sigma$ denotes the orthogonal space of $T_\rho\Lambda$ in $T_\rho T^*X$ with respect to the bilinear form σ_ρ, then we see that a submanifold $\Lambda \subset T^*X$ is Lagrangian if and only if $T_\rho\Lambda^\sigma = T_\rho\Lambda$ for every $\rho \in \Lambda$.

Theorem 1.4.1. *Let $\Lambda \subset T^*X$ be a submanifold with $\dim \Lambda = \dim X$ and such that $\pi|_\Lambda : \Lambda \to X$ is a local diffeomorphism (in the sense that every point $\rho \in \Lambda$ has a neighborhood in Λ which is mapped diffeomorphically by $\pi|_\Lambda$ onto a neighborhood of $\pi(\rho)$). Then Λ is Lagrangian iff for each point ρ of Λ we can find a (real) C^∞ function $\varphi(x)$ defined near the projection of ρ, such that Λ coincides near ρ with the manifold $\big\{(x, d\varphi(x)); x \in$ some neighborhood of $\pi(\rho)\big\}$.*

Proof. If ω is the canonical 1-form, we notice that $d(\omega|_\Lambda) = \sigma|_\Lambda$. Therefore the following three statements are equivalent:

(i) Λ is Lagrangian;

(ii) $\omega|_\Lambda$ is closed (i.e. $d(\omega|_\Lambda) = 0$);

(iii) locally on Λ we can find a smooth function φ with $\omega|_\Lambda = d\varphi$.

That (iii)$\Rightarrow$(i) is clear. We show (ii)$\Rightarrow$(iii). If $x_1, ..., x_n$ are local coordinates on X, we can also view them (or rather their compositions with π) as local coordinates on Λ, and represent Λ by $\xi = \xi(x)$ in the corresponding canonical coordinates. Then (iii) is equivalent to $\xi_j(x) = \partial\varphi(x)/\partial x_j$ i.e. $\sum \xi_j(x)dx_j = d\varphi$ $\square$

1.5 Hamilton-Jacobi equations

Hamilton-Jacobi equations are equations of the form $p(x, \varphi_x^{'}) = 0$, where p is a real-valued C^∞ function defined on some open subset of T^*X. Here we shall also assume that $dp \neq 0$ when $p = 0$. The basic idea is to construct a Lagrangian manifold Λ associated with φ, and to construct it inside the hypersurface $\Sigma = p^{-1}(0)$. If $\rho \in \Lambda$, we shall then have $T_\rho\Lambda \subset T_\rho\Sigma$ (considering these tangent spaces as subspaces of $T_\rho T^*X$), and hence $T_\rho\Sigma^\sigma \subset T_\rho\Lambda$ (since $T_\rho\Lambda^\sigma = T_\rho\Lambda$). Now $T_\rho\Sigma^\sigma = \mathbb{R}H_p$ so we must have $H_p \in T_\rho\Lambda$ at every point $\rho \in \Lambda$, or in other words, that H_p must be tangent to Λ at every point of Λ.

Proposition 1.5.1. *Let $\Lambda' \subset \Sigma$ be an isotropic submanifold (in the sense that $\sigma|_{\Lambda'} = 0$) of dimension $n-1$ passing through some given point $\rho_0 \in \Sigma$ and such that $H_p(\rho_0) \notin T_{\rho_0}\Lambda'$. Then in a neighborhood of ρ_0, we can find a Lagrangian manifold Λ such that $\Lambda' \subset \Lambda \subset \Sigma$ (in that neighborhood).*

Proof. According to the observation above it is natural to consider

$$\Lambda = \left\{ \exp(tH_p)(\rho); |t| < \epsilon, \rho \in \Lambda', |\rho - \rho_0| < \epsilon \right\} \tag{1.21}$$

for some sufficiently small $\epsilon > 0$. (Here $|\rho - \rho_0|$ is well defined if we choose some local canonical coordinates.) Then $\Lambda' \subset \Lambda$ (near ρ_0) and since H_p is tangent to Σ (by the relation $H_p p = 0$) we also have $\Lambda \subset \Sigma$. From the assumption $H_p(\rho_0) \notin T_{\rho_0}\Lambda'$ and the implicit function theorem, it also follows that Λ is a smooth manifold of dimension n. In order to verify that Λ is Lagrangian, we first take $\rho \in \Lambda'$ (with $|\rho - \rho_0| < \epsilon$) and consider $T_\rho\Lambda = T_\rho\Lambda \oplus \mathbb{R}H_p$. Then $\sigma_\rho|_{T_\rho\Lambda \times T_\rho\Lambda} = 0$ since $\sigma_\rho|_{T_\rho\Lambda' \times T_\rho\Lambda'} = 0$, $\sigma_\rho(H_p, H_p) = 0$, $\sigma_\rho(t, H_p) = 0$, $t \in T_\rho\Lambda'$ (this follows from $\sigma_\rho(t, H_p) = \langle t, dp\rangle = 0$, for all $t \in T_\rho H$). More generally, at the point $\rho_t := \exp(tH_p)(\rho), \rho \in \Lambda'$, we have $T_{\rho_t}\Lambda = \exp(tH_p)_*(T_\rho\Lambda)$ and for $u, v \in T_\rho\Lambda$ we get (using $\exp(tH_p)^*\sigma_{\rho_t} = \sigma_\rho$)

$$\sigma_{\rho_t}(\exp(tH_p)_*u, \exp(tH_p)_*v) = \sigma_\rho(u, v) = 0. \tag{1.22}$$

This conclues the proof in view of Thm. 1.4.1 $\square$

In what follows we write $x = (x', x_n) \in \mathbb{R}^n, x' = (x_1, ..., x_{n-1}) \in \mathbb{R}^{n-1}$. The following thm. gives a more local solution to the HJ eqt. once an initial value is fixed.

Theorem 1.5.2. *Let $p(x, \xi)$ be a real-valued C^∞ function, defined in a neighborhood of some point $(0, \xi_0) \in T^*\mathbb{R}^n$, such that $p(0, \xi_0) = 0$, $\frac{\partial p}{\partial \xi_n}(0, \xi_0) \neq 0$. Let $\psi(x')$ be a real-valued C^∞ function defined near 0 in $\mathbb{R}^{n-1}$ such that $\frac{\partial \psi}{\partial x'}(0) = \xi_0'$. Then there exists a real-valued smooth function $\varphi(x)$, defined in a neighborhood of $0 \in \mathbb{R}^n$, such that in that neighborhood*

$$p(x, \varphi_x'(x)) = 0, \quad \varphi(x', 0) = \psi(x'), \quad \varphi_x'(0) = \xi_0. \tag{1.23}$$

If $\tilde{\varphi}(x)$ is a second function with the same properties, then $\varphi(x) = \tilde{\varphi}(x)$ in some neighborhood of 0.

Proof. In a suitable neighborhood of $(0, \xi_0) \in \mathbb{R}^{n-1} \times \mathbb{R}^n$ we have $p(x', 0, \xi_0') = 0$ if and only if $\xi_n = \lambda(x', \xi')$, where λ is a real-valued C^∞ function, with $\lambda(0, \xi_0') = (\xi_0)_n$. Let

$$\Lambda' = \left\{(x, \xi); x_n = 0, \xi' = \frac{\partial \psi}{\partial x'}(x'), \xi_n = \lambda(x', \xi'), x' \in \text{neigh}(0)\right\} \tag{1.24}$$

(where "$x' \in \text{neigh}(0)$" means that x' belongs to some sufficiently small neighborhood of 0). Then $\Lambda \subset p^{-1}(0)$ is isotropic of dimension $n-1$ and H_p is nowhere tangent to Λ' since H_p has a component $\frac{\partial p}{\partial \xi_n} \frac{\partial}{\partial x_n}$ with $\frac{\partial p}{\partial \xi_n} \neq 0$. Let $\Lambda \subset p^{-1}(0)$ be a Lagrangian manifold as in Proposition 1.5.1. The differential of $\pi|_\Lambda : \Lambda \to \mathbb{R}^n$ is bijective at $(0, \xi_0)$ so if we restrict our attention to a sufficiently small neighborhood of $(0, \xi_0)$, Λ becomes of the form $\xi = \varphi'(x)$, $x \in \text{neigh}(0)$. Hence $p(x, \varphi_x'(x)) = 0$, $\varphi_x'(0) = \xi_0$. Since $\Lambda' \subset \Lambda$ we get $\frac{\partial \psi}{\partial x'}(x') = \frac{\partial \varphi}{\partial x'}(x', 0)$, and modifying φ by a constant gives $\varphi(x', 0) = \psi(x')$. For the uniqueness see [6, 26]. $\square$

We can view Λ as a union of integral curves of H_p passing through Λ'. The projection of such an integral curve is an integral curve of the field $\nu = \sum_1^n \frac{\partial p}{\partial \xi_j}(x, \varphi_x') \frac{\partial}{\partial x_j}$ (which, via $\pi|_\Lambda$, can be identified with $H_p|_\Lambda$). If $q(x, \xi) =$

$\sum_1^n \frac{\partial p}{\partial \xi_j}(x,\xi)\xi_j$, we have

$$(\sum_1^n \frac{\partial p}{\partial \xi_j}(x,\varphi'_x)\frac{\partial}{\partial x_j})\varphi = q(x,\varphi'_x). \tag{1.25}$$

Hence, if $x = x(t)$ is an integral curve of ν with $x_n(0) = 0$, then we get $\varphi(x(t)) = \psi(x'(0)) + \int_0^t q(x(s),\xi(s))ds$, where $\xi(s) = \varphi'(x(s))$, so that $s \mapsto (x(s),\xi(s))$ is the integral curve of H_p, with $x_n(0) = 0$, $\xi'(0) = \frac{\partial \psi}{\partial x'}(x'(0)), \xi_n(0) = \lambda(x'(0),\xi'(0))$. In particular, if p is positively homogeneous of degree $m > 0$, then by the Euler homogeneity relations, $q(x,\xi) = mp(x,\xi) = 0$ on Λ and we obtain $\varphi(x(t)) = \psi(x'(0))$.
If $\varphi = \varphi_\alpha$ depends smoothly on some parameter $\alpha \in \mathbb{R}$, then $\varphi = \varphi(x,\alpha)$ will be a smooth function of (x,α), and differentiating the equation $p(x,\varphi'_x) = 0$ with respect to α we get

$$\sum \frac{\partial p}{\partial \xi_j}(x,\varphi'_x)\frac{\partial}{\partial x_j}\frac{\partial \varphi}{\partial \alpha} = 0 \tag{1.26}$$

so that $\frac{\partial \varphi}{\partial \alpha}$ is constant along the bichacteristic curves (without any homogeneity assumption).

Recall that a characteristic curve is the x-space projection of a bicharacteristic curve, the latter being by definition an integral curve of H_p.

Chapter 2

Introduction to Solvability

2.1 The problem

This is an exposition of various results concerning the existence of solutions of pseudodifferential operators. In its most elementary form, the problem is the following. Let U be an open subset of $\mathbb{R}^n$, and let

$$P(x, D_x) = \sum_{|\alpha|<m} a_\alpha(x) D_x^\alpha, \quad x \in U, \tag{2.1}$$

$$0 \neq \sum_{|\alpha|=m} |a_\alpha(x)|^2 \quad \forall x \in U \tag{2.2}$$

Local solvability of P at $x_0 \in U$ is stated as follows.

Definition 2.1.1. For every $f \in C_0^\infty$ there exists a distribution u defined in U such that $Pu = f$ near x_0.

equazioni differenziali alle derivate parziali The distribution u is not required to be smooth. The validity of the assertion is referred to as the (local) solvability of P at x_0, and P is said to be (locally) solvable at x_0 if that holds. Solvable differential operators include operators with constant coefficients (by the Malgrange-Ehrenpreis Theorem [Chapter 3]) and elliptic operators [Chapter 4], that is, operators such as P in (2.1) for which the principal symbol, i.e. the function

$$p_m(x,\xi) = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha, \quad (x,\xi) \in U \times \mathbb{R}^n, \tag{2.3}$$

has the property that $p_m(x,\xi) \neq 0$ if $\xi \neq 0$.

The simplest class of non-elliptic solvable operators is the class of operators of *real principal type* (Hörmander [8]).
Principal type means that the differential forms

$$\sum \xi_j dx_j \quad \text{and} \quad \sum \Big(\frac{\partial p_m}{\partial x_j} dx_j + \frac{\partial p_m}{\partial \xi_j} d\xi_j\Big) \tag{2.4}$$

are *linearly independent* (over $\mathbb{C}$) at every point of

$$p^{-1}(0) = \mathrm{Char}(P) = \{(x,\xi) \in U \times \mathbb{R}^n : \xi \neq 0,\, p_m(x,\xi) = 0\ \}, \tag{2.5}$$

and *real principal type* means that, in addition, $p_m(x,\xi)$ is *real-valued.*

The first non-solvable operator,

$$L = D_{x_1} + D_{x_2} - 2i(x_1 + ix_2)D_{x_3} \quad (x_1, x_2, x_3) \in \mathbb{R}^3 \tag{2.6}$$

in $\mathbb{R}^3$, was discovered by Hans Lewy [20] in 1957. Elaborating on this example Hörmander [8] (see also [9]) found the first general necessary condition for solvability. Investigating further, Nirenberg and Treves [22] gave a weaker necessary condition for solvability of an operator of real principal type at every point of an open set. In this paper we find for the first time condition $(\mathscr{P})$ stated explicitly. The condition was shown to be sufficient for analytic differential operators by Nirenberg and Treves [23] and in full generality for pseudodifferential operators of principal type, by Beals and Fefferman [4].
A related condition, Condition (Ψ), also introduced by Nirenberg and Treves in [22], relevant for solvability of pseudodifferential operators, was shown to be necessary in dimension 2 by Moyer (unpublished) and in any dimension by Hörmander [12]. The sufficiency of the condition for solvability was proved by Lerner [17] in dimension 2, and by Dencker [5] in general case.

In order to state Hörmander's condition in [8] as well as conditions ($\mathscr{P}$) and (Ψ) we introduce some notation.
Suppose p is a smooth *complex-valued* function defined in an open subset of $\mathbb{R}^n \times \mathbb{R}^n$. The Hamiltonian vector field of p is (recall)

$$H_p = \sum_{j=1}^{n} \Big(\frac{\partial p}{\partial \xi_j} \frac{\partial}{\partial x_j} - \frac{\partial p}{\partial x_j} \frac{\partial}{\partial \xi_j} \Big). \tag{2.7}$$

If q is another such function, then the Poisson bracket of p and q is (recall)

$$\{p, q\} = H_p q. \tag{2.8}$$

Suppose $p : U \times (\mathbb{R}^n \backslash 0) \to \mathbb{C}$ is smooth and positively homogeneous of degree $m > 0$ in ξ. Recall that the latter means that $p(x, \lambda\xi) = \lambda^m p(x, \xi)$ for every $(x, \xi) \in U \times (\mathbb{R}^n \backslash 0)$ and $\lambda > 0$. Let

$$p^{-1}(0) = \mathrm{Char}(P) = \{(x, \xi) \in U \times \mathbb{R}^n : \xi \neq 0, \, p(x, \xi) = 0 \}, \tag{2.9}$$

Recall that p is elliptic if it vanishes nowhere. If p is real-valued, then the integral curves of its Hamiltonian vector field, the curves

$$\begin{gathered} \mathbb{R} \supset I \ni t \mapsto \chi(t; x, \xi) \in U \times \mathbb{R}^n \\ \dot{\chi}(t; x, \xi) = H_p(\chi(t; x, \xi)), \quad \chi(0; x, \xi) = (x, \xi) \end{gathered} \tag{2.10}$$

are well-defined and have the property that $p(\chi(t; x, \xi)) = p(x, \xi)$ (because $H_p p = 0$). An integral curve with a point in $\mathrm{Char}(P)$ (hence entirely contained in $\mathrm{Char}(P)$) is a null-bicharacteristic of p.
Let p be an arbitrary smooth complex-valued positively homogeneous function such that $dp \neq 0$ on $\mathrm{Char}(P)$. Hörmander's condition [8] is

(H) The Poisson bracket $\{p, \bar{p}\}$ vanishes at every point of $\mathrm{Char}(P)$

while Conditions (Ψ) and ($\mathscr{P}$) are, respectively,

(Ψ) For every elliptic homogeneous function q, the function $\mathsf{Im}(qp)$ does not change sign from - to + along any given oriented maximal integral curve of $H_{\mathsf{Re}(qp)}$ in $U \times \mathbb{R}^n$ passing through $\mathrm{Char}(P)$.

and

> ($\mathscr{P}$) For every elliptic homogeneous function q, the function $\mathsf{Im}(qp)$ does not change sign along any given maximal integral curve of $H_{\mathsf{Re}(qp)}$ in $U \times \mathbb{R}^n$ passing through $\mathrm{Char}(P)$.

Condition (Ψ) allows for $\mathsf{Im}(qp)$ to change sign from + to -. It also allows for $\mathsf{Im}(qp)$ to be negative at some point of a null-bicharateristic of $\mathsf{Re}(qp)$, then zero in an interval, and then again negative, as well as zero infinitely many times. Condition ($\mathscr{P}$) does not allow changes of sign at all. Returning to Lewy's example (2.6), for which the principal symbol is

$$p = \xi_1 + i\xi_2 - 2i(x_1 + ix_2)\xi_3,$$

we have

$$H_p = \frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2} - 2i(x_1 + ix_2)\frac{\partial}{\partial x_3} + 2i\xi_3\Big(\frac{\partial}{\partial \xi_1} + i\frac{\partial}{\partial \xi_2}\Big),$$

so

$$\{p, \bar{p}\} = \frac{\partial \bar{p}}{\partial x_1} + i\frac{\partial \bar{p}}{\partial x_2} - 2i(x_1 + ix_2)\frac{\partial \bar{p}}{\partial x_3} + 2i\xi_3\Big(\frac{\partial \bar{p}}{\partial \xi_1} + i\frac{\partial \bar{p}}{\partial \xi_2}\Big) = 8i\xi_3.$$

This vanishes if $\xi_3 = 0$. However,

$$\begin{aligned}\mathrm{Char}(P) &= \{(x_1, x_2, x_3; \xi_1, \xi_2, \xi_3) \in \mathbb{R}^3 \times (\mathbb{R}^3\backslash\{0\}) : \xi_1 = -2x_2\xi_3, \xi_2 = 2x_1\xi_3\} \\ &= \{(x_1, x_2, x_3; -2x_2\xi_3, 2x_1\xi_3, \xi_3) : \xi_3 \neq 0\},\end{aligned}$$

so $\{p, \bar{p}\}$ does not vanish on $\mathrm{Char}(P)$. Thus Hörmander's condition is violated.

Continuing with Lewy's example, the Hamiltonian of $\mathsf{Re}\, p = \xi_1 + 2x_2\xi_3$ (we are taking $q = 1$ here) is

$$H_{\mathsf{Re}\, p} = \frac{\partial}{\partial x_1} + 2x_2\frac{\partial}{\partial x_3} - 2\xi_3\frac{\partial}{\partial \xi_2}.$$

Its integral curves are

$$\chi(t; x_0, \xi_0) = (t + x_1^0, x_2^0, 2x_2^0 t + x_3^0; \xi_1^0, -2\xi_3^0 t + \xi_2^0, \xi_3^0),$$

and if $(x^0, \xi^0) \in \mathrm{Char}(P)$ then

$$\chi(t; x_0, \xi_0) = (t + x_1^0, x_2^0, 2x_2^0 t + x_3^0; -2x_2^0\xi_3^0, 2(x_1^0 - t)\xi_3^0, \xi_3^0).$$

Write $\gamma(t)$ for this curve. Evaluating $\mathsf{Im}\, p = \xi_2 - 2x_1\xi_3$ at $\gamma(t)$ we get

$$2(-t + x_1^0)\xi_3^0 - 2(t + x_1^0)\xi_3^0 = -4\xi_3^0 t.$$

So if $\xi_3 < 0$ then $\mathsf{Im}\, p$ changes sign from - to +, as t grows, at $t = 0$ along $\gamma(t)$. Thus Lewy's operator does not satisfy (Ψ) and neither does it satisfy $(\mathscr{P})$.
Note that

$$\{p, \bar{p}\} = -2i\{\mathsf{Re}p, \mathsf{Im}p\},$$

so the three conditions are related.
EXAMPLE 2.11. A simpler example of a non solvable operator is the Mizohata operator,

$$M_1 = D_{x_1} + ix_1 D_{x_2},$$

in $\mathbb{R}^2$. One may verify that $(\mathscr{P})$ is not satisfied. More generally,

$$M_k = D_{x_1} + ix_1^k D_{x_2},$$

does not satisfy the condition if k is odd.

2.2 An example of a proof of necessity

Generally speaking, it easier to find necessary conditions for solvability than sufficient conditions. The scheme for proving that a certain condition is necessary is to contradict an estimate, frequently referred to as Hörmander's estimate, which is equivalent to solvability. We will state and prove the estimate as a consequence of solvability, and then apply it to show, by way of contradiction that the Mizohata and Lewy operators are not solvable.

Lemma 2.2.1 ([11]). *Suppose P is differential operator defined in an open set U in $\mathbb{R}^n$ with the property that for every $f \in C_0^\infty(U)$ there is $u \in \mathcal{D}'(U)$*

such that $Pu = f$. Then, for any given $V \subset\subset U$ there are constants C, M and N such that

$$\Big|\int_U fv dx\Big| \leq C \sup_{|\alpha|\leqslant M, x\in U} |D_x^\alpha f| \sup_{|\beta|\leqslant N, x\in U} |D_x^\beta\, {}^tPv| \quad \text{for all}\, f, v \in C_0^\infty(V).$$

Proof. Let $X = C_0^\infty(\bar{V})$ with its standard topology and let $Y = C_0^\infty(V)$ with the topology determined by the seminorms

$$v \mapsto \sup_{|\beta|\leqslant N, x\in U} |D_x^\beta\, {}^tPv|, \tag{2.11}$$

for each N. The estimate is then seen to be equivalent to the continuity of the bilinear form

$$\begin{aligned} &B : C_0^\infty(\bar{V}) \times C_0^\infty(V) \to \mathbb{C} \\ &B(f, v) = \int_U fv dx. \end{aligned}$$

Note that for each N (2.11) is actually a norm, so Y is a metric space. To verify this, the only thing we need to check is that ${}^tPv = 0$ implies $v = 0$. So suppose ${}^tPv = 0$. Let $f \in C_0^\infty(U)$ be arbitrary. Since $Pu = f$ for some $u \in \mathcal{D}'(U)$,

$$\int fv dx = \langle Pu, v\rangle = \langle u, {}^tPv\rangle = 0.$$

So $v = 0$. We now show that B is separately continuous. Fix $v \in C_0^\infty(V)$. Then

$$|B(f, v)| \leqslant C \sup |v| \sup |f|, \tag{2.12}$$

so $X \ni f \to B(f, v) \in \mathbb{C}$ is continuous.
Next, fix $f \in C_0^\infty(\bar{V})$. There is $u \in \mathcal{D}'(U)$ such that $Pu = f$. Then

$$|B(f, v)| = |\langle Pu, v\rangle| = |\langle u, {}^tPv\rangle| \leqslant C \sup_{|\beta|\leqslant M, x\in U} |D_x^\beta\, {}^tPv|, \tag{2.13}$$

so by Theorem A.2.2 in Appendix A the map $Y \ni v \mapsto B(f, v)$ is continuous.

□

Let M be the Mizohata operator M_1. Suppose M is solvable near 0. Fix a neighborhood V of 0 and let N and M be the numbers in Hörmander's estimate. The general scheme is to find f_τ and $v_\tau \in C_0^\infty(V)$ with τ large such that

$$|\int f_\tau v_\tau dx|$$

is bounded from below by a positive number as $\tau \to \infty$ but

$$\sup_{|\alpha|\leqslant N, x\in U} |D_x^\alpha f_\tau| \sup_{|\beta|\leqslant N, x\in U} |D_x^\beta \, {}^tM v_\tau| \to 0 \quad \text{as} \quad \tau \to \infty.$$

In general, one does not have much control on

$$\sup_{|\alpha|\leqslant N, x\in U} |D_x^\alpha f_\tau|, \tag{2.14}$$

except that it is polynomially bounded. The function f_τ basically serves only as a localizer. The burden falls on finding v_τ, such that

$$\sup_{|\beta|\leqslant N, x\in U} |D_x^\beta \, {}^tM v_\tau|$$

decreases fast enough so as to compensate for the increase of the other factor. In general the family v_τ is of the form $v_\tau = e^{i\tau\phi} \sum_{j=0}^k v_j \tau^{-j}$ with ϕ and $v_j \in C_0^\infty(V)$, $j = 0, 1, ..., k$. (For the Mizohata operator and later the Lewy operator we will only need v_0). This will be achieved by first arranging for ϕ to have the property that $\mathsf{Im}\,\phi$ is strictly positive in a punctured neighborhood of 0. In finding ϕ we will take advantage of the fact that M has analytic coefficients. Note that ${}^tM$ is just $-M$. Below we write M rather than ${}^tM$ for this reason.

To find an equation for ϕ we apply M to $v_\tau = e^{i\tau\phi} \sum_{j=0}^N v_j \tau^{-j}$ and organize by powers of τ:

$$\begin{aligned} M(e^{i\tau\phi} \sum_{j=0}^k v_j \tau^{-j}) &= e^{i\tau\phi}(i\tau M\phi \sum_{j=0}^k v_j \tau^{-j} + \sum_{j=0}^k M v_j \tau^{-j}) \\ &= e^{i\tau\phi}(i\tau v_0 M\phi + \sum_{j=0}^{k-1}(i v_{j+1} M\phi + M v_j)\tau^{-j} + M v_k \tau^{-k}). \end{aligned}$$

In order for this to be at the very least bounded as $\tau \to \infty$ for arbitrary choices of v_j we need $M\phi = 0$. We focus on this equation for a while. We will find a solution of

$$M\phi = 0, \quad \phi|_{x_1=0} = \eta x_2 + i x_2^2.$$

This will ensure good bounds for the absolute value of $e^{i\tau\phi}$ at $\{x_1 = 0\}$ by the failure of condition $(\mathscr{P})$ once we make a choice for the real constant η. We will make a specific choice later on. To solve this Cauchy problem we use the complex version of the Hamilton-Jacobi method. We will go in some details through the various steps in the construction of the solution. The graph of the gradient of the initial condition is

$$\gamma_0 = \{y; \eta + 2iy\},$$

The initial strip is the subset of $\mathrm{Char}(P) = \{\xi_1 + i x_1 \xi_2 = 0\}$ consisting of points $(0, x_2; \xi_1, \xi_2)$ such that $(x_2, \xi_2) \in \gamma_0$, that is,

$$\Gamma_0 = \{(0, x_2; 0, \eta + 2ix_2)\}.$$

The Hamiltonian vector field of $p = \xi_1 + i x_1 \xi_2$ is

$$H_p = \frac{\partial}{\partial x_1} + i x_1 \frac{\partial}{\partial x_2} - i\xi_2 \frac{\partial}{\partial \xi_1}.$$

The integral curves are the solutions of

$$\dot{x_1} = 1, \dot{x_2} = i x_1, \dot{\xi_1} - i\xi_2, \dot{\xi_2} = 0.$$

The integral curve starting at the point $(0, y; 0, \eta + 2iy)$ is $\chi(t, y) = (X(t, y); \Xi(t, y))$ with

$$X(t, x) = (t, y + it^2/2), \quad \Xi(t, x) = (-i(\eta + 2iy)t, \eta + 2iy).$$

The equation $X(t, y) = (x_1, x_2)$ gives (t, y) in terms of (x_1, x_2), as expected:

$$t = x_1, \quad y = x_2 - i x_1^2/2.$$

From this we get the gradient of ϕ at (x_1, x_2) by using in $\Xi(t, y)$:

$$\frac{\partial \phi}{\partial x_1}(x_1, x_2) = -i(\eta + 2i(x_2 - ix_1^2/2))x_1$$
$$\frac{\partial \phi}{\partial x_2}(x_1, x_2) = \eta + 2i(x_2 - ix_1^2/2)$$

that is,

$$\frac{\partial \phi}{\partial x_1}(x_1, x_2) = -i\eta x_1 + 2x_1x_2 - ix_1^3$$
$$\frac{\partial \phi}{\partial x_2}(x_1, x_2) = \eta + 2ix_2 + x_1^2.$$

From this we get

$$\phi(x) = x_1^2x_2 + \eta x_2 + i(-\eta x_1^2/2 - x_1^4 + x_2^2), \quad \text{and} \quad M\phi = 0.$$

Choosing $\eta < 0$ we get that

$$\mathsf{Im}\, \phi \geq c(x_1^2 + x_2^2),$$

with some $c > 0$ in a disc D in V centered at 0 (of radius depending on η). Let $v_0 \in C_0^\infty(D)$, $v_0(x) = 1$ if $|x| \leq r$. Then

$$M(e^{i\tau\phi}v_0) = e^{i\tau\phi}Mv_0, \quad |x| < r.$$

Since $v_0(x) = 1$ if $|x| \leq r$,

$$|e^{i\tau\phi}Mv_0| = e^{-\mathsf{Im}\tau\phi}|Mv_0| \geq \sup|Mv_0|e^{-cr\tau},$$

which gives (recall that ${}^tM = -M$) $\|{}^tM(e^{i\tau\phi}v_0)\| \leq \text{Area}(D)^{1/2}\sup|Mv_0|e^{-cr\tau}$. Let $f \in C_0^\infty(V)$ have Fourier transform $\hat{f}$ such that $\hat{f}(0, -\eta) \neq 0$. Define $f_\tau(x) = \tau^2 f(\tau x)$. For a large τ we have $f_\tau \in C_0^\infty(V)$, so

$$\Big|\int_U f_\tau(x)v_\tau(x)dx\Big| \leq C \sup_{x\in V, |\alpha|\leq N} |D_x^\alpha f_\tau(x)|\|v_\tau\|. \tag{2.15}$$

But

$$\int_U f_\tau(x)v_\tau(x)dx = \int_U f(x)v_0(x/\tau)e^{i(x_1^2x_2/\tau^2 + \eta x_2 + i(-\eta x_1^2/2\tau - x_1^4\tau^3 + x_2^2/\tau))}dx.$$

Whose limits $\tau \to +\infty$ is

$$\int_U f(x) v_0(0) e^{i\eta x_2} dx = \hat{f}(0, -\eta),$$

which is not 0 by our choice of f. So the left hand side of (2.15) is uniformly bounded from below by a positive constant for all large τ. On the other hand,

$$D_x^\alpha f_\tau(x) = \tau^{|\alpha|+2} (D_x^\alpha f)(\tau x).$$

Therefore the right-hand side of (2.15) is bounded by

$$C \mathrm{Area}(D) \sup |M v_0| \sup_{x \in V, |\alpha| \le N} |(D_x^\alpha f(x)| \tau^{N+2} e^{-cr\tau},$$

which tends to 0 as τ tends to infinity. Thus M cannot be solvable. In the analysis just completed we chose $\eta < 0$ and $\hat{f}(0, -\eta) \neq 0$. The Hamiltonian vector field of $\mathsf{Re}\, p = \xi_1$ is $\partial / \partial x_1$ (in $\mathbb{R}_x^2 \times \mathbb{R}_\xi^2$). The integral curve of this vector field passing through $(0, 0; 0, -\eta)$, a point in $\mathrm{Char}(M)$, at time 0 is $t \to \gamma(t) = (t, 0; 0, -\eta)$, whence

$$\mathsf{Im}\, p(\gamma(t)) = -t\eta.$$

Thus $\mathsf{Im}\, p$ changes sign from $-$ to $+$ along γ at $t = 0$.

We will now use the same scheme to prove that the Lewy operator L is not solvable. We will assume that L is solvable near 0 (it is in fact non-solvable at any point, see below) and contradict Hörmander's inequality. We will again take advantage of the fact that L has analytic coefficients and look first for a function ϕ such that

$${}^tL\phi = 0, \quad \phi|_{x_1=0} = \eta x_3 + i(x_2^2 + x_3^2).$$

We find a solution using once more the holomorphic version of the Hamilton-Jacobi method. The principal symbol of ${}^tL$ is

$$p = -\xi_1 - i\xi_2 + 2i(x_1 + 2x_2)\xi_3.$$

The graph of the gradient of the initial condition is

$$\{(y_2, y_3; 2iy_2, \eta + 2iy_3)\},$$

so the initial strip is the subset

$$\{(0, y_2, y_3; \xi_1, 2iy_2, \eta + 2iy_3) : \xi_1 = 2y_2(1 - \eta - 2iy_3)\},$$

of Char(P). We already saw that

$$H_p = \frac{\partial}{\partial x_1} + i\frac{\partial}{\partial x_2} - 2i(x_1 + ix_2)\frac{\partial}{\partial x_3} + 2i\xi_3\Big(\frac{\partial}{\partial \xi_1} + i\frac{\partial}{\partial \xi_2}\Big).$$

The integral curves of the Hamiltonian vector field of p are the solutions of

$$\begin{aligned} \dot{x_1} &= 1 & \dot{\xi_1} &= 2i\xi_3 \\ \dot{x_2} &= i & \dot{\xi_2} &= -2\xi_3 \\ \dot{x_3} &= -2i(x_1 + ix_2) & \dot{\xi_3} &= 0. \end{aligned}$$

The integral curve $(X(t, y_2, y_3), \Xi(t, y_2, y_3))$ passing through

$$(0, y_2, y_3; 2y_2(1 - \eta - 2iy_3), 2iy_2, \eta + 2iy_3)$$

at time 0 is given by

$$\begin{aligned} x_1 &= t & \xi_1 &= 2i(\eta + 2iy_3)t + 2y_2(1 - \eta - 2iy_3) \\ x_2 &= it + y_2 & \xi_2 &= -2(\eta + 2iy_3)t + 2iy_2 \\ x_3 &= 2y_2t + y_3 & \xi_3 &= \eta + 2iy_3. \end{aligned}$$

The condition $X(t, y_2, y_3) = (x_1, x_2, x_3)$ gives

$$t = x_1, \quad y_2 = i(x_1 + ix_2), \quad y_3 = 2ix_1(x_1 + ix_2) + x_3.$$

Replacing this in $\Xi(t, y_2, y_3)$ we get the value of the gradient of ϕ at (x_1, x_2, x_3):

$$\frac{\partial \phi}{\partial x_1} = 4ix_1\eta - 16ix_1^3 + 24x_1^2 - 8x_1x_3 - 2ix_1 + 2x_2 - 2\eta x_2 + 8ix_1x_2^2 - 4ix_3x_2$$
$$\frac{\partial \phi}{\partial x_2} = -2x_1\eta + 8x_1^3 + 8ix_1^2x_2 - 4ix_1x_3 + 2x_1 + 2ix_2$$
$$\frac{\partial \phi}{\partial x_3} = \eta - 4x_1^2 - 4ix_1x_2 + 2ix_3.$$

After some computations one arrives at

$$\begin{aligned}\phi = \eta(-2x_1x_2 + x_3) - 4x_3x_1^2 + 8x_1^3x_2 + 2x_1x_2 + \\ i((2\eta - 1)x_1^2 + x_2^2 + x_3^2 - 4x_3x_1x_2 - 4x_1^4 + 4x_1x_2^2),\end{aligned}$$

and sees that if $\eta > 1/2$ then

$$\operatorname{Im} \phi \geq c(x_1^2 + x_2^2 + x_3^2)$$

for some $c > 0$ in a neighborhood of 0. Repeating the rest of the argument used for the Mizohata operator we get that L is not solvable near 0.
The Lewy operator is in fact non-solvable at any point of $\mathbb{R}^3$. To see this, define first, for arbitrary $y \in \mathbb{R}^3$,

$$\ell_y : \mathbb{R}^3 \to \mathbb{R}^3, \quad \ell_y(x) = (y_1 + x_1, y_2 + x_2, y_3 + x_3 + 2(y_2x_1 - y_1x_2)).$$

If $x_0 = (x_1^0, x_2^0, x_3^0)$ is given and u is any function defined near x_0, then $u(\ell_{x^0}(x))$ is defined near 0 and

$$L(u(\ell_{x^0}(x))) = (Lu)(\ell_{x^0}(x)).$$

Whence it follows that the non-solvability of L near 0 yelds the non-solvability of L near x_0.

2.3 The necessity of (H)

Theorem 2.3.1. *Let P be a differential operator of principal type defined in a neighborhood U of 0 in $\mathbb{R}^n$ and let p be its principal symbol. Suppose that*

P is solvable on U, that is, for every $f \in C_0^\infty$ there is $u \in \mathcal{D}'(U)$ such that $Pu = f$. Then $H_{\mathsf{Re}\, p}\mathsf{Im}\, p = 0$ on Char(P).

Since $p(x,\xi) = (-1)^m p(x,-\xi)$ (m is the order of P), $\nu \in \mathrm{Char}(P) \iff -\nu \in \mathrm{Char}(P)$. Also, $H_{\mathsf{Re}\, p}\mathsf{Im}\, p$ is a polinomial in ξ of order $2m-1$, so if $H_{\mathsf{Re}\, p}\mathsf{Im}\, p$ has one sign at $\nu_0 \in \mathrm{Char}(P)$, then it has the opposite sign at $-\nu_0$. Thus if the quantity $H_{\mathsf{Re}\, p}\mathsf{Im}\, p$ is not identically zero on $\mathrm{Char}(P)$ then we can assume that it is positive at some point of $\mathrm{Char}(P)$. The proof consists of assuming that P is solvable but $H_{\mathsf{Re}\, p}\mathsf{Im}\, p \neq 0$ at some $\nu_0 \in \mathrm{Char}(P)$ and reaching a contradiction to the estimate in Lemma 2.2.1. We continue to write U for a neighborhood of 0 in $\mathbb{R}^n$. In the following lemma, f takes the place of $\mathsf{Im}\, p$ or $\mathsf{Re}\, p$. It is stated in a way that at the same time emphasizes its local nature, and the invariant context in which it will be used.

Lemma 2.3.2. *Let $\nu_0 \in T^*U\backslash 0$ with $\pi(\nu_0) = 0$. Let f be a smooth real-valued function defined near ν_0 in T^*U such that $\pi_* H_f(\nu_0) \neq 0$. Then there are coordinates $x_1, ..., x_n$ centered at 0 such that, in the induced canonical coordinates on T^*U near ν_0,*

$$H_f(\nu_0) = \frac{\partial}{\partial x_1}\Big|_{\nu_0}.$$

Proof. Since $\pi_* H_f(\nu_0) \neq 0$, there are coordinates $y_1, ..., y_n$ centered at 0 such that

$$\pi_* H_f(\nu_0) = \frac{\partial}{\partial y_1}\Big|_{y=0}.$$

Since, in the canonical coordinates y_i, η_j,

$$H_f = \sum_j \Big(\frac{\partial f}{\partial \eta_j}\frac{\partial}{\partial y_j} - \frac{\partial f}{\partial y_j}\frac{\partial}{\partial \eta_j}\Big),$$

this means that

$$\frac{\partial f}{\partial \eta_1}(\nu_0) = 1, \quad \frac{\partial f}{\partial \eta_j}(\nu_0) = 0 \text{ for } j = 2, ..., n.$$

The task is to find coordinates $x_1, ..., x_n$ with respect to which the above formulas still hold but in addition all derivatives $\frac{\partial f}{\partial y_j}$ vanish at 0. The latter condition is achieved by using a change of variables of the form

$$x_j = y_j + \frac{1}{2} \sum_{k,l} b^j_{kl} y_k y_l + O(|y|^3). \tag{2.16}$$

Here we have the symmetry condition $b^j_{kl} = b^j_{lk}$. The inverse of this change of coordinates is of the form

$$y_k = x_k - \frac{1}{2} \sum_{i,j} b^k_{ij} x_i x_j + O(|x|^3).$$

We have

$$\begin{aligned}
\sum_k \eta_k dy_k &= \sum_{i,j} \eta_j (\delta_{ij} - \sum_k b^i_{kj} x_k) dx_j + O(|x|^2) \\
&= \sum_{i,j} \eta_j (\delta_{ij} - \sum_k b^i_{kj} y_k) dx_j + O(|y|^2),
\end{aligned}$$

so

$$\xi_j = \sum_{i,} \eta_i (\delta_{ij} - \sum_k b^i_{kj} y_k) + O(|y|^2).$$

Thus

$$\frac{\partial}{\partial y_i} = \frac{\partial}{\partial x_i} - \sum_{j,k} \eta_k b^k_{ij} \frac{\partial}{\partial \xi_j} + O(|x|),$$

and

$$\frac{\partial}{\partial \eta_i} = \frac{\partial}{\partial \xi_j} + O(|x|).$$

Thus, modulo terms of order $O(|x|)$,

$$\begin{aligned}
H_f &= \frac{\partial f}{\partial \eta_1} \frac{\partial}{\partial y_1} - \sum_i \frac{\partial f}{\partial y_i} \frac{\partial}{\partial \eta_i} \\
&= \frac{\partial f}{\partial \eta_1} \Big(\frac{\partial}{\partial x_1} - \sum_{j,k} \eta_k b^k_{1j} \frac{\partial}{\partial \xi_j} \Big) - \sum_i \frac{\partial f}{\partial y_i} \frac{\partial}{\partial \xi_i} \\
&= \frac{\partial f}{\partial \eta_1} \frac{\partial}{\partial x_1} - \sum_i \Big(\frac{\partial f}{\partial y_i} + \frac{\partial f}{\partial \eta_1} \sum_k \eta_k b^k_{1i} \Big) \frac{\partial}{\partial \xi_i}.
\end{aligned}$$

We seek the vanishing of the coefficients of $\frac{\partial}{\partial \xi_i}$ at ν_0:

$$\frac{\partial f}{\partial y_i} + \frac{\partial f}{\partial \eta_1} \sum_k \eta_k b_{1i}^k = 0 \text{ for all } i.$$

Thus we take

$$b_{1i}^k = -\frac{\eta_k}{|\eta|^2} \frac{\partial f / \partial y^i}{\partial f / \partial \eta_1} + \gamma_i^k,$$

with γ_i^k such that $\sum_k \gamma_i^k \eta_k = 0$ for all i but otherwise arbitrary. Thus, recalling that $\partial f / \partial \eta_1 = 1$ at ν_0 we get

$$H_f(\nu_0) = \frac{\partial}{\partial x_1}\Big|_{\nu_0}.$$

□

In the following theorem we view T^*U as a subset of the complexification $\mathbb{C} \otimes T^*U$. The latter just means that we allow the coefficients ξ_j in $\sum \xi_j dx_j$ to be complex.

Theorem 2.3.3. *Let p be a smooth complex valued function defined in a complex neighborhood of $\nu_0 \in T^*U \backslash 0$. Suppose $p(\nu_0) = 0$, $\pi_* H_{\mathsf{Re}\,p}(\nu_0) \neq 0$, and $H_{\mathsf{Re}\,p} \mathsf{Im}\, p > 0$ at ν_0. Let ${}^t p(\nu) = p(-\nu)$. Then there is, for any given positive integer N, a smooth function ϕ defined in a neighborhood of $x_0 = \pi(\nu_0)$ such that*

$$d\phi(0) = -\nu_0 \text{ and } {}^t p \circ d\phi = O(|x - x_0|^{N+1}) \text{ as } x \to x_0. \tag{2.17}$$

Furthermore, there is $c > 0$ such that

$$\mathsf{Im}\, \phi(x) \geq c|x - x_0|^2 \quad \text{in a neighborhood of } x_0. \tag{2.18}$$

Proof. Using a translation we may assume that $x_0 = 0$. We first prove the claim when $N = 1$. The proof splits along two possibilities: either $\pi_* H_{\mathsf{Re}\,p}(\nu_0)$ and $\pi_* H_{\mathsf{Im}\,p}(\nu_0)$ are linearly dependent, or they are not. These are coordinate-independent properties. The Mizohata operator illustrates the first case while the Lewy operator is an example of the second case.

We deal with the linearly dependent case first. Since $\pi_* H_{\mathsf{Re}\,p}(\nu_0) \neq 0$, there is $\mu \in \mathbb{R}$ such that $\pi_* H_{\mathsf{Im}\,p}(\nu_0) = \mu \pi_* H_{\mathsf{Re}\,p}(\nu_0)$. Thus

$$\pi_* H_{p/(1+i\mu)}(\nu_0) = \frac{1}{1+i\mu}(\pi_* H_{\mathsf{Re}\,p}(\nu_0) + i\pi_* H_{\mathsf{Im}\,p}(\nu_0)) = \pi_* H_{\mathsf{Re}\,p}(\nu_0)$$

is a real vector. Replacing p by $p/(1+i\mu)$ we may thus assume that $\pi_* H_p(\nu_0)$ is itself real. This implies that, using Lemma 2.3.2 with $f = \mathsf{Re}\,p$ we can find coordinates $x_1, ..., x_n$ such that in the induced canonical coordinates

$$H_{\mathsf{Re}\,p}(\nu_0) = \frac{\partial}{\partial x_1}.$$

Thus

$$\frac{\partial p}{\partial \xi_1}(\nu_0) = 1, \quad \frac{\partial p}{\partial \xi_j}(\nu_0) = 0 \text{ if } j \geq 2, \quad \frac{\partial \mathsf{Re}\,p}{\partial x_j}(\nu_0) = 0, j = 1, ..., n. \tag{2.19}$$

With respect to the coordinates x_j, ξ_k, the covector ν_0 is $(0, \xi^0)$. Let

$$\psi = -x \cdot \xi^0 + \frac{1}{2}\sum_{i,j=1}^{n} \alpha_{ij} x_i x_j, \quad \alpha_{ij} = \alpha_{ji}, \text{ all } i, j, \tag{2.20}$$

and write $\alpha(x)$ for the vector with components $\alpha_i(x) = \sum_j \alpha_{ij} x_j$. So $d\psi(x) = (x, -\xi^0 + \alpha(x))$ and

$${}^t p \circ d\psi(x) = {}^t p(x, -\xi^0 + \alpha(x)) = p(x, -\xi^0 + \alpha(x)).$$

Recalling that $p(\nu_0) = 0$ we have

$$p(x, -\xi^0 + \alpha(x)) = \sum_j \Big(\frac{\partial p}{\partial x_j}(0, \xi^0) - \sum_k \frac{\partial p}{\partial \xi_k}(0, \xi^0)\alpha_{kj}\Big) x_j + O(|x|^2),$$

so we will be done if

$$\frac{\partial p}{\partial x_j}(0, \xi^0) - \sum_k \frac{\partial p}{\partial \xi_k}(0, \xi^0)\alpha_{kj} = 0 \quad \text{for all } j. \tag{2.21}$$

Using (2.19) these conditions reduce to

$$i\frac{\partial \mathsf{Im}\,p}{\partial x_j}(0, \xi^0) - \alpha_{1,j} = 0, \tag{2.22}$$

for all j. Thus also the α_{j1} are determined, but we are free to choose the α_{kj} for $k, j \geq 2$. We set $\alpha_{kj} = i\mu\delta_{kj}$ for these indices, with μ to be determined later. The hypothesis that $H_{\mathsf{Re}\,p}\mathsf{Im}\,p > 0$ at ν_0 gives $\frac{\partial \mathsf{Im}\,p}{\partial x_1} > 0$ at ν_0, so

$$\mathsf{Im}\,\alpha_{11} > 0.$$

Thus

$$\mathsf{Im}\,\psi = \frac{1}{2}\mathsf{Im}\,\alpha_{11}(x_1)^2 + \sum_{j=2}^{n} \frac{\partial \mathsf{Im}\,p}{\partial x_j}(0, \xi^0)x_1 x_j + \frac{1}{2}\mu \sum_{j=2}^{n}(x_j)^2,$$

and choosing μ large enough we get that $\mathsf{Im}\,\psi > c|x|^2$ for some $c > 0$. This completes the proof of the case $N = 1$ when $\pi_* H_{\mathsf{Re}\,p}(\nu_0)$ and $\pi_* H_{\mathsf{Im}\,p}(\nu_0)$ are linearly dependent. Suppose now that $\pi_* H_{\mathsf{Re}\,p}(\nu_0)$ and $\pi_* H_{\mathsf{Im}\,p}(\nu_0)$ are linearly independent. We choose coordinates $y_1, ..., y_n$ such that

$$\pi_* H_{\mathsf{Re}\,p}(\nu_0) = \frac{\partial}{\partial y_1}, \quad \pi_* H_{\mathsf{Im}\,p}(\nu_0) = \frac{\partial}{\partial y_2}, \tag{2.23}$$

and use the proof of Lemma 2.3.2 to get new coordinates (2.16) such that the conclusion of the lemma holds for $H_{\mathsf{Re}\,p}$. With the same kind of change of coordinates we can simultaneously ask that in the expression of $H_{\mathsf{Im}\,p}$ in the new coordinates,

$$\begin{aligned} H_{\mathsf{Im}\,p} &= \frac{\partial \mathsf{Im}\,p}{\partial \eta_2}\frac{\partial}{\partial y_2} - \sum_i \frac{\partial \mathsf{Im}\,p}{\partial y_i}\frac{\partial}{\partial \eta_i} \\ &= \frac{\partial \mathsf{Im}\,p}{\partial \eta_2}\Big(\frac{\partial}{\partial x_2} - \sum_{j,k} \eta_k b^k_{2j}\frac{\partial}{\partial \xi_j}\Big) - \sum_i \frac{\partial \mathsf{Im}\,p}{\partial y_i}\frac{\partial}{\partial \xi_i} \\ &= \frac{\partial \mathsf{Im}\,p}{\partial \eta_2}\frac{\partial}{\partial x_2} - \sum_i \Big(\frac{\partial \mathsf{Im}\,p}{\partial y_i} + \frac{\partial \mathsf{Im}\,p}{\partial \eta_2}\sum_k \eta_k b^k_{2i}\Big)\frac{\partial}{\partial \xi_i}, \end{aligned}$$

(see the proof of Lemma 2.3.2), the coefficients of $\frac{\partial}{\partial \xi_i}$ vanish at ν_0 when $i \geq 2$. So in the new coordinates,

$$H_{\mathsf{Re}\,p} = \frac{\partial}{\partial x_1}, \quad H_{\mathsf{Im}\,p} = \frac{\partial}{\partial x_2} - \frac{\partial \mathsf{Im} p}{\partial x_1}\frac{\partial}{\partial \xi_1}.$$

Writing H_p at ν_0 using these formulas gives

$$\frac{\partial p}{\partial \xi_1} = 1 \quad \frac{\partial p}{\partial \xi_2} = i, \quad \frac{\partial p}{\partial \xi_j} = 0 \text{ for } j > 2, \frac{\partial \mathsf{Re}\, p}{\partial x_1} = 0, \quad \frac{\partial p}{\partial x_j} = 0 \text{ for } j > 1$$

at $(0, \xi^0)$. We proceed as before with ψ given by (2.20). The condition that the linear terms vanish (see (2.21)) gives

$$\frac{\partial p}{\partial x_j}(0, \xi^0) - (\alpha_{1j} + i\alpha_{2j}) = 0 \quad \text{for all } j,$$

so $\alpha_{1j} + i\alpha_{2j}$ is determined (but not yet the individual coefficients) while the α_{jk} with $k, j \geq 3$ can be chosen arbitrarily. We take advantage of the latter fact by choosing, for these indices, $\alpha_{jk} = i\mu\delta_{kj}$ with positive μ. Since $\partial p/\partial x_j(\nu_0) = 0$ if $j \geq 2$ we may further take $\alpha_{jk} = 0$ for $j = 1, 2$ and $k \geq 3$. The matrix α is thus a block matrix whose top left 2×2 block we now specify. The conditions for $j = 1, 2$ are, respectively

$$i\frac{\partial \mathsf{Im}\, p}{\partial x_1}(0, \xi^0) - (\alpha_{11} + i\alpha_{21}) = 0 \quad \text{and } \alpha_{12} + i\alpha_{22} = 0.$$

The first of these equations gives

$$\mathsf{Im}\, \alpha_{11} = \frac{\partial \mathsf{Im}\, p}{\partial x_1}(0, \xi^0) - \mathsf{Re}\, \alpha_{21},$$

and the second,

$$\mathsf{Im}\, \alpha_{22} = \mathsf{Re}\, \alpha_{12}.$$

Note that $\mathsf{Im}\, \alpha_{12}$ is irrelevant, so we choose it to be zero. We pick α_{22} so that

$$0 < \mathsf{Im}\, \alpha_{22} < \frac{\partial \mathsf{Im}\, p}{\partial x_1}(0, \xi^0),$$

and α_{11} purely imaginary, with

$$\mathsf{Im}\, \alpha_{11} < \Big(\frac{\partial \mathsf{Im}\, p}{\partial x_1}(0, \xi^0) - \mathsf{Im}\, \alpha_{22}\Big).$$

Thus

$$\alpha_{ij} = \begin{pmatrix} i\mathsf{Im}\, \alpha_{11} & \mathsf{Im}\, \alpha_{22} & 0 \\ \mathsf{Im}\, \alpha_{22} & i\mathsf{Im}\, \alpha_{22} & 0 \\ 0 & 0 & i\mu I \end{pmatrix}$$

has positive definite imaginary part. This concludes the proof of the theorem when $N = 1$.

We now show that the proof for general $N \geq 1$ can be reduced to the case where p is analytic. We use the coordinates $x_1, ..., x_n$ centered at 0 obtained in either of the two cases discussed above and let, as before, x_j, ξ_k denote the canonical coordinates near $\nu_0 = (0, \xi^0)$. Let p_N be the Taylor polynomial of p of degree $N \geq 1$ based at ν_0 in these coordinates:

$$p_N(x,\xi) = \sum_{|\alpha|+|\beta|\leq N} \frac{1}{\alpha!\beta!}\frac{\partial^{|\alpha|+|\beta|}p}{\partial x_\alpha \partial \xi_\beta}(0,\xi^0)x^\alpha(\xi-\xi^0)^\beta.$$

For the error, $\tilde{p}_{N+1} = p - p_N$, we naturally have

$$\tilde{p}_{N+1}(z,\zeta) = \sum_{|\alpha|+|\beta|=N+1} p_{N+1,\alpha,\beta}(x,\xi)x^\alpha(\xi-\xi^0)^\beta,$$

for some functions $p_{N+1,\alpha,\beta}$. If ϕ is a function defined near 0 in U with $d\phi(0) = \nu_0$ and ${}^t\!p_N \circ d\phi = O(|x|^{N+1})$ then also ${}^t\!p \circ d\phi = O(|x|^{N+1})$. Indeed, in the coordinates x_j, ξ_k we have $d\phi(x) = (x, \nabla\phi(x))$, $\nabla\phi(0) = -\xi^0$, so $\nabla\phi(x) = -\xi^0 + \alpha(x)$ where $\alpha(x) = O(|x|)$, whence

$$\tilde{p}_{N+1}(x,\nabla\phi(x)) = \sum_{|\alpha|+|\beta|=N+1} p_{N+1,\alpha,\beta}(x,\nabla\phi(x))x^\alpha(-\alpha(x))^\beta = O(|x|^{N+1}).$$

Note that since $N \geq 1$, $H_{\mathsf{Re}\, p}\mathsf{Im}\, p(\nu_0) = H_{\mathsf{Re}\, p_N}\mathsf{Im}\, p_N(\nu_0)$. Thus we may work with p_N instead of p and add to the hypotheses of the theorem that p is real-analytic. Consider the Cauchy problem

$${}^t\!p_N(x,\nabla\phi(x)) = 0 \quad \phi\Big|_{x_1=0} = \psi\Big|_{x_1=0},$$

where ψ is the function (2.20) previously obtained. Write x' for $(x_2, ..., x_n)$ and let $\Xi_1(x')$ be defined near $x' = 0$ and satisfy

$$p(0,x',\Xi_1(x'),\nabla'\psi(0,x')) = 0, \quad \Xi_1(0) = -\xi_1^0. \tag{2.24}$$

Since $\nabla'\psi(0,0) = (-\xi_2^0, ..., -\xi_n^0)$ and $p(0,\xi^0) = 0$, and furthermore $\partial p/\partial\xi_1(\nu_0) = 1 \neq 0$, the holomorphic version of the Implicit Function Theorem gives the existence, uniqueness, and analyticity of the function Ξ_1 in a neighborhood of 0. Using the holomorphic version of the Hamilton-Jacobi method as before we get a solution ϕ of the Cauchy problem stated above. By construction, $\nabla\phi(0) = -\xi^0$. We now verify, making full use of the special coordinates we chose, that the Hessian of ϕ at 0 is the matrix α. This will imply (2.18) and conclude the proof of the theorem. Note in the first place that the initial condition already gives

$$\frac{\partial^2\phi}{\partial x_i \partial x_j}(0) = \alpha_{ij}, \quad i, j > 1.$$

To obtain these formulas for $j = 1$ we use the fact that

$$0 = {}^t p(x, \nabla\phi) = \sum_j \Big(\frac{\partial p}{\partial x_j}(0, \xi^0) - \sum_{k=1}^n \frac{\partial p}{\partial \xi_k}(0, \xi^0)\frac{\partial^2\phi}{\partial x_k \partial x_j}(0)\Big)x_j + O(|x|^2),$$

to conclude first that

$$\frac{\partial p}{\partial x_j}(0, \xi^0) - \sum_{k=1}^n \frac{\partial p}{\partial \xi_k}(0, \xi^0)\frac{\partial^2\phi}{\partial x_k \partial x_j}$$

for each j. The argument now splits as before. In the first case we discussed (where $\pi_* H_{\mathsf{Im}\, p}(\nu_0) = 0$) this gives, exactly as in (2.22) and with the same consequence, that

$$\frac{\partial^2\phi}{\partial x_k \partial x_j} = i\frac{\partial p}{\partial x_j}(0, \xi^0) = \alpha_{1j} \quad \text{for all } j.$$

In the second case (where $\pi_* H_{\mathsf{Im}\, p}(\nu_0)$ and $\pi_* H_{\mathsf{Re}\, p}(\nu_0)$ are linearly independent) we get

$$\frac{\partial p}{\partial x_1}(0, \xi^0) - \Big(\frac{\partial^2\phi}{\partial x_1 \partial x_1}(0) + i\frac{\partial^2\phi}{\partial x_1 \partial x_2}(0)\Big) = 0$$
$$\frac{\partial^2\phi}{\partial x_1 \partial x_j}(0) + i\frac{\partial^2\phi}{\partial x_2 \partial x_j}(0) = 0 \quad \text{for } j > 1,$$

which give

$$\frac{\partial^2\phi}{\partial x_1 \partial x_j}(0) = -i\frac{\partial^2\phi}{\partial x_2 \partial x_j}(0) = \alpha_{2j} \quad \text{for } j > 1,$$

and, as a consequence

$$\frac{\partial^2 \phi}{\partial x_1 \partial x_1}(0) = \frac{\partial p}{\partial x_1}(0, \xi^0) - i \frac{\partial^2 \phi}{\partial x_1 \partial x_2}(0) = \frac{\partial p}{\partial x_1}(0, \xi^0) - i\alpha_{12} = \alpha_{11}. \quad (2.25)$$

This completes the proof of the theorem. □

Suppose P is a differential operator of order m and principal symbol p and ϕ and v_0 are smooth. Then

$$^tP(x, D_x)(e^{i\tau\phi} v_0) = e^{i\tau\phi} \sum_{j=0}^{m} \tau^{m-j}\, {}^tP_{(m-j)}(x, D_x) v_0, \quad (2.26)$$

where ${}^tP_{(m-j)}(x, D_x)$ is of order j, ${}^tP_0(x, D_x) = {}^tP(x, D_x)$, ${}^tP_m(x, D_x)$ is a multiplication by $p(x, \nabla\phi(x))$, and

$$P_{m-1} = -\frac{\partial p}{\partial \xi_j}(x, -\nabla\phi(x)) D_{x_j} + c, \quad (2.27)$$

where c is a function. If ϕ satisfies (2.17)-(2.18), then

$$|\tau^m e^{i\tau\phi}\, {}^tP_m(x, D_x) v_0| \le C\tau^{m-(N+1)/2},$$

near x_0 so with N such that $m - (N + 1)/2 = -r$ with r a positive integer we get that this term decreases to 0 as $\tau \to \infty$. We wish to get the same kind of behavior for all terms $\tau^{m-j} e^{i\tau\phi}\, {}^tP_{m-j}(x, D_x) v_0$. We can get $\tau^{m-1} e^{i\tau\phi}\, {}^tP_{m-1}(x, D_x) v_0$ to have the right behavior if we can arrange that

$${}^tP_{m-1}(x, D_x) v_0 = O(|x - x_0|^{2(m+r-1)}),$$

but in general v_0 needs to be replaced by a polynomial in τ^{-1} (with smooth coefficients) in order to achieve estimates with τ^{-r}. We have in fact the following lemma.

Lemma 2.3.4. *Let P be a differential operator of principal type, of order m and principal symbol p. Suppose the hypotheses of Theorem 2.3.3 hold for p, and let ϕ satisfy (2.17)-(2.18), in a neighborhood V of x_0. Let N be a positive integer. There are $v_0, ..., v_k \in C_0^\infty$ such that*

$$\sup_{x \in V} |e^{-i\tau\phi t} P(x, D_x)\big(e^{i\tau\phi} \sum_{k=0}^{K} \tau^{-k} v_k\big)| \le \tau^{-r}$$

for some $C > 0$.

Proof. Without loss of generality we assume $x_0 = 0$. Using (2.26) we get

$$e^{-i\tau\phi t}P(x, D_x)\big(e^{i\tau\phi}\sum_{k=0}^{\ell}\tau^{-k}v_k\big) = \sum_{\ell=0}^{m+K}\tau^{m-l}\sum_{k+j=\ell}{}^{t}P_{(m-j)}(x, D_x)v_k,$$

where for simplify the formulas we have defined ${}^{t}P = 0$ if $j > m$ and $v_k = 0$ if $k > K$. We will specify ϕ and find the v_k so that

$$\sum_{k+j=\ell}{}^{t}P_{m-j}(x, D_x)v_k = O(|x|^{2(m+r-\ell)}) \quad \text{as } x \to 0, \tag{2.28}$$

without concern about their support. From this it will follow that

$$|\tau^{m-\ell}e^{i\tau\phi}\sum_{k+j=\ell}{}^{t}P_{m-j}(x, D_x)v_k| \le C\tau^{-r},$$

uniformly in a neighborhood of 0. We then replace each v_k by χv_k where $\chi \in C_0^\infty(V)$ and $\chi(x) = 1$ near 0 to arrange for the condition of the support of the v_k in the statement of the lemma. When $\ell = 0$, the left-hand side of (2.28) reduces to $p(x, -\nabla\phi(x))v_0$, so we pick $N = 2(m + r) - 1$ in Theorem 2.3.3 to obtain (2.28). Next, with $\ell = 1$ we get

$$\sum_{k+j=\ell}{}^{t}P_{m-j}(x, D_x)v_k = {}^{t}P_{m-1}(x, D_x)v_0 + {}^{t}P_m(x, D_x)v_1.$$

The second term on the right is $p(x, -\nabla\phi(x))v_1 = O(|x|^{2(m+r-\ell)})$ which is better than needed, so we dismiss it and focus on finding v_0 such that

$${}^{t}P_{m-1}(x, D_x)v_0 = O(|x|^{2(m+r-\ell)}).$$

We work in the coordinates of the proof of Theorem 2.3.3. Since $d\phi(0) = \nu_0$ and $\partial p/\partial x_1 \neq 0$ at ν_0, ${}^{t}P_{m-1}(x, D_x)$ is noncharacteristic for $x_1 = 0$. We find a solution of our problem by first replacing the coefficients of ${}^{t}P_{(m-1)}(x, D_x)$ by their Taylor polynomials of order $N_0 = 2(m-r)-3$ to reduce the problem to the analytic situation, as in the proof of Theorem 2.3.3; the reminders will then be $O(|x|^{2(m+r-\ell)})$ which is all that is needed. Letting ${}^{t}P_{m-1,N_0}(x, D_x)$ be the resulting operator, we then solve

$${}^{t}P_{m-1,N_0}(x, D_x)v_0 = 0, \quad v_0\Big|_{x_1=0} = 1,$$

in a neighborhood of 0 taking advantage of the analiticity of the coefficients. We now proceed by induction. Suppose that $k_0 \geq 1$ and that $v_0, ..., v_{k_0-1}$ have been found so that (2.28) holds for $\ell \leq k_0$. With $\ell = k_0 + 1$ the left hand side of (2.28) is

$$\sum_{k=0}^{k_0+1} {}^tP_{m-(k_0+1-k)}(x, D_x)v_k.$$

The term with $k = k_0 + 1$ is $p(x, -\nabla\phi(x))v_{k_0+1}$ which is already of order $O(|x|^{2(m+r)})$, so after dismissing it the problem becomes to ensure that

$$\sum_{k=0}^{k_0+1} {}^tP_{m-(k_0+1-k)}(x, D_x)v_k = O(|x|^{2(m+r-k_0-1)}).$$

Here only v_{k_0} is not yet known. We replace $\sum_{k=0}^{k_0+1} {}^tP_{m-(k_0+1-k)}(x, D_x)v_k$ by its Taylor polynomial of sufficiently high order, call it f_{k_0-1}, and then solve

$${}^tP_{(m-1,N_0)}(x, D_x)v_0 = -f_{k_0}, \quad v_{k_0}\Big|_{x^1=0} = 1.$$

This concludes the proof. $\square$

2.4 On the change of sign from + to −

The condition $H_{\mathsf{Re}\,p}\mathsf{Im}\,p > 0$ at some point $\nu_0 \in \mathrm{Char}(P)$ implies that $\mathsf{Im}\,p$ changes sign form − to + along the oriented integral curve $\gamma(t)$ of $H_{\mathsf{Re}\,p}$ through ν_0. We have seen how this change of sign enters in the proof of non-solvability. We may ask, on the other hand, what can happen when the opposite change of sign occurs, or if there is no change of sign at all. In this section we explore these possibilities through examples.

Consider $P = D_{x_1} + ix_1^k D_{x_2}$, first with k odd. The principal symbol is $p = \xi_1 + ix_1^k\xi_2$ with characteristic set

$$\mathrm{Char}(P) = \{(0, x_2; 0, \xi_2) : \xi_2 \neq 0\},$$

and $H_{\mathsf{Re}\,p}\mathsf{Im}\,p < 0$ on $\xi_2 < 0$. We show that the equation

$$Pu = f,$$

with arbitrary $f \in C_0^\infty(\mathbb{R}^2)$ has a solution "on $\xi_2 < 0$" in the sense that there is a Schwartz distribution $u \in \mathscr{S}'(\mathbb{R}^2)$ such that the partial Fourier transform $\widehat{Pu}(x_1, \xi_2)$ is equal to $\hat{f}(x_1, \xi_2)$ on $\xi_2 < 0$. Indeed, with f as specified, consider the equation

$$(D_{x_1} + i x_1^k \xi_2)\hat{u}(x_1, \xi_2) = \hat{f}(x_1, \xi_2).$$

The method of variation of the coefficients gives

$$\hat{u}(x_1, \xi_2) = i \int_0^{x_1} e^{(x_1^{k+1} - y_1^{k+1})\xi_2/(k+1)} \hat{f}(y_1, \xi_2) dy_1.$$

Evidently the quantity $(x_1^{k+1} - y_1^{k+1})\xi_2$ is nonpositive in $\xi_2 < 0$ when $y_1^{k+1} < x_1^{k+1}$, which is the case if $0 \le y_1 \le x_1$ or $x_1 \le y_1 \le 0$ because $k+1$ is even. But this is the case in the integral above.
Define

$$v(x_1, x_2) = \frac{i}{2\pi} \int_{-\infty}^{+\infty} e^{i x_2 \xi_2} \int_0^{x_1} e^{(x_1^{k+1} - y_1^{k+1})\xi_2/(k+1)} \hat{f}(y_1, \xi_2) dy_1 d\xi_2.$$

Then v is a Schwartz distribution such that

$$\widehat{Pv}(x_1, \xi_2) - \hat{f}(x_1, \xi_2) = 0 \quad \text{on} \quad \xi_2 < 0,$$

as required. We say that $Pv = f$ microlocally on $\xi_2 < 0$.
Consider now P as above but with k even. Then $\mathsf{Im}\, p$ does not change sign along any of the integral curves of $H_{\mathsf{Re}\, p}$ so we should expect P to be solvable. This is indeed the case: the equation $Pu = f$ with $f \in C_0^\infty(\mathbb{R}^2)$ has a solution, namely

$$u(x) = \frac{i}{2\pi} \int\int e^{i\phi(x_1, y_1, \xi_2)} K(x_1, y_1, \xi_2) \hat{f}(y_1, \xi_2) dy_1 d\xi_2,$$

where

$$\phi(x_1, y_1, \xi_2) = x_2 \xi_2 - i(x_1^{k+1} - y_1^{k+1})\xi_2/(k+1),$$
$$K(x_1, y_1, \xi_2) = H(-\xi_2)H(x_1 - y_1) - H(\xi_2)H(y_1 - x_1),$$

and H being the Heaviside function.

2.5 Estimates and solvability

The estimate in Lemma 2.2.1 is in terms of C^∞ seminorms. For some purposes it is better to use Sobolev norms and in particular it is convenient to remove the function f from the statement. We restate the lemma as follows.

Define for each $k \in \mathbb{Z}_+$, $\|f\|_k = \sum_{|\alpha| \leq k \|D^\alpha f\|}$

Proposition 2.5.1. *Suppose P is a differential operator defined in an open set U in $\mathbb{R}^n$ with the property that for every $f \in C_0^\infty(U)$ and open $W \Subset U$ with supp $f \subset W$ there is $u \in \mathcal{D}'(U)$ such that $Pu = f$ in W. Then, for any $V \Subset U$ there are C and N such that*

$$\forall v \in C_0^\infty(V) : \|v\|_{-N} \leq C\| {}^tPv\|_N. \tag{2.29}$$

Proof. The proof is essentially the same as for Lemma 2.2.1. Let $Y = C_0^\infty(V)$ with the topology determined by the seminorms

$$v \mapsto \sum_{|\beta| \leq N} \|D^\beta \, {}^tPv\|, \tag{2.30}$$

for each N. Here the norm is the L^2 norm. These seminorms are actually norms because $v \to {}^tPv$ is injective on $C_0^\infty(V)$. Let $X = C_0^\infty(\bar{V})$ with its standard topology. This topology is the same as that defined by the family of norms

$$f \to ||f||_N = \sum_{|\alpha| \leq N} \|D_x^\alpha f\|.$$

As in the proof of Lemma 2.2.1, the bilinear form

$$B : C_0^\infty(\bar{V}) \times C_0^\infty(V) \to \mathbb{C}, \quad B(f, v) = \int f v dx,$$

is separately continuous. Indeed, on the one hand the Cauchy-Schwarz inequality,

$$|B(f, v)| \leq \|f\|\|v\|,$$

gives the continuity in the first variable. On the other, if $f \in C_0^\infty(V)$ and $Pu = f$ ($u \in \mathcal{D}'(U)$ with $W \Subset U$ open containing $\bar{V}$), then

$$B(f, v) = \langle Pu, v\rangle = \langle u, {}^tPv\rangle. \tag{2.31}$$

But the restriction of u to a neighborhood $W \Subset U$ of $\bar{V}$ belongs to some Sobolev space $H^{-N}(W)$, therefore

$$|B(f, v)| \leq \|u\|_{-N} \| {}^tPv\|_N.$$

Since X is a Fréchet space, B is continuous: there are C, N, and M such that

$$|\int fvdx| \leq C||f||_M||^tPv||_N \quad \forall f \in C_0^\infty(\bar{V}), v \in C_0^\infty(V). \tag{2.32}$$

Thus

$$\|v\|_{-N} = \sup_{f \in H^M(\bar{V}), \|f\|_M = 1} |\int fvdx| \leq C\|^tPv\|_N.$$

Replacing M and N by $\max(M, N)$ we get the estimate in the form (2.29).

□

The estimate (2.29) is in fact equivalent to solvability:

Proposition 2.5.2. *Let P be a differential operator defined on an open set $U \subset \mathbb{R}^n$. Suppose that for any $V \Subset U$ there are C and N such that (2.29) holds. Then P is solvable on U.*

Proof. Let $f \in C_0^\infty(U)$ be arbitrary and let $V \Subset U$ be a neighborhood of supp f. Using (2.29) we obtain

$$|\langle f, v\rangle| \leq \|f\|_N\|v\|_{-N} \leq C\|f\|_N\|^tPv\|_N \quad \forall v \in C_0^\infty(V),$$

which shows that the linear form

$${}^tPC_0^\infty(V) \ni {}^tPv \mapsto \langle f, v\rangle \in \mathbb{C}$$

is continuous (recall that ${}^tP$ is injective) in the topology of $H_0^N(V)$.
By the Hahn-Banach Theorem this functional has an extension to a continuous linear functional $u : H_0^N(V) \to \mathbb{C}$, that is, there is $u \in H^{-N}(V)$ such that

$$\langle u, {}^tPv\rangle = \langle f, v\rangle \quad v \in C_0^\infty(V).$$

Thus $Pu = f$ in V □

Chapter 3

Constant-coefficient PDEs

In this chapter we present two different proofs of the Malgrange-Ehrenpreis theorem about the solvability of PDE with constant coefficients. Malgrange-Ehrenpreis theorem says that every constant coefficient linear partial differential equations have a fundamental solution E, i.e. there exists a distribution E s.t. $P(D)E = \delta$ and so there exists a solution of the equation $P(D)u = f$ with $f \in \mathscr{E}'$, $u = E * f$ is a solution of it.
In the last part of the chapter we give an elementary proof due to D. Jerison [14], of the L^2 local solvability, in which use is made of the SAK principle by C. Fefferman and D. H. Phong (see [7]).

3.1 Atiyah's proof of the Malgrange-Ehrenpreis theorem

In this proof, Atiyah uses the Hironaka theorem on the resolution of singularities in order to prove the Hörmander-Lojasiewicz theorem on the division of distributions and hence the existence of temperate fundamental solutions for constant-coefficient differential operators.

Theorem 3.1.1. *Let X be a real analytic manifold (paracompact and connected) and let $f, g_1, g_2, ..., g_p$ be real analytic functions on X with f nonnegative and not identically 0. Let Γ denote the characteristic function of the* *set*

$$G = \{x \in X | g_i(x) \geq 0 \textit{ for all } i\}$$

Then the function $f^s\Gamma$, *which is locally integrable for* $\mathsf{Re}s > 0$, *extends as an analytic function to a distribution on* X *which is a meromorphic function of* s *in the whole complex plane. Any given relatively compact open set* U *in* X *the poles of* $f^s\Gamma$ *occur at points of the form* $-r/N, r = 1, 2, ...,$ *where* N *is a fixed integer (depending on* f *and* U*) and the order of every pole does not exceed the dimension of* X*. For* $s = 0$ *we have* $f^0\Gamma = \Gamma$.

Before Proving the theorem let us first deduce the corollaries on the division of distributions.

Corollary 3.1.2. *Let* X *be a real analytic manifold,* $f : X \to \mathbb{C}$ *be an analytic function (*$f \not\equiv 0$*). Then there exists a distribution* T *on* X *such that* $fT = 1$.

Proof. It is enough to prove the corollary for $f \geq 0$ because, if S is an inverse of $|f|^2 = \bar{f}f$, then $T = \bar{f}S$ is an inverse of f. Applying the theorem we can expand f^s ,over U, around $s = -1$ in the form

$$f^s = \sum_{-n}^{\infty} a_k(s+1)^k, \quad n = \dim X, \tag{3.1}$$

where each a_k is a distribution. But $f \cdot f^s = f^{s+1}$ cannot have a pole at $s = -1$ (since $f^0 = 1$) and so we must have

$$fa_k = 0 \quad \text{for} \quad k < 0$$
$$fa_0 = f^0 = 1.$$

Thus, over U, $T = a_0$ is the required inverse of f. If $V \supset U$ is another open set, the expansions of f^s for U and V are necessarily compatible, though the region of convergence (around $s = -1$) may be smaller for V than for U. The distribution $T = a_0$ therefore exists on the whole of X. $\square$

Corollary 3.1.3. *Let f be a polynomial on $\mathbb{R}^n$ with complex coefficients ($f \not\equiv 0$). Then there exists a temperate distribution T on $\mathbb{R}^n$ such that $fT = 1$.*

Proof. Let $m = \deg f$; then the function g, defined by

$$g(x_1, ..., x_n) = \frac{f(x_1, ..., x_n)}{\left(1 + \sum_1^n x_i^2\right)^m},$$

extends to an analytic function on the n-sphere $\mathbb{S}^n = \mathbb{R}^n \cup \{\infty\}$ (it is enough to compose g with the stereographic projection). By Corollary 3.1.2, there is a distribution Q on $\mathbb{S}^n$ such that $gQ = 1$. The restriction of Q to $\mathbb{R}^n \subset \mathbb{S}^n$ is then a temperate distribution. Now put

$$T = \frac{Q}{\left(1 + \sum_1^n x_i^2\right)^m} \quad \text{on} \quad \mathbb{R}^n.$$

Then T is also temperate and

$$fT = \frac{fQ}{\left(1 + \sum_1^n x_i^2\right)^m} = gQ = 1$$

□

Taking the Fourier transform in Corollary 3.1.3 we obtain in the well-known way that

Corollary 3.1.4. *Every constant-coefficient partial differential operator which is not identically* 0 *has a temperate fundamental solution.*

Having explained these corollaries we now return to the main theorem, and make a number of preliminary remarks.

(i) The theorem is of a local character so that it is sufficient to prove it for small neighborhoods of the origin in $\mathbb{R}^n$.

(ii) The theorem is classical when $X = \mathbb{R}$, $f(x) = x^2$ and $G = \mathbb{R}$ or $\mathbb{R}^+$: the poles occur at points $-r/2$.

(iii) For $X = \mathbb{R}$, $f(x) = x^N$, N even, and $G = \mathbb{R}$ or $\mathbb{R}^+$ the theorem follows from Remark (ii) by re-indexing ($f(x) = (x^{N/2})^2$).

(iv) Taking products, the theorem now follows for $X = \mathbb{R}^n$, $f(x) = \Pi x_i^{N_i}$, N_i even, and $G = \prod G_i$, where each $G_i = \mathbb{R}$ or $\mathbb{R}^+$.

We also need for the proof of the thm. the following version of Hironaka's theorem.

Theorem 3.1.5. *Let F be a real analytic function ($F \not\equiv 0$), defined in a neighborhood of $0 \in \mathbb{R}^n$. Then there exists an open set $U \ni 0$, a real analytic manifold $\tilde{U}$ and a proper analytic map $\phi : \tilde{U} \longrightarrow U$ such that*

(a) $\phi : \tilde{U} \backslash \tilde{A} \longrightarrow U \backslash A$ is an isomorphism, where $A = F^{-1}(0)$ and $\tilde{A} = \phi^{-1}(A) = (F \circ \phi)^{-1}(0)$,

(b) for each $P \in \tilde{U}$ there are local analytic coordinates $(y_1, ..., y_n)$ centered at P so that, locally near P, we have

$$F \circ \phi = \epsilon \cdot \prod_{i=1}^{n} y_i^{k_i},$$

where ϵ is an invertible analytic function and $k_i \geq 0$.

The basic idea, for the proof of our theorem, is to use the Hironaka theorem to reduce the problem to the simple cases described earlier in Remark (iv). Before proceding further, however, we must make a few general remarks about distributions on a manifold.

For any given n-dimensional C^∞ manifold X let $\Omega(X)$ be the space of C^∞ *exterior differential n-forms* with compact support and $\mathscr{D}(X)$ the space of C^∞ *densities* with compact support. In local coordinates, $\omega \in \Omega(X)$ and $\mu \in \mathscr{D}(X)$ can be respectively written as

$$\omega = f(x)dx, \quad \mu = g(x)|dx|,$$

where f, g are C^∞ functions, $dx = dx_1 \wedge dx_2 \wedge ... \wedge dx_n$ and $|dx|$ denotes the Lebesgue measure on $\mathbb{R}^n$. If α is a local choice of orientation, then μ

determines locally an n-form which we write as $\alpha\mu$. If X is itself orientable, then a global orientation enables us to identify $\Omega(X)$ and $\mathscr{D}(X)$.

The space $\mathscr{D}'(X)$ of distributions is defined as the dual of $\mathscr{D}(X)$. Since a locally integrable function f defines a linear form on $\mathscr{D}(X)$, by $\mu \mapsto \int f\mu$, we have $f \in \mathscr{D}'(X)$. The space $\Omega'(X)$, dual of $\Omega(X)$, may be called the space of *"twisted" distributions.* If $\phi : X \to Y$ is a proper C^∞ map of manifolds, it induces a map $\phi^* : \Omega(Y) \to \Omega(X)$ and hence by duality a direct image homomorphism $\phi_* : \Omega'(X) \to \Omega'(Y)$.

In order to prove the theorem it is sufficient to consider the local situation, so we can take X to be a neighborhood of $0 \in \mathbb{R}^n$. Put $F = f\prod_1^n g_i$ and take $U, \tilde{U}, \phi$ as given by the Hironaka theorem. Let $\tilde{f} = f \circ \phi$, $\tilde{g}_i = g_i \circ \phi$ be the induced functions on $\tilde{U}$. For any given point $P \in \tilde{U}$ the local factorization

$$\tilde{F} = \tilde{f}\prod_1^p g_i = \epsilon \prod_1^n y_i^{k_j},$$

implies a corresponding local factorization of $\tilde{f}$ and each $\tilde{g}_i$. Moreover, since $\tilde{f} \geq 0$ the exponents for $\tilde{f}$ are necessarly even. Now let α denote the standard orientation of U (inherited from $\mathbb{R}^n$), $\tilde{\alpha}$ the corresponding orientation of $\tilde{U}\backslash\tilde{A}$. Then, for $\mathsf{Re}s > 0$,

$$\tilde{f}^s\tilde{\Gamma}\tilde{\alpha} \in \Omega(\tilde{U}\backslash\tilde{A}),$$

extends to a locally integrable n-form on $\tilde{U}$ (here $\tilde{\Gamma} = \Gamma\circ\phi$ is the characteristic function of the set defined by $\tilde{g}_i \geq 0$ for all i). In the neighborhood of $P \in \tilde{U}$ the orientation $\tilde{\alpha}$ must be of the form

$$\tilde{\alpha} = \Big(\prod_{j\in S} \operatorname{sgn}\, y_i\Big)\beta, \tag{3.2}$$

where β is the orientation given by the local coordinates $(y_1, ..., y_n)$ and S is a subset of $(1, ..., n)$ defined as follows: $j \in S$ if $\tilde{\alpha}$ changes as we cross the hyperplane $y_i = 0$. Since $\operatorname{sgn}\, y_i = 2\Gamma_j - 1$ (where Γ_j is the characteristic function of $y_i \geq 0$), it follows that $\tilde{f}^s\tilde{\Gamma}\tilde{\alpha}$ is locally a sum of expressions of the form

$$\epsilon \cdot \prod y_i^{2sM_i} \cdot \prod \Gamma_j \cdot \beta,$$

where ϵ is an invertible analytic function and j runs over some subset of $(1,..,n)$. By Remark (iv) this implies that $\tilde{f}^s\tilde{\Gamma}\tilde{\alpha}$ extends analytically in a neighborhood of P, and has poles as specified in the theorem. Since this holds for any $P \in \tilde{U}$, we get on the whole $\tilde{U}$ a twisted distribution $\tilde{f}^s\tilde{\Gamma}\tilde{\alpha}$ depending meromorphically on s. Applying the direct immage $\phi_* : \Omega'(\tilde{U}) \to \Omega'(U)$ we obtain a twisted distribution $\phi_*(\tilde{f}^s\tilde{\Gamma}\tilde{\alpha})$ on U, depending meromorphically on s and with poles as in the theorem. Since we have an orientation α on U which induces an isomorphism $\Omega'(U) \cong \mathscr{D}'(U)$, we get

$$\phi_*(\tilde{f}^s\tilde{\Gamma}\tilde{\alpha}) = T(s)\alpha,$$

where $T(s) \in \mathscr{D}'(U)$. To complete the proof of the theorem it remains to check that, for $\mathsf{Re}\, s > 0$, $T(s) = f^s\Gamma$. As we have already observed, for $\mathsf{Re}\, s > 0$, $\tilde{f}^s\tilde{\Gamma}\tilde{\alpha}$ is a locally integrable n-form on $\tilde{U}$, which is determined by its restriction to $\tilde{U}\backslash\tilde{A}$. Similarly, $f^s\Gamma$ is locally integrable and determined by its restriction to $U\backslash A$. Since ϕ induces an isomorphism $\tilde{U}\backslash\tilde{A} \to U\backslash A$, it follows that

$$\phi_*(\tilde{f}^s\tilde{\Gamma}\tilde{\alpha}) = f^s\Gamma\alpha, \quad \mathsf{Re} s > 0,$$

and the proof is complete.

3.2 Hörmander's version of the Malgrange-Ehrenpreis theorem

3.2.1 Temperate weight functions

Definition 3.2.1. A positive function k defined in $\mathbb{R}^n$ will be called a *temperate weight function* if there exist positive constants C and N such that

$$k(\xi + \eta) \leq (1 + C|\xi|)^N k(\eta) \quad \xi, \eta \in \mathbb{R}^n. \tag{3.3}$$

The set of all such functions k will be denoted by $\mathscr{K}$.

From the inequality (3.3) it follows that

$$(1 + C|\xi|)^{-N} \leq k(\xi + \eta)/k(\eta) \leq (1 + C|\xi|)^N. \tag{3.4}$$

If we let $\xi \to 0$ in (3.4) it follows that k is continuous, and when $\eta = 0$ we obtain the estimates

$$k(0)(1 + C|\xi|)^{-N} \leq k(\xi) \leq k(0)(1 + C|\xi|)^N. \tag{3.5}$$

If $k \in \mathscr{K}$ we shall write

$$M_k(\xi) = \sup_{\eta} k(\xi + \eta)/k(\eta). \tag{3.6}$$

This means that M_k is the smallest function such that

$$k(\xi + \eta) \leq M_k(\xi)k(\eta). \tag{3.7}$$

It also follows immediately that M_k is submultiplicative,

$$M_k(\xi + \eta) \leq M_k(\xi)M_k(\eta), \tag{3.8}$$

and since $M_k(\xi) \leq (1 + C|\xi|)^N$ this implies that $M_k \in \mathscr{K}$.

Example 3.2.1. The example of a function in $\mathscr{K}$ which occurs frequently is

$$k_s(\xi) = (1 + |\xi|^2)^{s/2}, \quad s \in \mathbb{R}.$$

To prove that $k_s \in \mathscr{K}$ it is sufficient to prove that $k_2 \in \mathscr{K}$. In fact if $k \in \mathscr{K}$ then $k^s \in \mathscr{K}$ for every real s, and this follows from the estimates

$$1 + |\xi + \eta|^2 \leq 1 + |\xi|^2 + 2|\xi||\eta| + |\eta|^2 \leq (1 + |\xi|^2)^2(1 + |\eta|)^2.$$

Example 3.2.2. The basic example of a function in $\mathscr{K}$, which is the reason for the definition of this class, is the function $\tilde{P}$ defined by

$$\tilde{P}(\xi)^2 = \sum_{|\alpha| \geq 0} |P^{(\alpha)}(\xi)|^2, \quad P^\alpha = \partial_\xi^\alpha P, \tag{3.9}$$

where P is a polinomial, which yields that the sum is finite. It follows immediately from Taylor's formula that

$$\tilde{P}(\xi + \eta) \leq (1 + C|\xi|)^m \tilde{P}(\eta), \tag{3.10}$$

where m is the degree of P and C a constant depending only on m and the dimension n.

3.2.2 The space $\mathscr{B}_{p,k}$

Definition 3.2.2. If $k \in \mathscr{K}$ and $1 \leq p \leq \infty$, we denote by $\mathscr{B}_{p,k}$ the set of all temperate distribution $u \in \mathscr{S}'$ such that $\hat{u}$ is a function and

$$\|u\|_{p,k} = ((2\pi)^{-n} \int |k(\xi)\hat{u}(\xi)|^p d\xi)^{1/p} < \infty. \tag{3.11}$$

Theorem 3.2.3. *Let $P(D)$ be a differential operator. If $u \in \mathscr{B}_{p,k}$ it follows that $P(D)u \in \mathscr{B}_{p,k/\tilde{P}}$.*

Proof. Since the Fourier transform of $P(D)u$ is $P(\xi)\hat{u}(\xi)$ and since $|P(\xi)\hat{u}(\xi)| \leq |\tilde{P}(\xi)\hat{u}(\xi)|$, the statement is trivial. □

Theorem 3.2.4. *If $u \in \mathscr{B}_{p,k}$ and $\phi \in \mathscr{S}$, it follows that $\phi u \in \mathscr{B}_{p,k}$ and that*

$$\|\phi u\|_{p,k} \leq \|\phi\|_{1,M_k} \|u\|_{p,k}. \tag{3.12}$$

Proof. We know that the Fourier transform of $v = \phi u$ is the convolution

$$\hat{v}(\xi) = (2\pi)^{-n} \int \hat{\phi}(\xi - \eta)\hat{u}(\eta)d\eta, \tag{3.13}$$

when $\hat{\phi} \in C_0^\infty$. Multiplying (3.13) by $k(\xi)$ and noting that $k(\xi) \leq M_k(\xi - \eta)k(\eta)$, we obtain $|k\hat{v}| \leq (2\pi)^{-n}|M_k\hat{\phi}| * |k\hat{u}|$. Hence Minkowski's inequality in integral form gives $\|k\hat{v}\|_p \leq (2\pi)^{-n}\|M_k\hat{\phi}\|_1 * \|k\hat{u}\|_p$, which is equivalent to the estimate (3.12). Since C_0^∞ is dense in $\mathscr{S}$, the result immediately extends to an arbitrary $\phi \in \mathscr{S}$. □

3.2.3 The space $\mathscr{B}_{p,k}^{loc}$

Definition 3.2.5. Let Ω be an open subset of $\mathbb{R}^n$. A linear subspace $\mathscr{F}$ of $\mathscr{D}'(\Omega)$ is called *semi-local* if $\phi u \in \mathscr{F}$ when $u \in \mathscr{F}$ and $\phi \in C_0^\infty(\Omega)$.
It is called *local* if, in addition, $\mathscr{F}$ contains every distribution u such that $\phi u \in \mathscr{F}$ for every $\phi \in C_0^\infty(\Omega)$.

Example 3.2.3. $\mathscr{D}'(\Omega), C^k(\Omega), L^p_{loc}(\Omega)$ are local space, whereas $\mathscr{D}'_F(\Omega)$ (distributions with finite order), $\mathscr{E}'(\Omega), L^p(\Omega)$ are semi-local but not local.

Example 3.2.4. It follows from Theorem 3.2.4 that the set of restrictions to Ω of distributions in $\mathscr{B}_{p,k}$ is semi-local.

Theorem 3.2.6. *If $\mathscr{F}$ is semi-local, the smallest local space containing $\mathscr{F}$ is the space*

$$\mathscr{F}^{loc} = \{u; u \in \mathscr{D}'(\Omega), \phi u \in \mathscr{F} \text{ for every } \phi \in C_0^\infty(\Omega)\}.$$

Proof. Since $\mathscr{F}$ is semi-local, we have $\mathscr{F} \subset \mathscr{F}^{loc}$. It is also clear that $\mathscr{F}^{loc}$ is semi-local. To prove that $\mathscr{F}^{loc}$ is local, we take a distribution u such that $\phi u \in \mathscr{F}^{loc}$ for every $\phi \in C_0^\infty(\Omega)$. Choose $\psi \in C_0^\infty(\Omega)$ so that $\psi = 1$ in the support of ϕ. Then it follows that $\phi u = \psi(\phi u) \in \mathscr{F}$, in view of the definition of $\mathscr{F}^{loc}$. Hence $u \in \mathscr{F}^{loc}$, so that $\mathscr{F}^{loc}$ is a local space. It is obvious that it is the smallest local space containing $\mathscr{F}$. □

Theorem 3.2.7. *If $u \in \mathscr{B}^{loc}_{p,k}(\Omega)$ we have $P(D)u \in \mathscr{B}^{loc}_{p,k/\tilde{P}}(\Omega)$.*

Proof. For any given $\phi \in C_0^\infty(\Omega)$ we can choose $\psi \in C_0^\infty(\Omega)$ so that $\psi = 1$ in a neighborhood of the support of ϕ. Since $\psi u \in \mathscr{B}_{p,k}$ it then follows from Theorems 3.2.3 and 3.2.4 that, using $\psi\phi = \phi$

$$\phi P(D)u = \phi P(D)(\psi u) \in B_{p,k/\tilde{P}}, \tag{3.14}$$

which proves the theorem. □

3.2.4 Existence of fundamental solutions

Definition 3.2.8. A distribution $E \in \mathscr{D}'(\mathbb{R}^n)$ is called a *fundamental solution* for the differential operator $P(D)$ with constant coefficients if

$$P(D)E = \delta \tag{3.15}$$

where δ is the Dirac measure at 0.

Theorem 3.2.9. *To every differential operator $P(D)$ there exists a fundamental solution $E \in \mathscr{B}^{loc}_{\infty,\tilde{P}}(\mathbb{R}^n)$. More precisely, to every $\epsilon > 0$ there exists a* *fundamental solution E such that $E/\cosh(\epsilon|x|) \in \mathscr{B}_{\infty,\tilde{P}}$ and $\|E/\cosh(\epsilon|x|)\|_{\infty,\tilde{P}}$ is bounded by a constant depending only on ϵ, the dimension n and the degree m of P.*

Proof. The main step in the proof is the estimate given by the following lemma.

Lemma 3.2.10. *For every $\epsilon > 0$ there exists a constant C depending only on ϵ, n and m such that*

$$|u(0)| \leq C\|\cosh(\epsilon|x|)P(D)u\|_{1,1/\tilde{P}}, \quad u \in C_0^\infty(\mathbb{R}^n). \tag{3.16}$$

We shall first prove that Theorem 3.2.9 follows from Lemma 3.2.10. Note that Definition 3.2.8 says that the distribution E is a fundamental solution if the linear form $\breve{E}(v) = E * v(0)$ on $C_0^\infty(\mathbb{R}^n)$ satisfies the identity

$$u(0) = \breve{E}(P(D)u), \quad \forall u \in C_0^\infty(\mathbb{R}^n). \tag{3.17}$$

In fact

$$\langle\delta|u\rangle = u(0) = \langle\delta|E * Pu\rangle = \langle\delta|PE * u\rangle = \langle(PE)\check{} * \delta|u\rangle = \\ \langle PE * \check{\delta}|u\rangle = \langle PE|u\rangle, \quad \forall u \in C_0^\infty(\mathbb{R}^n),$$

where $\check{v}(x) = v(-x)$, (for a review of the inequalities that we have used above one can see [33]).

In other words E is a fundamental solution if $\breve{E}$ is an extension of the linear form $P(D)u \mapsto u(0)$, $u \in C_0^\infty(\mathbb{R}^n)$. In view of the Hahn-Banach theorem and (3.16), a linear form $\breve{E}$ on $C_0^\infty(\mathbb{R}^n)$ satisfying (3.17) can thus be constructed so that

$$|\breve{E}(v)| \leq C\|(\cosh\epsilon|x|)v)\|_{1,1/\tilde{P}}, \quad v \in C_0^\infty(\mathbb{R}^n). \tag{3.18}$$

If we write $E_\epsilon = E/\cosh\epsilon|x|$, this means that

$$|\breve{E}(v)| \leq C\|v\|_{1,1/\tilde{P}}, \quad v \in C_0^\infty(\mathbb{R}^n). \tag{3.19}$$

Hence $E_\epsilon \in \mathscr{B}_{\infty,\tilde{P}}$ (since $\mathscr{B}'_{1,1/\tilde{P}} = \mathscr{B}_{\infty,\tilde{P}}$), which proves Theorem 3.2.9. $\square$

The proof of Lemma 3.2.10 will be obtained as a result of a few lemmas concerning analytic functions.

Lemma 3.2.11. *If f is an analytic function of a complex variable t when $|t| \leq 1$, and p is a polinomial in which the coefficient of the highest order term is A we have the inequality*

$$|Af(0)| \leq (2\pi)^{-1} \int_0^{2\pi} |f(e^{i\theta})p(e^{i\theta})| d\theta. \tag{3.20}$$

Proof. Let m be the degree of p and let q be the polynomial $q(t) = t^m \bar{p}(1/t)$ where $\bar{p}$ is obtain by taking complex conjugates of the coefficients of p. Then we have $q(0) = \bar{A}$ and $|q(e^{i\theta})| = |p(e^{i\theta})|$ so that 3.20 reduces to the familiar inequality

$$|f(0)q(0)| \leq (2\pi)^{-1} \int_0^{2\pi} |f(e^{i\theta})q(e^{i\theta})| d\theta.$$

□

Lemma 3.2.12. *With the notation of Lemma 3.2.11 we have, if the degree of p is $\leq m$ and $C_{m,k} = m!/(m-k)!$,*

$$|f(0)p^{(k)}(0)| \leq C_{m,k}(2\pi)^{-1} \int_0^{2\pi} |f(e^{i\theta})p(e^{i\theta})| d\theta. \tag{3.21}$$

Proof. We may assume that the degree of p is equal to m and write

$$p(t) = \prod_1^m (t - t_j). \tag{3.22}$$

Applying the previous lemma to the polynomial $\prod_1^m (t - t_j)$ and the analytic function $f(t) \prod_1^m (t - t_j)$, we obtain

$$\left| f(0) \prod_1^m t_j \right| \leq (2\pi)^{-1} \int_0^{2\pi} |f(e^{i\theta})p(e^{i\theta})| d\theta. \tag{3.23}$$

A similar inequality holds for any $(m-k)$-fold product of the numbers t_j on the left-hand side, and since $p^{(k)}(0)$ is the sum of $C_{m,k}$ terms, the inequality (3.21) follows.

Note that (3.21) reduces to (3.20) when $k = m$ and is trivial when $k = 0$. □

Before extending Lemma 3.2.12 in several variables we shall give it a slightly more general form. Suppose that f is entire and apply (3.21) to the function $f(rt)$ and the polynomial $p(rt)$ where $r > 0$. This gives

$$|f(0)p^{(k)}(0)|2\pi r^k \leq C_{m,k} \int_0^{2\pi} |f(re^{i\theta})p(re^{i\theta})|d\theta.$$

Let $\psi(r)$ be a nonnegative integrable function with compact support. Multiplying by $r\psi(r)$ and integrating w.r.t. r, we obtain

$$|f(0)p^{(k)}(0)| \int_{\mathbb{C}} |t^k||\psi(|t|)|L(dt) \leq C_{m,k} \int_{\mathbb{C}} |f(t)p(t)|\psi(|t|)L(dt), \tag{3.24}$$

where $L(dt)$ stands for the Lebesgue measure $rdrd\theta$ and the integrals are extended over the whole complex plane. The following generalization to several variables follows immediately by applying (3.24) to the variables $\xi_1, ..., \xi_n$.

Lemma 3.2.13. *Let $F(\xi)$ be an entire function and $P(\xi)$ a polynomial of degree $\leq m$ in $\xi = (\xi_1, ..., \xi_n) \in \mathbb{C}^n$. Let $\Psi(\xi)$ be a nonnegative integrable function with compact support, depending only on $|\xi_1|, ..., |\xi_n|$. Then*

$$|F(0)P^{(\alpha)}(0)| \int_{\mathbb{C}^n} |\xi^\alpha|\Psi(\xi)L(d\xi) \leq C_{m,|\alpha|} \int_{\mathbb{C}^n} |F(\xi)P(\xi)|\Psi(\xi)L(d\xi), \tag{3.25}$$

where $L(d\xi)$ is the Lebesgue measure in $\mathbb{C}^n$.

Proof of Lemma 3.2.10. Let $u \in C_0^\infty(\mathbb{R}^n)$ and write $P(D)u = v$. We then have $P(\xi)\hat{u}(\xi) = \hat{v}(\xi)$. With fixed ζ we apply Lemma 3.2.13 to $F(\xi) = \hat{u}(\xi + \zeta)$ and to the polynomial $P(\xi + \zeta)$, taking $\Psi(\xi) = 1$ if $|\xi| < \epsilon/2$ and $\Psi(\xi) = 0$ otherwise. Adding over α and noting that $\tilde{P}(\zeta) \leq \sum |P^{(\alpha)}(\zeta)|$, we obtain with a constant C_1 depending only on ϵ, n and m

$$\begin{aligned} |\hat{u}(\zeta)|\tilde{P}(\zeta) \leq & C_1 \int_{\mathbb{C}^n} |\hat{u}(\zeta + \xi)P(\zeta + \xi)|\Psi(\xi)L(d\xi) \\ = & C_1 \int_{\mathbb{C}^n} |\hat{v}(\zeta + \xi)|\Psi(\xi)L(d\xi). \end{aligned}$$

Integration of this estimate w.r.t. ζ after division by $\tilde{P}(\zeta)$ now gives

$$\begin{aligned}
|u(0)| =& \left|(2\pi)^{-n}\int \hat{u}(\zeta)L(d\zeta)\right| \leq \\
\leq& (2\pi)^{-n}C_1 \int\int |\hat{v}((\xi+\zeta)|/\tilde{P}(\zeta)L(d\zeta)\Psi(\xi)L(d\xi) \\
=& C_1 \int \|e^{-i\langle\cdot,\xi\rangle}v\|_{1,1/\tilde{P}}\Psi(\xi)L(d\xi) \leq \\
\leq& \int C_1\Psi(\xi)L(d\xi) \sup_{|\xi|<\epsilon/2} \|e^{-i\langle\cdot,\xi\rangle}v\|_{1,1/\tilde{P}}.
\end{aligned}$$

Using Theorem 3.2.2 we obtain the estimate (3.16) with

$$C = C_1 \int \Psi(\xi)d\xi \sup_{|\xi|<\epsilon/2} \|e^{-i\langle\cdot,\xi\rangle}/\cosh\epsilon|\cdot|\|_{1,M_{1/\tilde{P}}}. \tag{3.26}$$

The right-hand side is finite since the set formed by the function $x \mapsto e^{-i\langle x,\xi\rangle}/\cosh\epsilon|x|$ with $\xi \in \mathbb{C}^n$ and $|\xi| < \epsilon/2$ is bounded in $\mathscr{S}$. This complete the proof of the Lemma 3.2.10.

$\square$

3.3 An elementary approach to local solvability in $L^2(\Omega)$ (Jerison [14])

Let Ω be a bounded, open subset of $\mathbb{R}^n$. The Hilbert space $L^2(\Omega)$ with inner product $\langle f,g\rangle = \int_\Omega f(x)\bar{g(x)}dx$ and norm $\|f\| = \sqrt{\langle f,f\rangle}$ is the space in which we will find a solution to $P(D)u = f$ i.e L^2 solvability

Theorem 3.3.1. *Let $P(D)$ be a constant-coefficient partial differential operator (but not the zero operator). For any given $\phi \in L^2(\Omega)$ there exists $u \in L^2(\Omega)$ such that $P(D)u = \phi$ in the sense of distributions.*

Proof. We begin the proof by showing that it suffices to prove

$$\|f\| \leq C\|\bar{P}(D)f\|, \quad \forall f \in C_0^\infty(\Omega). \tag{3.27}$$

Indeed, given $\phi \in L^2(\Omega)$, define a linear functional u_0 on the subspace $V = \{\bar{P}(D)f; f \in C_0^\infty(\Omega)\}$ by $u_0(\bar{P}(D)f) = \langle f, \phi\rangle$. Note that u_0 is well-defined because $\bar{P}(D)$ is injective (and $\bar{P}(D) = P(D)^*$). Moreover the linear functional is bounded beacuse (3.27) implies

$$|u_0(\bar{P}(D)f)| = |\langle f, \phi\rangle| \leq \|f\|\|\phi\| \leq C\|\bar{P}(D)f\|\|\phi\|.$$

By the Hahn-Banach theorem, there is an extension of u_0 to a linear functional on $L^2(\Omega)$. We can identify this functional with an element u of $L^2(\Omega)$. Thus $\langle f, \phi\rangle = \langle \bar{P}(D)f, u\rangle$ for all $f \in C_0^\infty$, as desired. □

Plancherel's formula now implies that (3.27) is equivalent to

$$\int_{\mathbb{R}^n} |\hat{f}(y)|^2 dy \leq C^2 \int_{\mathbb{R}^n} |P(y)|^2 |\hat{f}(y)|^2 dy, \tag{3.28}$$

for all $f \in C_0^\infty(\Omega)$.

We will suppose for convenience that $\Omega \subset B_1$, where B_1 is the ball of radius 1 centered at the origin. Here is the version of the *uncertainty principle* that we need.

Theorem 3.3.2. *Let $\mathscr{F}_1$ be the family of unit cubes of $\mathbb{R}^n$ with integer lattice point corners. Let P be a polynomial of degree m and denote*

$$P_1^*(y) = \sum_{Q \in \mathscr{F}_1} \max_Q |P| \chi_Q(y),$$

where χ_Q is the characteristic function of Q.

(i) There is a constant $\sigma(P) > 0$ such that $P_1^*(y) \geq \sigma(P)$.

(ii) There is a constant $C = C(n, m)$ such that for every $f \in C_0^\infty(B_1)$,

$$\int_{\mathbb{R}^n} |\hat{f}(y)|^2 dy \leq C^2 \int_{\mathbb{R}^n} \frac{|P(y)|^2}{P_1^*(y)^2} |\hat{f}(y)|^2 dy.$$

In particular,

$$\int_{\mathbb{R}^n} |\hat{f}(y)|^2 dy \leq \frac{C}{\sigma(P)^2} \int_{\mathbb{R}^n} |P(y)|^2 |\hat{f}(y)|^2 dy.$$

Remark 1. The role played by tiling $\mathscr{F}_1$ is somewhat arbitrary. One can, for instance, rotate and traslate $\mathscr{F}_1$ without changing the result. In fact, if Q_1 and Q_2 are unit cubes whose distance apart is at most 1, then there is a constant $c = c(n, m) > 0$ such that $c \leq \max_{Q_1} |P| / \max_{Q_2} |P| \leq c^{-1}$. This follows from the fact that all norms on the finite dimensional space of polynomials of degree $\leq m$ are equivalent. Moreover, a similar argument shows that there is a constant $c = c(n, m, r) > 0$ such that

$$c \leq P_r^*(y)/P_1^*(y) \leq c^{-1}, \quad \forall y \in \mathbb{R}^n, \tag{3.29}$$

where $P_r^*(y) = \sum_{Q \in \mathscr{F}_r} \max_Q |P| \chi_Q(y)$ and $\mathscr{F}_r = \{rQ; Q \in \mathscr{F}_1, r > 0\}$.

Proof. Let Q be a unit cube. Equivalence of norm implies there is a constant $C = C(n, m)$ such that

$$\max_Q |(\partial^\alpha / \partial y^\alpha) P(y)| \leq C \max_Q |P|.$$

Therefore, $|(\partial^\alpha / \partial y^\alpha) P(y)| \leq C P_1^*(y)$. Choose a non-zero coefficient of P, a_α, of highest order $|\alpha| = m$. Then $|(\partial^\alpha / \partial y^\alpha) P(y)| = \alpha! |a_\alpha|$ is a constant independent of y, and we have the positive lower bound of part (i).

For part (ii) we observe that by Plancherel's theorem, for all $f \in C_0^\infty(B_1)$, $\|\nabla \hat{f}\|^2 = (2\pi)^n \| |x| f \|^2 \leq (2\pi)^n \|f\|^2 = \|\hat{f}\|^2$ (it is here that we use the restriction on the support of f).

Lemma 3.3.3 (Fefferman, SAK [7] p.146)**.** *Assume that $V(y)$ is a nonnegative polynomial of degree $\leq d$ on a cube Q of side-length r in $\mathbb{R}^n$. Suppose that $\max_Q V \geq r^{-2}$. There is a constant $c_1 = c_1(n, d)$ such that for all $u \in C^\infty(Q)$,*

$$\int_Q (|\nabla u(y)|^2 + V(y)|u(y)|^2) dy \geq c_1 r^{-2} \int_Q |u(y)|^2 dy$$

Proof. First of all, a change of variables $x \to rx$ shows that it suffices to consider the case $r = 1$. Thus we take the unit cube $Q = \{y \in \mathbb{R}^n : 0 \leq y_i \leq 1\}$ and V a polynomial of degree $\leq d$ such that $\max_Q V \geq 1$.

The family of functions $\phi_\alpha(y) = \prod_{j=1}^n \cos(\pi\alpha_j y_j)$ indexed by $\alpha \in \mathbb{Z}_+^n$ forms an orhogonal basis for $L^2(Q)$. If we write the series for $u \in C^\infty(Q)$ in this basis, it is easy to check that

$$\int_Q |\nabla u(y)|^2 dy \geq \pi^2 \int_Q |u(y) - u_0|^2 dy, \tag{3.30}$$

with $u_0 = \int_Q u(y)dy$, the average of u on Q. Therefore, we need only bound $|u_0|^2$ from above. To do this we will first show that $V(y) \geq 1/2$ on an wide portion of Q.

By equivalence of norms, there is a constant $C = C(d,n)$ such that $\max_Q |\nabla V| \leq C \max_Q |V| = C \max_Q V$. Then $V(y) \geq \frac{1}{2}\max_Q V \geq \frac{1}{2}$ for all $y \in B \cap Q$, where B is a ball of radius $1/2C$ centered at a point of Q at which V takes on its maximum.

Next,

$$\begin{aligned}
&\int_Q (|\nabla u(y)|^2 + V(y)|u(y)|^2)dy \\
&\qquad \geq \int_Q (\pi^2|u(y) - u_0|^2 + V(y)|u(y)|^2)dy \\
&\qquad \geq \int_Q \min(V(y), \pi^2)(|u(y) - u_0|^2 + |u(y)|^2)dy \\
&\qquad \geq \frac{1}{2}\int_Q \min(V(y), \pi^2)|u_0|^2 dy \geq \frac{1}{4}\int_{B\cap Q} |u_0|^2 dy = c_0|u_0|^2,
\end{aligned}$$

where $c_0 = Vol(B \cap Q)/4$ depends only on n and d. Combining this with (3.30), we have

$$\begin{aligned}
&\int_Q (|\nabla u(y)|^2 + V(y)|u(y)|^2)dy \\
&\qquad = (\frac{1}{2} + \frac{1}{2})\int_Q (|\nabla u(y)|^2 + V(y)|u(y)|^2)dy \\
&\qquad \geq \frac{\pi^2}{2}\int_Q |u(y) - u_0|^2 dy + \frac{c_0}{2}|u_0|^2 \\
&\qquad \geq \frac{c_0}{2}\int_Q (|u(y) - u_0|^2 + |u_0|^2)dy = \frac{c_0}{2}\int_Q |u(y)|^2 dy.
\end{aligned}$$

□

We will apply the lemma with $d = 2m$. Choose r so that $c_1 r^{-2} = 2$. Fix $Q \in \mathscr{F}_r$, and let $V(y) = |P(y)|^2 / r^2 \max_Q |P|^2$. Then $\max_{y \in Q} V(y) = r^{-2}$, and the lemma implies

$$\int_Q (|\nabla \hat{f}(y)|^2 + V(y)|\hat{f}(y)|^2) dy \geq 2 \int_Q |\hat{f}(y)|^2 dy.$$

Summing over $Q \in \mathscr{F}_r$,

$$\int_{\mathbb{R}^n} \left(|\nabla \hat{f}(y)|^2 + \frac{|P(y)|^2}{r^2 P_r^*(y)^2} |\hat{f}(y)|^2 \right) dy \geq 2 \int_{\mathbb{R}^n} |\hat{f}(y)|^2) dy.$$

But $\|\nabla \hat{f}\|^2 \leq \|\hat{f}\|^2$, so that

$$\int_{\mathbb{R}^n} \frac{|P(y)|^2}{r^2 P_r^*(y)^2} |\hat{f}(y)|^2 dy \geq \int_{\mathbb{R}^n} |\hat{f}(y)|^2) dy.$$

Finally, since r depends only on n and m, (3.29) implies

$$\frac{P_1^*(y)^2}{r^2 P_r^*(y)^2} \leq C = C(n, m),$$

and hence (ii). $\square$

Chapter 4

Construction of a parametrix for elliptic operators

4.1 The process of "inverting" an elliptic differential operator

A differential operator is called *elliptic* if its principal symbol does not vanish when $\xi \neq 0$. We will now describe the process of "inverting" an elliptic differential operator.

Let us assume first that P has constant coefficients. Since $p_m(\xi) \neq 0$ on the unit sphere $\mathbb{S}^{n-1} = \{\xi; |\xi| = 1\}$, there is a constant $C_1 > 0$ such that $p_m(\xi) \geq C_1$ on $\mathbb{S}^{n-1}$ and, by homogeneity, $|p_m(\xi)| \geq C_1|\xi|^m$ on $\mathbb{R}^n$. On the other hand, there is another constant, C_2, such that $|p(\xi) - p_m(\xi)| \leq C_2|\xi|^{m-1}$ if $|\xi|$ is large enough, since $p(\xi) - p_m(\xi)$ is a polynomial of degree $m-1$. Then

$$\begin{aligned} |p(\xi)| \geq & |p_m(\xi)| - |p(\xi) - p_m(\xi)| \\ \geq & C_1|\xi|^m - C_2|\xi|^{m-1} \\ = & |\xi|^m (C_1 - \frac{C_2}{|\xi|}), \end{aligned}$$

so $|p(\xi)| > 0$ if $|\xi|$ is large enough. Let then $\phi \in C_0^\infty(\mathbb{R}^n)$ be a function which is equal to 1 in a neighborhood of the zeros of $p(\xi)$ and define $q(\xi) =$

$(1-\phi)/p$. Then q is a smooth, tempered function with the property that $pq = 1-\phi$,which is a function equal to 1 in the complement of a compact set. Define the operator Q by

$$Qu(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} q(\xi)u(y)dyd\xi, \quad u \in \mathscr{S}(\mathbb{R}^n).$$

It is then immediate that $Q : \mathscr{S}(\mathbb{R}^n) \longrightarrow \mathscr{S}(\mathbb{R}^n)$ continuously. Moreover,

$$\begin{aligned} PQu(x) &= (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} p(\xi)q(\xi)u(y)dyd\xi \\ &= 2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} (1-\phi(\xi))u(y)dyd\xi \\ &= u(x) - Ru(x), \end{aligned}$$

with $Ru(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi}\phi(\xi)u(y)dyd\xi$. Since $\phi \in C_0^\infty(\mathbb{R}^n)$, the operator R maps $\mathscr{S}'(\mathbb{R}^n)$ into $\mathscr{S}(\mathbb{R}^n)$ and it is continuous. We call such an operator a *smoothing* operator. Thus P is *invertible* modulo smoothing operators.

Let us now consider the case where P does not have constant coefficients. Let us assume that the coefficients are defined on all of $\mathbb{R}^n$ and that they are bounded along with and all their derivatives by all orders. Let us also assume P is elliptic, i.e. $\exists C > 0$ such that $|p_m(x,\xi)| > C|\xi|^m$ for all (x,ξ) with $\xi \neq 0$. By our assumptions, one sees that the zeros of $p(x,\xi)$ are contained in a set $\mathbb{R}^n \times \Omega$ where Ω is a bounded neighborhood of 0. Choose $\phi \in C_0^\infty(\mathbb{R}^n)$ with $\phi = 1$ near Ω and define $q_{-m}(x,\xi) = (1-\phi)/p(x,\xi)$. The operator Q_{-m} defined by

$$Q_{-m}u(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} q_{-m}(x,\xi)u(y)dyd\xi,$$

(integrate first in y, then in ξ) is a continuous operator from $\mathscr{S}(\mathbb{R}^n)$ to $C^\infty(\mathbb{R}^n)$. But now we do not have $PQ_{-m} = I + R$ with R smoothing.

Rather,

$$PQ_{-m}u(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} p(x,\xi) q_{-m}(x,\xi) u(y) dy d\xi$$

$$+(2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} \Big(\sum_{|\alpha|\le m} a_\alpha(x) \sum_{\beta<\alpha} \binom{\alpha}{\beta} D_x^{\alpha-\beta} q_{-m}(x,\xi)\xi^\beta \Big) u(y) dy d\xi, \tag{4.1}$$

using the Leibnitz rule. Since $pq_{-m} = 1+\phi$, the first expression on the right equals $u+Ru$, where R is an error like the one obtained earlier in the constant coefficient case. But the second term (absent in the constant coefficient case) is not of that form. Call it $R_{-1}u(x)$, and let $r_{-1}(x,\xi)$ be the expression in bracket in that integral, so that

$$R_{-1}u(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} r_{-1}(x,\xi) u(y) dy d\xi.$$

Note that for large $|\xi|$,the term $q_{-m}(x,\xi)$ and its x-derivatives are controlled by $|\xi|^{-m}$, so that $r_{-1}(x,\xi)$ is controlled by $|\xi|^{-1}$. We also note that (4.1) is true regardless q_m. So, if we set $q_{-m-1} = -r_{-1}q_{-m}$ and define

$$Q_{-m-1}u(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} q_{-m-1}(x,\xi) u(y) dy d\xi,$$

we have

$$PQ_{-m-1}u(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} p(x,\xi) q_{-m-1}(x,\xi) u(y) dy d\xi$$

$$+(2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} \Big(\sum_{|\alpha|\le m} a_\alpha(x) \sum_{\beta<\alpha} \binom{\alpha}{\beta} D_x^{\alpha-\beta} q_{-m-1}(x,\xi)\xi^\beta \Big) u(y) dy d\xi.$$

The first expression equals $-R_{-1}$ plus a smoothing operator, and the second one can be written as

$$R_{-2}u(x) = (2\pi)^{-n} \iint_{\mathbb{R}^{2n}} e^{i(x-y)\cdot\xi} r_{-2}(x,\xi) u(y) dy d\xi,$$

so that $P(Q_{-m} + Q_{-m-1}) = I + R_{-2}$+smoothing. The symbol $r_{-2}(x,\xi)$ is controlled by $|\xi|^{-2}$ and q_{-m-1} by $|\xi|^{-m-1}$. We repeat the process, with the purpose of getting rid of R_{-2}, defining $q_{-m-2} = -r_{-2}q_{-m}$ and the corresponding operator. In this way we get sequences of operators Q_{-m-j}, R_{-1-j}, respectively defined by symbols of order $m-j$ and $1-j$, with the property that

$$P(Q_{-m} + Q_{-m-1} + ... + Q_{-m-N}) = I + R_{-N-1} + \text{smoothing}.$$

The regularity of the Schwartz kernel of R_{-N} increases with N, so if you could add all the Q_{-m-j}, so as to be able to define $Q = \sum_j Q_{-m-j}$ in such a way that the operators $\sum_{j>N} Q_{-m-j}$ have kernels with increasing regularity, we would have $PQ - I = P(\sum_{j\leq N} Q_{-m-j} + \sum_{j>N} Q_{-m-j}) = R_{-N-1} + P(\sum_{j>N} Q_{-m-j})$, which is an operator whose kernel has arbitrarily high regularity, that is, is smoothing.

All of the above can actually be done in a suitable way. One cannot directly add the Q_{-m-j} but there is a perfectly good substitute for that.

4.2 Parametrix for elliptic operators

Definition 4.2.1. Let $X \subset \mathbb{R}^n$ be an open set, if $a \in S^{m'}(X \times \mathbb{R}^n)$, $b \in S^{m''}(X \times \mathbb{R}^n)$, we can define $a\sharp b \in S^{m'+m''}(X \times \mathbb{R}^n)$ uniquely up to some element of $S^{-\infty}(X \times \mathbb{R}^n)$ by

$$(a\sharp b)(x,\xi) \sim \sum \frac{1}{\alpha!}\partial_\xi^\alpha a(x,\xi) D_x^\alpha b(x,\xi),$$

where $\sim$ means the *asymptotic sum.*

This gives a biliner map

$$S^{m'}/S^{-\infty} \times S^{m''}/S^{-\infty} \ni (a,b) \mapsto a\sharp b \in S^{m'+m''}/S^{-\infty}, \tag{4.2}$$

and the "product" $\sharp$ is associative.

Theorem 4.2.2. *If $P \in L^m(X)$ is elliptic, then there exists $Q \in L^{-m}(X)$, properly supported, such that $P \circ Q \equiv Q \circ P \equiv I \mod L^{-\infty}(X)$. Moreover Q is unique modulo $L^{-\infty}(X)$.*

Proof. Using a partition of unity we can first find a function $q_0 \in C^\infty(X \times \mathbb{R}^n)$ such that for every compact $K \subset X$, there is a constant $C_K > 0$ such that

$$q_0(x,\xi) = \frac{1}{p(x,\xi)} \quad \text{for} \quad x \in K, |\xi| \geq C_K.$$

Lemma 4.2.3. $q_0(x,\xi) \in S^{-m}(X \times \mathbb{R}^n)$.

We have $p\sharp q_0 = 1 - r$, $q_0 \sharp p = 1 - t$ with $r, t \in S^{-1}/S^{-\infty}$. Define

$$q_r = q_0 \sharp (1 + r + r\sharp r + r\sharp r\sharp r + ...) \in S^m/S^{-\infty}$$
$$q_l = (1 + t + t\sharp t + t\sharp t\sharp t + ...)\sharp q_0 \in S^m/S^{-\infty}.$$

Then $p\sharp q_r = 1$, $q_l \sharp p = 1$ (in $S^0/S^{-\infty}$), and, furthermore $q_l \sharp (p \sharp q_r) = (q_l \sharp p)\sharp q_r = q_r$.

Let $Q \in L^m(X)$ be properly supported with symbol $q_l = q_r \mod S^{-\infty}$. Then $P \circ Q \equiv Q \circ P \equiv I \mod L^{-\infty}(X)$. If $Q' \in L^m(X)$ is a second operator with the same properties, we get $P \circ (Q - Q') \equiv 0$ and composing with Q to the left gives $Q - Q' \equiv 0$.

□

The operator Q is called a *parametrix.*

Corollary 4.2.4. *Let A be an elliptic differential operator with smooth coefficients on an open set $X \subset \mathbb{R}^n$ and let $x_0 \in X$. Then there exists an open neighborhood $V \subset X$ of x_0 such that for every $v \in \mathscr{D}'(V)$ and every open $W \Subset V$, there exists $u \in \mathscr{D}'(V)$ such that $Au = v$ in W.*

Proof. For every compact $K \subset X$,and every $s \in \mathbb{R}$, there exists $C = C_{K,s} > 0$ such that

$$\|u\|_{s+m} \leq C(\|A^* u\|_s + \|u\|_s), \quad \forall u \in \mathscr{E}'(K) \cap H^{s+m}(\mathbb{R}^n).$$

In fact, let $B \in L^{-m}(X)$ be a properly supported parametrix of A^*. Then $u = BA^*u + Ru$ where $R \in L^{-\infty}(X)$ is properly supported and both B and R are continuous $H^s_{\text{comp}}(X) \longrightarrow H^{s+m}_{\text{comp}}(X)$.

We assume $m \geq 1$ (the case $m = 0$ is trivial). By Poincaré's lemma we know that for every $\epsilon > 0$ we have $\|u\|_0 \leq \|u\|_m$ for all $u \in \mathscr{E}'(K) \cap H^m(\mathbb{R}^n)$, provided the diameter of the support of u is sufficiently small depending on ϵ and m only. Hence, if V is a sufficiently small open neighborhood of x_0, we have in addition to the previous estimate with $s = 0$, that $C\|u\|_0 \leq \frac{1}{2}\|u\|_m$ and hence

$$\|u\|_m \leq 2C\|A^*u\|_0, \quad \forall u \in \mathscr{E}'(K) \cap H^m(\mathbb{R}^n).$$

If $v \in \mathscr{D}'(V)$, we first put $\tilde{u} = \tilde{B}v$, where $\tilde{B} \in L^{-m}(V)$ is a properly supported parametrix of A. Then $A\tilde{u} = v + \tilde{v}$ where $\tilde{v} \in C^\infty(V)$, and the problem of local solvability is reduced to the case when $v \in C^\infty(V)$. For such a v, we let $W \Subset V$ be open and consider the linear form

$$\ell : H^m(\mathbb{R}^n) \cap \mathscr{E}'(W) \ni \phi \mapsto \langle \phi | v \rangle \in \mathbb{C}.$$

Then $|\ell(\phi)| \leq C(v, W)\|\phi\|_m \leq \tilde{C}(v, W)\|A^*\phi\|_0$. Hence $\ell(\phi) = k(A^*\phi)$, where k is a bounded linear form on $L = \{A^*\phi \in L^2 \cap \mathscr{E}'(W); \phi \in H^m(\mathbb{R}^n) \cap \mathscr{E}'(W)\}$.

By the Hahn-Banach theorem, k has a bounded extension to $L^2(\mathbb{R}^n)$, whence there exists $u \in L^2(\mathbb{R}^n)$ such that

$$k(A^*\phi) = \langle A^*\phi, u \rangle, \quad \forall \phi \in L^2 \cap \mathscr{E}'(W),$$

and therefore $Au = v$ in W. □

Chapter 5

Hypoelliptic operators

5.1 Hypoellipticity and local solvability

Let P be a *properly supported* pseudodifferential operator of order m, and let $X \subset \mathbb{R}^n$ be open.

Definition 5.1.1. The operator P is *hypoelliptic* if

$$\text{sing supp } u = \text{sing supp } Pu, \quad \forall u \in \mathscr{D}'(X).$$

Equivalently: $\forall Y \subset X$, Y open,

$$u \in \mathscr{D}'(X), \quad Pu \in C^\infty(Y) \Longrightarrow u \in C^\infty(Y). \tag{5.1}$$

Set now

$$H^s(K) := H^s(\mathbb{R}^n) \cap \mathscr{E}'(K),$$

$K \subset X$ a compact. Then

$$H^s(K) \text{ is closed in } H^s(\mathbb{R}^n).$$

Theorem 5.1.2. *Suppose that P is a differential operator, such that*

$$u \in \mathscr{E}'(K), \quad Pu \in C^\infty \Longrightarrow u \in C_0^\infty(K). \tag{5.2}$$

Then P^* is locally solvable at any $x_0 \in \mathrm{int}(K)$.

Proof. Define, for a given $s \in \mathbb{R}$,

$$F := \{u \in H^s(K); Pu \in C^\infty\},$$

endowed with the family of seminorms

$$\|u\|_r^F := \|u\|_s + \|Pu\|_r, \quad r \in \mathbb{Z}_+.$$

We claim that F is a Fréchet space. In fact, suppose $\|u_k - u_k\|_r^F \to 0$ as $k, k' \to +\infty$ for all $r \in \mathbb{Z}_+$. Then $\|u_k - u_{k'}\|_s \to 0$, whence $u_k \to u$ in $H^s(K)$. But then $Pu_k \to Pu$ in $\mathscr{E}'(K)$ as $k \to \infty$, and since $\|Pu_k - Pu_{k'}\|_r \to 0$ as $k, k' \to +\infty$ for all $r \in \mathbb{Z}_+$, for each $r \in \mathbb{Z}_+$ there exists $v_r \in H^r(K)$ such that $\|Pu_k - v_r\|_r \to 0$ as $k \to \infty$. Hence also $Pu_k \to v_r$ in $\mathscr{E}'(K)$ as $k \to \infty$, so that $Pu \in H^r(K)$ for all $r \in \mathbb{Z}_+$, which proves the claim.

Now, the inclusion $j : C_0^\infty(K) \longrightarrow F$ is continuous ($C_0^\infty(K)$ endowed with the family of H^s-norms), injective and onto by virtue of the hypothesis (5.1). Hence, by the open mapping theorem the inverse is continuous, whence: *For any given $s, k \in \mathbb{R}$ there exists $r \in \mathbb{Z}_+$ and a constant $C = C(K, s, k) > 0$, such that*

$$\|u\|_k \leq C(\|Pu\|_r + \|u\|_s), \quad \forall u \in C_0^\infty(K).$$

Now fix $k = 1$, $s = 0$. It follows that for some $t > 0$

$$\|u\|_1 \leq C(\|Pu\|_t + \|u\|_0), \quad \forall u \in C_0^\infty(K).$$

By the Poincaré inequality, for any given $\epsilon > 0$ we may find a relatively compact neighborhood $V \subset K$ of $x_0 \in \mathrm{int}(K)$ with diameter sufficiently small, so as to have

$$\|u\|_0 \leq \epsilon\|\nabla u\|_0 \leq \epsilon(\|u\|_0^2 + \|\nabla u\|_0^2)^{1/2} = \epsilon\|u\|_1, \quad \forall u \in C_0^\infty(V),$$

whence (for a new constant $C > 0$)

$$\|u\|_1 \leq C\|Pu\|_t, \quad \forall u \in C_0^\infty(V).$$

If necessary, we increase $t > 0$ so that

$$\|u\|_{-t} \leq \|u\|_1 \leq C\|Pu\|_t, \quad \forall u \in C_0^\infty(V).$$

Consider then

$$M := \overline{\{P\varphi; \varphi \in C_0^\infty(V)\}}^{H^t(\mathbb{R}^n)} \subset H^t(\mathbb{R}^n).$$

Given any $f \in C^\infty(X)$, we consider $f|_{\overline{V}}$ and extend it as a compactly supported smooth function (with support contained in some relatively compact subset of X that contains the compact $\overline{V}$) so that we may suppose also that $f \in H^t(\mathbb{R}^n)$. So, consider the linear form

$$L : P\varphi \longmapsto \langle \bar{f}|\varphi\rangle, \quad \varphi \in C_0^\infty(V).$$

Since

$$|L(P\varphi)| \leq \|f\|_t \|\varphi\|_{-t} \leq \|f\|_t \|P\varphi\|_t,$$

L can be extended as a continuous linear form $L : M \longrightarrow \mathbb{C}$, and it can therefore be further extended to a continuous linear form to the whole $H^t(\mathbb{R}^n)$, whence the existence of $u \in H^{-t}(\mathbb{R}^n)$ such that

$$\langle \bar{f}|\varphi\rangle = L(P\varphi) = \langle \bar{u}|P\varphi\rangle = \langle \overline{P^* u}|\varphi\rangle, \quad \forall \varphi \in C_0^\infty(V), \tag{5.3}$$

that is $P^* u = f$ (in the distribution sense) in V.

□

5.2 An example of an unsolvable hypoelliptic operator

We have seen in Theorem 5.1.1. that the formal adjoint of an hypoelliptic differential operator is locally solvable. This does not imply that the operator itself is locally solvable. In this section we will show an example of a hypoelliptic second order differential operator in two variables for which there exists a line such that the operator is not solvable at any point of this line.

The differential operator is

$$A = D_1 + i x_1 D_2^2, \tag{5.4}$$

where $D_j = \frac{1}{i}(\partial/\partial x_j)$, $j = 1, 2$.

5.2.1 Proof of the hypoellipticity of A

The nature of A becomes more trasparent if the coordinates $x_1^2/2 = t$, $x_2 = x$ are used. Then

$$A = -i\sqrt{2t}\Big(\frac{\partial}{\partial t} + \frac{\partial^2}{\partial x^2}\Big), \tag{5.5}$$

and it follows that on each of the half planes $\{x_1 > 0\}$ and $\{x_1 < 0\}$, A is equal to a C^∞ function times the backward heat operator (on each half plane separately). Thus, A is certainly hypoelliptic in the complement of the line $x_1 = 0$. So, let u be a solution of the equation

$$Au = f, \tag{5.6}$$

where $u \in \mathscr{D}'(\Omega)$, $f \in C^\infty(\Omega)$ and Ω is an open subset of the plane. We hence know that $u \in C^\infty(\Omega \cap \{(x_1, x_2); x_1 \neq 0\})$. Hence, we have only to show that u is infinitely differentiable also in the neighborhood of the line $x_1 = 0$. We may assume, therefore, that Ω is an open disk whose center lies on the line $x_1 = 0$, and since A is invariant under translations in the direction of the x_2 variable we may assume that the center of Ω is at the origin. It suffices to prove that u is infinitely differentiable in a neighborhood V of the origin, where $\overline{V} \subset \Omega$. Let $\phi \in C_0^\infty(\Omega)$ be identically equal to 1 in a neighborhood V_1 of $\overline{V}$. Then

$$\begin{aligned} A\phi u &= \phi A u + (D_1\phi + ix_1 D_2^2\phi)u + 2ix_1(D_2\phi)(D_2 u) \\ &= g + (A\phi)u + 2ix_1(D_2\phi)(D_2 u), \end{aligned}$$

(where $g = \phi f$). Let us first introduce new coordinates $-(x_1^2/2) = t$, $x_2 = x$ for the open half plane $\{(x_1, x_2); x_1 > 0\}$. Then

$$\Big(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}\Big)\phi u = \frac{-ig}{\sqrt{-2t}} - \frac{-i}{\sqrt{-2t}}(A\phi)u + 2(D_2\phi)(D_2 u). \tag{5.7}$$

The function $(\phi u)(\sqrt{-2t}, x)$ vanishes identically on the line $t = -K$ if K is sufficiently large and is bounded in the set $\{(t, x); -\infty < x < +\infty, -K \leq$

$t \leq -\epsilon\}$ for every $\epsilon > 0$. Hence, we may use the usual fundamental solution of the heat equation (5.6) and conclude that

$$\begin{aligned}(\phi u)(\sqrt{-2t}, x) &= \frac{H(t)}{\sqrt{4\pi t}} e^{-x^2/4t} * \Big(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}\Big)\phi u \\ &= \int_{-\infty}^{t} \frac{1}{\sqrt{4\pi(t-\tau)}} \int e^{-(x-y)^2/4(t-\tau)} \Big(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}\Big)\phi u(\sqrt{-2\tau}, y)\,dyd\tau,\end{aligned}$$

for negative value of t. Using (5.6), we see that

$$\begin{aligned}(\phi u)(\sqrt{-2t}, x) = -i \int_{-\infty}^{t} &\frac{1}{\sqrt{4\pi(t-\tau)}} \int e^{-(x-y)^2/4(t-\tau)} \frac{g(\sqrt{-2\tau}, y)}{\sqrt{-2\tau}} dyd\tau \\ &+ \int_{-\infty}^{t} \frac{1}{\sqrt{4\pi(t-\tau)}} \int e^{-(x-y)^2/4(t-\tau)} \Big(\frac{-i((A\phi)u)(\sqrt{-2\tau}, y)}{\sqrt{-2\tau}} \\ &+ 2(D_2\phi)(D_2 u)\Big) dyd\tau.\end{aligned}$$

Changing back to the original coordinates (x_1, x_2) (where $x_1 > 0$) we find that

$$(\phi u)(x_1, x_2) = v(x_1, x_2) + w(x_1, x_2), \tag{5.8}$$

where

$$v(x_1, x_2) = -i \int_{x_1}^{\infty} \frac{1}{\sqrt{2\pi(y_1^2 - x_1^2)}} \int e^{-(x_2-y_2)^2/2(y_1^2-x_1^2)} g(y_1, y_2)\,dy_2 dy_1, \tag{5.9}$$

and

$$\begin{aligned}w(x_1, x_2) = \int_{x_1}^{\infty} &\frac{1}{\sqrt{2\pi(y_1^2 - x_1^2)}} \int e^{-(x_2-y_2)^2/2(y_1^2-x_1^2)} \times \\ &\times(-i((A\phi)u)(y_1, y_2) + 2y_1((D_2\phi)(D_2 u))(y_1, y_2))\,dy_2 dy_1.\end{aligned} \tag{5.10}$$

Taking the partial Fourier transform of v we get

$$\hat{v}(x_1, \xi) = -i \int_{x_1}^{\infty} e^{-(y_1^2 - x_1^2)\xi^2/2} \hat{g}(y_1, \xi)\,dy_1. \tag{5.11}$$

Since $g = \phi f$ has compact support, the integration is actually perfomed on a finite interval (of length at most equal to $\sqrt{2K}$). Moreover, for every positive number N there exists a constant C_N such that $|\hat{g}(x_1,\xi)| = |\widehat{\phi f}(x_1,\xi)| \leq C_N(1+|\xi|)^{-N}$ since $\phi f \in C_0^\infty(\Omega)$. Hence

$$|\hat{v}(x_1,\xi)| = \int_{x_1}^{\infty} |\hat{g}(y_1,\xi)| dy_1 \leq \sqrt{2K} C_N (1+|\xi|)^{-N}. \tag{5.12}$$

It follows that $v(x_1,x_2)$ is infinitely differentiable w.r.t. x_2, and each of the derivatives $D_2^k v$ is uniformly bounded as $x_1 \to 0^+$. Noting that (5.10) or (5.8) imply that

$$D_1 v = g - i x_1 D_2^2 v, \tag{5.13}$$

we see that the function $D_1 v$ along with each of its derivatives w.r.t. x_2 are bounded as $x_1 \to 0^+$. Differentiating (5.12) w.r.t. x_1 we find that $D_1^2 v = D_1 g - i x_1 D_2^2 v - D_2^2 v$ is uniformly bounded as $x_1 \to 0^+$, and that the same holds for each of its derivatives w.r.t. x_2. Iteration of this procedure leads us to the conclusion that each of the derivatives of v is uniformly bounded as $x_1 \to 0^+$ (and therefore v has in fact an infinitely differentiable extension to the closed half plane $\{(x_1,x_2); x_1 \geq 0\}$).

Turning now our attention to $w(x_1,x_2)$, we note in the first place that $\mathrm{dist}(\mathrm{supp}\nabla\phi, V) > 0$, since $\phi \equiv 1$ on the set V_1 which contains $\overline{V}$ in its interior. Moreover, the functions $A\phi$ and $D_2\phi$ have compact support, and the fundamental solution $E(x,t)$ defined by the equations

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

is infinitely differentiable except at the point $x = t = 0$. Hence, the functions $\psi(y_1,y_2)$ defined by

$$\psi(y_1,y_2,x_1,x_2) = \begin{cases} (A\phi)(y_1,y_2) E\big(x_2 - y_2, \frac{y_1^2 - x_1^2}{2}\big) & y_1 \geq 0 \\ 0 & y_1 \leq 0 \end{cases}$$

are in fact test function in $C_0^\infty(\Omega)$ (of the variables y_1,y_2) and depend in an infinitely differentiable manner (as vector valued functions of y_1,y_2 with values in $C_0^\infty(\Omega)$) on x_1 and x_2, where $(x_1,x_2) \in V$ and $x_1 \geq 0$. Since u

is a distribution and thus is continuous on $C_0^\infty(\Omega)$ it follows that the scalar function

$$\int_{x_1}^{\infty}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi(y_1^2-x_1^2)}}\int e^{-(x_2-y_2)^2/2(y_1^2-x_1^2)}((A\phi)(y_1,y_2)u(y_1,y_2))dy_2dy_1$$
$$= u(E(x_2-\cdot,\frac{\cdot^2-x_1^2}{2})(A\phi)(\cdot))$$

is infinitely differentiable w.r.t. x_1 and to x_2 in the intersection of V with the closed half plane $\{(x_1,x_2);x_1\geq 0\}$. Since D_2u is also a distribution in $\mathscr{D}'(\Omega)$ we may treat the second term in (5.9) in a similar way and conclude that $w(x_1,x_2)$ is infinitely differentiable in the closed half plane $\{(x_1,x_2);x_1\geq 0\}$. Using (5.9) we thus see that the function $u(x_1,x_2)$ and every derivative of it are uniformly bounded (in V) as $x_1\to 0^+$.

In a similar fashion, let us define $v(x_1,x_2)$ and $w(x_1,x_2)$ for $x_1<0$ by

$$v(x_1,x_2)=i\int_{-\infty}^{x_1}\frac{1}{\sqrt{2\pi(y_1^2-x_1^2)}}\int e^{-(x_2-y_2)^2/2(y_1^2-x_1^2)}g(y_1,y_2)dy_2dy_1, \quad (5.14)$$

and

$$w(x_1,x_2)=\int_{\infty}^{x_1}\frac{1}{\sqrt{2\pi(y_1^2-x_1^2)}}\int e^{-(x_2-y_2)^2/2(y_1^2-x_1^2)}\times \quad (5.15)$$
$$\times(i((A\phi)u)(y_1,y_2)-2y_1((D_2\phi)(D_2u))(y_1,y_2))dy_2dy_1.$$

It follows once again that $\phi u=v+w$ for $x_1<0$ and that the function $u(x_1,x_2)$ as well as each of its derivatives are bounded as $x_1\to 0^-$.

We have proved that the function $u(x_1,x_2)$ possesses C^∞ boundary values as x_1 tends to zero from either the right or the left. In order to finish the proof that $u\in C^\infty(V)$ one has to show that the boundary values of u and its derivatives actually match up and that u has no singular part with support on the line $x_1=0$. For details look at [15].

5.3 Proof that A is not locally solvable on the line $x_1 = 0$

We show that the operator A does not satisfies the Hörmander's condition: $p = 0 \Longrightarrow \{p, \bar{p}\} = 0$.
We have

$$p = \xi_1 + ix_1\xi_2^2 \quad \bar{p} = \xi_1 - ix_1\xi_2^2.$$

Now the points $(0, x_2; 0, \xi_2) \in \mathrm{Char}(P)$, but

$$\{p, \bar{p}\} = -2i\xi_2^2$$

does not vanish at these points. Hence A is not solvable on $\{x_1 = 0\}$.

5.4 Hypoelliptic operators with loss of derivatives

Definition 5.4.1. The operator P properly supported of order m is said to be hypoelliptic with a loss of r of derivatives, $r \geq 0$, if

$$u \in \mathscr{D}'(X), \quad Pu \in H^s_{\mathrm{loc}} \Longrightarrow u \in H^{s+m-r}_{\mathrm{loc}}. \tag{5.16}$$

Note that (5.16) measures the extent to which P fails to be elliptic, i.e. hypoelliptic with loss of 0 derivatives. One can also see that (5.16) implies (5.1).

When P is hypoelliptic with a certain loss of $r \geq 0$ derivatives, the a priori estimate is easier. In fact, one has the following lemma (Hörmander [13] Lemma 22.4.2, Vol.III).

Lemma 5.4.2. *Let P be properly supported and of order m. Suppose that for some $r \geq 0$, whatever $s \in \mathbb{R}$,*

$$u \in \mathscr{E}'(X), \quad Pu \in H^s \Longrightarrow u \in H^{s+m-r}.$$

Then for every compact $K \subset X$ and for every $s \in \mathbb{R}$ there exists $C_{K,s} > 0$ such that

$$(HE) \quad \|u\|_{s+m-r} \leq C_{K,s}(\|Pu\|_s + \|u\|_{s+m-r-1}), \quad \forall u \in C_0^\infty(K).$$

Proof. Set

$$F := \{u \in H^{s+m-r-1}(K); Pu \in H^s\},$$

with norm

$$\|u\|_F := \|u\|_{s+m-r-1} + \|Pu\|_s.$$

Then $(F, \|\cdot\|_F)$ is a Banach space. By (HE) we have that F is embedded into $H^{s+m-r}(K)$ and the embedding is a closed operator. In fact, take a sequence $\{u_k\}_k \subset F$ such that $u_k \longrightarrow u$ in F and $u_k \longrightarrow v$ in $H^{s+m-r}(K)$. Then $u = v$, i.e. $u \in H^{s+m-r}(K)$ and since $\{Pu_k\}_k$ is a Cauchy sequence in H^s, we get $Pu_k \longrightarrow v_0$ in H^s. Thus, since $Pu_k \longrightarrow Pu$ in $\mathscr{E}'(X)$, we also get $v_0 = Pu \in H^s$. Hence $u \in F$.

Finally, by the closed graph theorem, we obtain that the sought for inequality holds, and this conclude the proof. □

It is worth noting that inequality (HE) with $r \geq 1$ is not sufficient in order to have hypoellipticity. Let us consider in fact, $P = \partial_t^2 - \triangle_x$, the wave-operator. We have the following lemma.

Lemma 5.4.3. *For any given compact $K \subset \mathbb{R}_t \times \mathbb{R}_x^n$ and any given $s \in \mathbb{R}$ there exists a constant $C = C(K, s) > 0$ such that (HE) holds for $r = 1$. However, because of propagation of singularities, P is not hypoelliptic (not even with a loss of* 1 *derivative).*

Proof. Since the operators $(1 + |D_t|^2 + |D_x|^2)^s$ all commute with P, it suffices to prove the inequality when $s = 0$. Let then K be a compact of $\mathbb{R}^{n+1}$. We may hence suppose that for $T, R > 0$ we have $K \subset [-T, T] \times \overline{D_R(0)}$. Let for $u \in C_0^\infty(K)$

$$E(t) := \frac{1}{2} \int_{\mathbb{R}^n} \big(|\partial_t u(t, x)|^2 + |\nabla_x u(t, x)|^2\big) dx.$$

We may suppose, P being real, that u is real-valued. Moreover, we have that for $|t| > T$ the energy $E(t) = 0$. One computes

$$\frac{d}{dt}E(t) = \int_{\mathbb{R}^n} \big(\partial_t^2 u(t,x)\partial_t u(t,x) + \nabla_x \partial_t u(t,x) \cdot \nabla_x \partial_t u(t,x)\big)dx = \\ = \int_{\mathbb{R}^n} \partial_t u(t,x)(Pu)(t,x)dx.$$

We thus have, for some $\lambda \in \mathbb{R}$ to be picked,

$$(e^{\lambda t}\partial_t u, Pu) = \int_{-T}^{T} e^{\lambda t}\frac{dE}{dt}(t)dt = -\lambda \int_{-T}^{T} e^{\lambda t}E(t)dt + [e^{\lambda t}E(t)]_{t=-T}^{t=T} = \\ = -\lambda \int_{-T}^{T} e^{\lambda t}E(t)dt.$$

Choose then $\lambda = -1$, whence

$$\int_{-T}^{T} e^{-t}E(t)dt = (e^{-t}\partial_t u, Pu).$$

Since $E(t) \geq 0$ we have on the one hand

$$\int_{-T}^{T} e^{-t}E(t)dt \geq e^{-T}\int_{-T}^{T} E(t)dt = \frac{e^{-T}}{2}\big(\|\partial_t u\|^2 + \|\nabla_x u\|^2\big),$$

and on the other

$$|\int\int_{\mathbb{R}^{1+n}} e^{-t}\partial_t u(t,x)(Pu)(t,x)dtdx| \leq \frac{\epsilon e^{T}}{2}\|\partial_t u\|^2 + \frac{e^{T}}{2\epsilon}\|Pu\|^2,$$

where $\epsilon > 0$ is to be picked. Therefore

$$\frac{1}{2}(e^{-T} - \epsilon e^{T})\|\partial_t u\|^2 + \frac{e^{-T}}{2}\|\nabla_x u\|^2 \leq \frac{e^{T}}{2\epsilon}\|Pu\|^2.$$

We choose $\epsilon = e^{-2T}/2$. Hence

$$\frac{e^{-T}}{4}\|\partial_t u\|^2 + \frac{e^{-T}}{4}\|\nabla_x u\|^2 \leq \frac{e^{-T}}{4}\|\partial_t u\|^2 + \frac{e^{-T}}{2}\|\nabla_x u\|^2 \leq e^{3T}\|Pu\|^2.$$

Thus

$$\|u\|_1^2 \leq 4e^{4T}\big(\|Pu\|_0^2 + \|u\|_0^2\big), \tag{5.17}$$

with space-time Sobolev norms. □

For details about hypoellipticity with loss of derivatives one can see [29], [13], [24] (and for a recent work on hypoellipticity with a big loss of derivatives see [16], [25], [27]).

Chapter 6

Invariance of condition (Ψ) and a proof of local solvability in two dimension under condition (Ψ)

6.1 Flow invariant sets and the invariance of condition (Ψ)

> (Ψ) For every elliptic homogeneous function q, the function $\mathsf{Im}(qp)$ does not change sign from - to + along any given oriented maximal integral curve of $H_{\mathsf{Re}(qp)}$ in $U \times \mathbb{R}^n$ passing through $\mathrm{Char}(P)$.

A surprising feature of condition (Ψ) is that it involves the bicharacteristics of $\mathsf{Re}(qp)$ although they depend very much on q except where H_p is proportional to a real vector. In spite of this it was shown by Nirenberg and Treves [23] that the choice of q is not very important in condition (Ψ). The main point in their proof is the application of results on flow invariant sets due to Bony [2] and Brézis [3].

Let X be a C^2 manifold, $F \subset X$ be a closed subset and ν be a Lipschitz continuous vector field in X. We want to describe the conditions on ν required for integral curves starting in F to remain in F for all later times.

First note that if $x_0 \in F$ and $f \in C^1$, $f(x_0) = 0$ and $f \leq 0$ in F, then we must have $\nu f(x_0) \leq 0$. In fact, let γ be an integral curve of ν s.t. $\gamma(0) = x_0$, then $\frac{d}{dt}\big|_{t=0} f(\gamma(t)) = (\nu f)(x_0)$, if $\nu f(x_0) \geq 0$ then $f(\gamma(t)) > f(x_0) = 0$ on the right of x_0, a contradiction since $f \leq 0$ on F.

Definition 6.1.1. We define $N(F)$ as the set of all $(x, \xi) \in T^*(X)\backslash 0$ s.t. one can find $f \in C^1$ with $f(x) = 0$, $df(x) = \xi$ and $f \leq 0$ in a neighborhood of x in F. Note that necessarily $x \in \partial F$.

Theorem 6.1.2 (Bony [2])**.** *Let ν be a Lipschitz continuous vector field in X. Then the following conditions are equivalent:*

(i) Every integral curve $x(t)$, $0 \leq t \leq T$, of ν with $x(0) \in F$ is contained in F.

(ii) $\langle \nu(x), \xi \rangle \leq 0$ for all $(x, \xi) \in N(F)$.

We have already proved that $(i) \implies (ii)$. Since the statement is local, in proving the converse we may assume that $X = \mathbb{R}^n$. We need the following lemma.

Lemma 6.1.3. *Let F be a closed set in $\mathbb{R}^n$ and set*

$$f(x) = \min_{z \in F} |x - z|^2,$$

where $|\cdot|$ is the Euclidean norm. Then we have

$$\begin{aligned} &f(x + y) = f(x) + g(x, y) + o(|y|), \quad \text{where} \\ &g(x, y) = \min\{\langle 2y, x - z \rangle; z \in F, |x - z|^2 = f(x)\}. \end{aligned}$$

Proof. We may assume in the proof that $x = 0$. Set, for $\epsilon > 0$,

$$q_\epsilon(y) = \min\{-2\langle y, z \rangle; z \in F, |z| \leq \sqrt{f(0)} + \epsilon\}.$$

Then q_ϵ is a homogeneous function of degree 1, and $q_\epsilon \uparrow q_0$ as $\epsilon \downarrow 0$. The limit is therefore uniform on the unit sphere (by Dini's theorem on uniform convergence), so that

$$|q_\epsilon(y/|y|) - q_0(y/|y|)| \le c_\epsilon$$

gives

$$q_0(y) \ge q_\epsilon(y) \ge q_o(y) - c_\epsilon|y|, \quad c_\epsilon \to 0, \quad \text{as} \quad \epsilon \to 0.$$

Now $|y-z|^2 = |z|^2 - 2\langle y, z\rangle + |y|^2$ whence

$$f(y) \le f(0) + q_0(y) + |y|^2.$$

On the other hand, when $|y| \le \epsilon$ the minimum in the definition of $f(y)$ is assumed for some z with $|z| \le \sqrt{f(0)} + \epsilon$, hence

$$f(y) \ge f(0) + q_\epsilon(y) + |y|^2, \quad |y| \le \epsilon,$$

which proves the lemma. □

Proof of Theorem 6.1.2. With the notation in (i) and Lemma 6.1.3 we have if $t < T$

$$\lim_{s \to t^+} \frac{f(x(s)) - f(x(t))}{s-t} = g(x(t), \nu(x(t))).$$

Since the result to be proved is local we may assume that for all x, y

$$|\nu(x) - \nu(y)| \le C|x-y|.$$

When $z \in F$ and $|x(t) - z|^2 = f(x(t))$ we have

$$2\langle \nu(x(t)), x(t) - z\rangle = 2\langle \nu(z), x(t) - z\rangle - 2\langle \nu(z) - \nu(x(t)), x(t) - z\rangle.$$

The last term in absolute value is controlled by $2Cf(x(t))$. Since $f(x(t)) - |x(t) - \tilde{z}|^2 \le 0$ for all $\tilde{z} \in F$, we have $(z, x(t) - z) \in N(F)$ if $x(t) \ne z$, so the first term on the right is ≤ 0 by condition (ii). Hence the right-derivative of $f(x(t))$ is $\le 2Cf(x(t))$ so that $f(x(t))e^{-2Ct} \le 0$. Hence $f(x(t))e^{-2Ct}$ is decreasing in every interval where it is positive, and if $f(x(0)) = 0$ it then follows that $f(x(t)) = 0$ for $0 \le t \le T$. □

Corollary 6.1.4 (Brézis [3]). *Let $q \in C^1(X)$ where X is a C^2 manifold and let ν be a Lipschitz continuous vector field in X such that for every integral curve $t \longmapsto x(t)$ of ν we have*

$$q(x(0)) < 0 \Longrightarrow q(x(t)) \leq 0 \quad \text{for} \quad t > 0. \tag{6.1}$$

Let w be another C^1 vector field such that

$$\langle w, \nabla q \rangle \leq 0 \quad \textit{when} \quad q = 0 \tag{6.2}$$

$$w = \nu \quad \textit{when} \quad q = dq = 0. \tag{6.3}$$

Then (6.1) *remains valid if $x(t)$ is replaced by any given integral curve of w.*

Proof. Let F be the closure of the union of all forward orbits of ν starting at points x with $q(x) < 0$. By (6.1) we have $q \leq 0$ in F, and F contains the closure of the set where $q < 0$. Therefore orbits of ν which start in F must remain in F. If now $(x, \xi) \in N(F)$, then x is in the boundary of F so $q(x) = 0$. If $dq(x) \neq 0$ then F is bounded by the surface $q = 0$ in a neighborhood of x, and ξ is a positive multiple of $dq(x)$, thus $\langle w(x), \xi \rangle \leq 0$ by (6.2). If $dq(x) = 0$ we have $\langle w(x), \xi \rangle = \langle \nu(x), \xi \rangle \leq 0$ by (6.3) since ν satisfies condition (ii) in Theorem 6.1.2. Hence w satisfies condition (ii) in Theorem 6.1.2 and therefore condition (i) also, which proves the corollary. $\square$

Lemma 6.1.5. *Let I be a point or a compact interval on $\mathbb{R}$, and let $\gamma : I \longrightarrow M$ be an embedding of I in a sympletic manifold M as a one dimensional bicharacteristic of $p_1 + ip_2$ if I is not reduced to a point, and any characteristic point otherwise. Let*

$$f_j = \sum_{k=1}^{2} a_{jk} p_k, \quad j = 1, 2,$$

where $\det(a_{jk}) > 0$ on $\gamma(I)$. Assume that $H_{p_1} \neq 0$ and that $H_{f_1} \neq 0$ on $\gamma(I)$. If $\gamma(I)$ has a neighborhood U such that p_2 does not change sign from $-$ to $+$ along any bicharacteristic for p_1 in U, then U can be so chosen that f_2 has no such sign change along the bicharacteristics of f_1 in U.

Proof. First note that if $p_1 = p_2 = 0$ at a point in U then

$$\{p_1, p_2\} = H_{p_1} p_2 \le 0,$$

Hence, at the same point,

$$\{f_1, f_2\} = \{a_{11}p_1 + a_{12}p_2, a_{21}p_1 + a_{22}p_2\} = (a_{11}a_{22} - a_{12}a_{21})\{p_1, p_2\} \le 0.$$

The proof is now divided into two steps.

(i) Assume first that $a_{12} = 0$. Since $a_{11}a_{22} > 0$ either a_{11} and a_{22} are both positive or both negative. Thus the bicharacteristics of $f_1 = a_{11}p_1$ are equal to those of p_1 with preserved and reserved orientation respectively, and $f_2 = a_{22}p_2$ when $p_1 = 0$ so f_2 has the same and opposite sign, as p_2, respectively. This proves the lemma in this case.

(ii) By a canonical change of variables we can make $M = \mathbb{R}^{2n}$, $p_1 = \xi_1$ and $\Gamma = \gamma(I)$ equal to an interval on the x_1 axis (Darboux's theorem). Let $T \in \mathbb{R}^{2n}$ be a vector with

$$\langle T, dp_1 \rangle = 1 \quad \text{and} \quad \langle T, df_1 \rangle \ne 0 \quad \text{on} \quad \Gamma.$$

Since dp_1 and df_1 do not vanish on Γ, the existence of T is obvious if Γ consists of a single point. Otherwise dp_2 is proportional to dp_1 (since this is the case for H_{p_2} and H_{p_1}) on Γ so df_1 is proportional to dp_1. Hence we just have to take $T = (\underbrace{0, ..., 0}_{n}, \underbrace{1, 0, ..., 0}_{n})$.

Set

$$q_2(x, \xi) = p_2((x, \xi) - \xi_1 T).$$

Then $p_2 = q_2$ when $\xi_1 = 0$ (i.e. $p_1 = 0$) and q_2 is constant in the direction T. Then there is a smooth function φ such that

$$q_2 = \varphi p_1 + p_2.$$

It therefore follows from step (i) that the hypotheses in the lemma are fulfilled for $p_1 + iq_2$. We have

$$f_1 = (a_{11} - a_{12}\varphi)p_1 + a_{12}q_2,$$

hence

$$0 \neq \langle T, df_1 \rangle = (a_{11} - a_{12}\varphi) \quad \text{on} \quad \Gamma.$$

In a neighborhood of Γ we can therefore divide f_1 by $a_{11} - a_{12}\varphi$ and set

$$q_1 = f_1/(a_{11} - a_{12}\varphi) = p_1 + a_{12}(a_{11} - a_{12}\varphi)^{-1} q_2,$$

which implies

$$f_j = \sum_{k=1}^{2} b_{jk} q_k, \quad j = 1, 2,$$

where $b_{11} = a_{11} - a_{12}\varphi$, $b_{12} = 0$ and $\det b = \det a > 0$. Thus it follows from step (i) that it suffices to prove that (q_1, q_2) satisfies the hypothesis made on (p_1, p_2) in the lemma. The difficulty here is that the surfaces $p_1 = 0$ and $q_1 = 0$ are not the same. However, they may be identified by projecting in the direction T.

Let U be a neighborhood of Γ where q_2 does not change sign from $-$ to $+$ on the bicharacteristics of p_1. Since T is transversal to the surface $f_1 = q_1 = 0$, we can choose U so small that $Y = \{(x, \xi) \in U; q_1(x, \xi) = 0\}$ is mapped diffeomorphically by the projection π_T along the direction T to $X = \{(x, \xi) \in U; \xi_1 = 0\}$. When $q_1 = q_2 = 0$, i.e. $p_1 = p_2 = 0$, we have $H_{q_1} q_2 = H_{p_1} p_2 \leq 0$, so (6.2) are fulfilled in Y by the restriction q of q_2 to Y and $w = H_{q_1}$. At a point in Y where $q = 0$ and dq vanishes on the tangent space of Y, we have $dq_2 = 0$ since $\langle T, dq_2 \rangle = 0$. Hence $w = H_{q_1} = H_{p_1}$ there and thus $\pi_{T*} w = H_{p_1}$. If we apply Corollary 6.1.4 to $q = \pi_T^* q_2$ and the vector fields $\nu = \pi_{T*}(H_{p_1})$ and w we conclude that q_2 cannot change sign from $-$ to $+$ along a bichacteristic of q_1 in Y. This completes the proof. $\square$

6.2 Proof of local solvability in two dimension under condition (Ψ)

In n dimensions, the principal symbol of P can be microlocally conjugated to the form ([13] thm. 21.3.6.)

$$\xi_1 + iq(x, \xi'),$$

where $q(x, \xi')$ is real, positively homogeneous of degree one, and $\xi' = (\xi_2, ..., \xi_n)$. Condition (Ψ), in this case, is equivalent to the following: *For each x', ξ', the real function $q(x_1, x', \xi')$ nowhere changes sign from $-$ to $+$ as x_1 increases.* The special feature of two dimensions is that the positively homogeneous (of degree 1) function $q(x, \xi_2)$ can be written for $\xi_2 > 0$ as $b(x)\xi_2$, where $b(x) = q(x, 1)$ is a function of x only. Local solvability is then implied by the following energy estimate (6.4), which is Theorem 1.2.3 of [17]. Note that in the energy estimate (6.4) the symbol $\xi_1 + iq$ satisfies condition $(\bar{\psi})$ (i.e. q does not change from $+$ to $-$ as x_1 increases). The estimate implies a local solvability result for an operator with principal symbol $\xi_1 - iq$, the lower order terms being unimportant by virtue of of the large constant T^{-1}.

Theorem 6.2.1. *Suppose that $b(t, x) \in C^\infty([-1, 1] \times \mathbb{R})$ nowhere changes sign from $+$ to $-$ as t increases. Then there exist $C, T_0 > 0$ such that for all $T \leq T_0$,*

$$\|(\partial_t - b(t, x)|D_x|)u\|_2 \geq C^{-1}T^{-1}\|u\|_2, \tag{6.4}$$

for all Schwartz function u on $\mathbb{R}^2$ such that $u(t, x) = 0$ when $|t| \geq T$.

Proof. Consider the vector field on $\mathbb{R}^3$,

$$L = \partial_t + b(t, x)\partial_y,$$

and the smooth extension of $u(t, x)$ from $\mathbb{R}^2$ to $\mathbb{R}^2 \times \{y \geq 0\}$,

$$f(t, x, y) = e^{-y|D_x|}u(t, x).$$

The sign-change condition on b is equivalent to saying that, for each (t, x), the integral curve $(t + s, x, \int_t^{t+s} b(r, x)dr)$ of L through $(t, x, 0)$ remains in the region $y \geq 0$ for either all $s \geq 0$ or all $s \leq 0$: the past of a point (t, x) such $b(t, x) \leq 0$ is made of points $(t' = t + s, x)$, $s \leq 0$, with $b(t', x) \leq 0$, analogously for the future with $b(t, x) \geq 0$, this gives that the integral curves remains in the region $y \geq 0$. Since $f(t, x, y) = 0$ for $|t| \geq T$, we can express $u(t, x) = f(t, x, 0)$ as the integral of Lf along the integral curve of L through $(t, x, 0)$, to either $\pm T$, the choice depending on (t, x). This yields for all (t, x) the inequality

$$|u(t,x)| \leq \int_{-T}^{+T} \sup_{0<y<M} |(Lf)(s,x,y)| ds,$$

where M depends on the bounds for b.

Now observe that $Lf\big|_{y=0} = (\partial_t - b(t,x)|D_x|)u$ and that $\triangle_{x,y}(Lf) = [\triangle_{x,y}, Ll]f + L\underbrace{\triangle_{x,y}f}_{=0}$, (due to $\widehat{f}(t,\xi,y) = e^{-y|\xi|}\widehat{u}(t,\xi)$ and $-\partial_y^2\widehat{f}(t,\xi,y) = |\xi|^2\widehat{f}(t,\xi,y)$.) We thus write

$$Lf = e^{-y|D_x|}(\partial_t - b(t,x)|D_x|)u + G_{x,y}[\triangle_{x,y}, l]f,$$

where $G_{x,y}$ is the Green's kernel in x, y. Since $[\triangle_{x,y}, L]$ is of second order and involves derivatives in x and y only, the second term depends continuously on u in the L^2 norm. If T is small enough, the integral of this term along L will have a small L^2 norm compared to u.

Directly from the definition of f we have

$$(Lf)(t,x,y) = e^{-y|D_x|}(\partial_t - b(t,x)|D_x|)u(t,x) - \left[b(t,x), e^{-y|D_x|}\right]|D_x|u(t,x) \tag{6.5}$$

and we seek to control the L^2 norm of the supremum over y.

Now we want to recall some properties of the *Poisson kernel* (P_{x_n}) and the *maximal function* (M). (See, ref Stein Singular Integrals and Differentiability Properties of Function). Let

$$u(x', x_n) = P_{x_n} * f(x'), \quad x' \in \mathbb{R}^{n-1}, \quad x_n > 0, \quad f \in L^p(\mathbb{R}^{n-1})$$

and

$$M(f)(x) = \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f(y) dy.$$

We have

$$\sup_{x_n>0} |u(x', x_n)| \leq M(f)(x').$$

Furthermore, note that

$$M : L^p(\mathbb{R}^{n-1}) \longrightarrow L^p(\mathbb{R}^{n-1}), \quad \forall p \in (1, \infty], \quad \text{isboundend},$$

and that

$$\mathcal{F}_{x'\to\xi}(P_{x_n})(\xi) = e^{-x_n|\xi|}.$$

Then we obtain that the supremum over y of the first term in (6.5) is dominated by the *maximal function* in x of $(\partial_t - b(t,x)|D_x|)u(t,x)$ and thus has an L^2 norm bounded by $\|(\partial_t - b(t,x)|D_x|)u(t,x)\|_2$.

For the commutator term we fix a cutoff function $\varphi \in C_0^\infty(\mathbb{R})$ such that

$$\varphi(\xi) = \begin{cases} 1 & |\xi| \leq 1 \\ 0 & |\xi| \geq 2 \end{cases} \quad , 0 \leq \varphi \leq 1.$$

Thus we can write the multiplier $e^{-y|\xi|}$ as a compactly supported term, plus a term supported in $|\xi| > 1$. The compactly supported term, $e^{-y|\xi|}\varphi(\xi)$, yelds $R_1 u(t,x,y)$ such that both $\|R_1 u(\cdot, y)\|_{L^2(x,t)}$ and $\|\partial_y R_1 u(\cdot, y)\|_{L^2()x,t}$ are bounded by $\|u\|_2$ uniformly in y. The supremum over $0 < y < M$ thus has an L^2 norm bounded by $\|u\|_2$.

The term $a_1(\xi) = e^{-y|\xi|}(1 - \varphi(\xi)) \in S^0$. In fact, to fix ideas, with no loss of generality, consider $\xi > 0$. We have

$$\partial_\xi a_1(\xi) = -y(1 - \varphi(\xi))e^{-y|\xi|} - e^{-y|\xi|}\varphi'(\xi),$$

the second term is supported in $1 \leq \xi \leq 2$ and the first in $\xi \geq 1$, so that

$$\partial_\xi a_1(\xi) = \frac{-y\xi(1 - \varphi(\xi))e^{-y|\xi|}}{|\xi|} + S^{-\infty}. \tag{6.6}$$

Note that

$$|y\xi e^{-y|\xi|}| \leq C \quad \forall y \geq 0, \tag{6.7}$$

and

$$\partial_\xi^k e^{-y|\xi|}(1-\varphi(\xi)) = (-y)^k e^{-y|\xi|}(1-\varphi(\xi)) + S^{-\infty} = \tag{6.8}$$

$$\frac{(-y|\xi|)^k e^{-y|\xi|}}{|\xi|^k}(1-\varphi(\xi)) + S^{-\infty}, \tag{6.9}$$

where the first term in (6.9) is $\leq C'|\xi|^{-k}$.
Hence the $Kohn-Nirenberg$ formula for pseudodifferential operator in x applies to the commutator with the remainder estimates uniform over y and t:

$$\sigma([b,a_1]) \sim \frac{1}{i}\{b,a_1\} + S^{-2}, \quad \{b,a_1\} = i\frac{\partial b}{\partial x}\frac{\partial a_1}{\partial \xi} = \tag{6.10}$$

$$= i\frac{\partial b}{\partial x}\Big(-\frac{y\xi}{|\xi|}\Big)(1-\varphi(\xi))e^{-y|\xi|} + S^{-\infty} = \tag{6.11}$$

$$= -i\frac{\partial b}{\partial x} y e^{-y|\xi|}\frac{\xi}{|\xi|} + S^{-\infty}. \tag{6.12}$$

The leading term combined with the $|D_x|$ equals,

$$\begin{aligned} &i\frac{\partial b}{\partial x}(-y\xi)e^{-y|\xi|} + (\text{compactly supported}) = \\ &= i\frac{\partial b}{\partial x}\Big(y\frac{\xi}{|\xi|}\Big)\frac{\partial}{\partial y}e^{-y|\xi|} + (\text{compactly supported}) = \\ &= i\frac{\partial b}{\partial x} y\,\mathrm{sgn}(\xi)\frac{\partial}{\partial y}e^{-y|\xi|} + (\text{compactly supported}), \end{aligned}$$

then, modulo a term compactly supported in ξ which behaves as R_1 above, we get

$$(\partial_x b) y \int \mathrm{sgn}(\xi)\partial_y e^{-y|\xi|} e^{ix\xi}\widehat{u}(t,\xi)d\xi$$

Noting that $\mathrm{sgn}(\xi)\widehat{u}(t,\xi)$ is the fourier transform of $H_x u$, we obtain

$$(\partial_x b) y \partial_y e^{-y|D_x|} H_x u(t,x)$$

where H_x is the Hilbert transform in x.
The maximal function for the kernel $y\partial_y e^{-y|D_x|}$ is bounded on L^2 by the same argument as for the Poisson kernel since, as for the Poisson kernel, we have

$$y\partial_y P_y(x/\epsilon)\Big|_{y=1} = \epsilon y \partial_y P_y(x)\Big|_{y=\epsilon} \quad \forall \epsilon > 0.$$

The second-order remainder term $R_2(y, t, x, |D_x|)$ (that is, the term related to the S^{-2} in (6.10)) is a pseudodifferential operator in x of order -2, with its bounds uniform in y and t. Furthermore $\partial_y R_2(y, t, x, |D_x|)$ is the remainder term for the commutator of $b(t, x)$ and the cutoff of $\partial_y e^{-y|D_x|}$; hence it is uniformly of order -1. We thus have similar estimates for $R_2|D_x|u$ as for $R_1 u$.

The first inequality now yelds

$$\|u\|_2 \leq C_1 T \|(\partial_t - b(t, x)|D_x|)u(t, x)\|_2 + C_2 T \|u\|_2,$$

for T sufficiently small. $\square$

Appendix A

Locally convex topological vector spaces

A.1 Some topological vector spaces

A topological vector space is a vector space together with a Hausdorff topology with respect to which the algebraic operations are continuous. We collect here some definitions and results concerning such vector spaces.

A seminorm on a complex vector space X is a function $\wp : X \to \mathbb{R}$ satisfying:

(1) $\wp(x) \geq 0 \ \forall x \in X$,

(2) $\wp(x) = 0$ if $x = 0$,

(3) $\wp(x + y) \leq \wp(x) + \wp(y) \ \forall x, y \in X$,

(4) $\wp(tx) = |t|\wp(x) \ \forall x \in X, t \in \mathbb{C}$.

Example A.1.1. Relevant examples of seminorms are as follows. Let $K \Subset \mathbb{R}^n$ be a compact subset and N be a nonnegative integer. For $f \in C^\infty(\mathbb{R}^n)$ let

$$\wp_{N,K}(f) = \sup_{x \in K, |\alpha| \leq N} |D_x^\alpha f(x)| \tag{A.1}$$

A family $\mathscr{P}$ of seminorms on X defines a topology on X by way of declaring a set $U \subset X$ to be open if for every $x_0 \in U$ there are $\wp_1, ..., \wp_k \in \mathscr{P}$ and positive numbers $r_1, ..., r_k$ such that

$$\{x \in X; \wp_j(x - x_0) < r_j, j = 1, ..., k\} \subset U.$$

The family $\mathscr{P}$ is *separating* if for every $x \in X$, $x \neq 0$, there is $\wp \in \mathscr{P}$ such that $\wp(x) \neq 0$. If the family $\mathscr{P}$ is separating then X with the induced topology is a Hausdorff topological vector space admitting a local base of neighborhood at 0 consisting of convex open sets. Such a space is a locally convex topological vector space.

Let $N = \{y \in X; \wp(y) = 0 \quad \forall \wp \in \mathscr{P}\}$. Then N is a subspace of X and the topology defined by $\mathscr{P}$ is Hausdorff if and only if $N = 0$. The space N is in any case closed, and for any $x \in X$, $\wp(x + y) = \wp(x)$ for every $y \in N$. So the seminorms $\wp$ determine seminorms on the quotient X/N giving a Hausdorff topology.

Example A.1.2. Recall that convergence of a sequence $\{f_k\}_{k=1}^{\infty}$ in $C^\infty(\mathbb{R}^n)$ to a function $f \in C^\infty$ means that for any given α, $D_x^\alpha f_k \to D_x^\alpha f$ uniformly on compact subsets of $\mathbb{R}^n$. This notion of convergence is precisely the convergence in the topology defined by the seminorms $\wp_{N,K}$.

The seminorms may in fact be norms and still produce an interesting topology. Two examples are the following:

Example A.1.3. Let $K \Subset \mathbb{R}^n$ be compact. Then $C_0^\infty(K)$ is, by definition the space

$$\{f \in C^\infty(\mathbb{R}^n); \quad \mathrm{supp} f \subset K\},$$

is a closed subset of $C^\infty(\mathbb{R}^n)$. The seminorms $\wp_{K,N}$ of $C^\infty(\mathbb{R}^n)$ become norms on $C_0^\infty(K)$.

Example A.1.4. Recall that the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ is the subspace of $C^\infty(\mathbb{R}^n)$ whose elements have the property that $x^\beta D_x^\alpha f$ is bounded for each α, β. Define

$$\wp_{\alpha,\beta}(f) = \sup_{x \in \mathbb{R}^n} |x^\beta D_x^\alpha f(x)|.$$

Then $\wp_{\alpha,\beta}$ is a norm on $\mathscr{S}(\mathbb{R}^n)$. The collection of these seminorms for $\alpha, \beta \in \mathbb{Z}_+^n$ defines the standard topology of $\mathscr{S}(\mathbb{R}^n)$.

A topological vector spaces is metrizable if its topology is determined by a metric. If the topology of X is Hausdorff defined by countably many seminorms then it is metrizable. For if $\{\wp_k\}_{k=1}^{\infty}$ is an enumeration of the seminorms, then

$$d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{\wp_k(x - y)}{1 + \wp_k(x - y)}$$

is a metric defining the same topology as the seminorms. Evidently this metric is invariant under translations.

Example A.1.5. The topology of $C^\infty(\mathbb{R}^n)$ is metrizable. Indeed, the topology $C^\infty(\mathbb{R}^n)$ is determined by the seminorms $\{\wp_{\bar{B}(0,N),N}; N \in \mathbb{N}\}$, where $\bar{B}(0, N)$ is the closed ball of radius N with center 0. Likewise, if $K \Subset \mathbb{R}^n$, then the topology of $C_0^\infty(K)$ is metrizable, as is that of $\mathscr{S}(\mathbb{R}^n)$.

A.2 Complete topological spaces

In order to define completeness of a topological vector space we need to recall the notion of net. First, a directed set is a set $\mathcal{D}$ together with an order relation $\leq$ such that

1. for all $\nu \in \mathcal{D}$, $\nu \leq \nu$;
2. for all $\eta, \mu, \nu \in \mathcal{D}$, $\eta \leq \mu$ and $\mu \leq \nu$ implies $\eta \leq \nu$;
3. for all $\eta, \mu \in \mathcal{D}$ there is $\nu \in \mathcal{D}$ such that $\eta \leq \nu$ and $\mu \leq \nu$.

For example, the natural numbers with the usual order relation form a directed set. A net in a set X is a family $\{x_\nu\}_{\nu\in\mathcal{D}}$ indexed by a directed set. If X is a topological vector space, then $\{x_\nu\}_{\nu\in\mathcal{D}}$ is a Cauchy net if for every neighborhood U of 0 there is $\nu_0 \in \mathcal{D}$ such that $\nu_0 \leq \mu, \nu$ implies $x_\mu - x_\nu \in U$. When $\mathcal{D} = \mathbb{N}$ and the topology is defined by a norm this coincides with the notion of Cauchy sequence.

A topological vector space X is complete if every Cauchy net converges. If X is metrizable, then completeness needs to be checked only for sequences.

A *Fréchet space* is a locally convex complete metrizable topological vector space. Examples of such are Banach spaces, as well as the spaces $C^\infty(\mathbb{R}^n)$, $C_0^\infty(K)$ if $K \Subset \mathbb{R}^n$, and $\mathscr{S}(\mathbb{R}^n)$.

Let X, Y be topological vector spaces and Γ be a family of continuous linear maps $X \to Y$. The *orbit* of $x \in X$ by Γ is the set

$$\Gamma(x) = \{\gamma(x) \in Y; \gamma \in \Gamma\}$$

Theorem A.2.1 (Banach-Steinhaus, [28])**.** *Let X, Y be topological vector spaces and Γ be a family of continuous linear maps $X \to Y$. Let*

$$B = \{x; \Gamma(x) \text{ is bounded}\}.$$

If B is of second category (i.e. it isn't a countable union of nowhere dense sets), then $B = X$ and Γ is equicontinuous.

This the basic ingredient for proving the following result

Theorem A.2.2. *Suppose X is a Fréchet space, Y is a metrizable topological vector space and Z is a topological vector spaces. If $\beta : X \times Y \longrightarrow Z$ is separately continuous bilinear form, then it is continuous.*

For details see [28], [32].

Bibliography

[1] M. F. Atiyah. Resolution of singularities and division of distributions, Comm. Pure Appl. Math. 23 1970, 145-150.

[2] J. M. Bony. Principe du maximum, inégalité de Harnack et unicité du probléme de Cauchy pour les opérateurs elliptiques dégenerés. Ann. Inst. Fourier. Grenoble 19 (1969),277-304.

[3] H. Brézis. On a characterization of flow-invariant sets. Comm. Pure Appl. Math. 23 (1970), 261-263.

[4] R. Beals and C. Fefferman. On local solvability of linear partial differential equations. Ann. of Math. (2) 97 (1973), 482-498.

[5] N. Dencker. The resolution of the Nirenberg-Treves conjecture. Ann. of Math. (2) 163 (2006), 405-444.

[6] A. Grigis and J. Sjöstrand. Microlocal Analysis for Differential Operators. An Introduction. Cambridge University Press (1994).

[7] C. Fefferman. The uncertainty principle. Bull. Amer. Math. Soc. (N.S.) 9 (1983), no. 2, 129-206.

[8] L. Hörmander. Differential operators of principal type. Math. Ann. 140 (1960), 124-146.

[9] L. Hörmander. On existence of solutions of partial differential equations. In Partial differential equations and continuum mechanics. Univ. of Wisconsin Press, Madison, Wis. (1961) 233-240 .

[10] L. Hörmander. On the range of convolution operators, Ann. of Math. 76 (1962), 148-170

[11] L. Hörmander. Linear partial differential operators. Third revised printing. Die Grundlehren der mathematischen Wissenschaften, Band 116 Springer-Verlag New York Inc.. New York 1969.

[12] L. Hörmander. Pseudodifferential operators of principal type. In Singularities in boundary value problems (Proc. NATO Adv. Study Inst., Maratea, 1980) 69-96, NATO Adv. Study Inst. Ser. C: Math. Phys. Sci., 65, Reidel, Dordrecht-Boston, Mass., 1981.

[13] L. Hörmander. The analysis of linear partial differential operators. III. Pseudo-differential operators. Reprint of the 1994 edition. Classics in Mathematics. Springer, Berlin (2007).

[14] D. Jerison. An Elementary Approach to Local Solvability for Constant Coefficient Partial Differential Equations. Forum Mathematicum 2 (1990), 45-50.

[15] Y. Kannai. An unsolvable hypoelliptic differential operator. Israel J. Math. 9 1971 306-315.

[16] J. J. Kohn. Hypoellipticity and loss of derivatives. Ann. of Math. (2) 162 (2005), 943-986

[17] N. Lerner. Sufficiency of condition (Ψ) for local solvability in two dimensions. Ann. of Math. (2) 128 (1988), 243-258.

[18] N. Lerner. Cutting the loss of derivatives for solvability under condition (Ψ). Bull. Soc. Math. France 134 (2006),4, 559-631.

[19] N. Lerner. Metrics on the phase space and non-selfadjoint pseudo-differential operators. Pseudo-Differential Operators. Theory and Applications, 3. Birkhäuser Verlag, Basel, (2010).

[20] H. Lewy. An example of a smooth linear partial differential equation without solution. Ann. of Math. (2) 66 (1957), 155-158.

[21] B. Malgrange. Existence et approximation des solutions des équations aux dérivées partielles et des équations de convolution. Ann. Inst. Fourier Grenoble, 6 (1955-1956), 271-355.

[22] L. Nirenberg and F. Treves. On local solvability of linear partial differential equations, I: Necessary conditions. Comm. Pure Appl. Math. 23 (1970), 1-38

[23] L. Nirenberg and F. Treves. On local solvability of linear partial differential equations, II: Sufficient conditions. Comm. Pure Appl. Math. 23 (1970), 459-509 with correction in Comm.Pure Appl. Math 24 (1971), 279-288.

[24] C. Parenti. The hypoellipticity problem. (Italian) Boll. Un. Mat. Ital. A (5) 15 (1978), no. 2, 300-326.

[25] C. Parenti and A. Parmeggiani. On the hypoellipticity with a big loss of derivatives. Kyushu J. Math. 59 (2005), no. 1, 155-230.

[26] C. Parenti and A. Parmeggiani. Algebra lineare ed equazioni differenziali ordinarie. Springer-Verlag Italia (2010).

[27] A. Parmeggiani. Hypoellipticity with loss of many derivatives. (Italian) "Bruno Pini" Mathematical Analysis Seminar, University of Bologna Department of Mathematics: Academic Year 2008/2009 (Italian), 61-80, Tecnoprint, Bologna, 2009.

[28] W. Rudin. Functional Analysis. Second edition. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York. (1990).

[29] J. Sjöstrand.Parametrices for pseudodifferential operators with multiple characteristics. Ark. Mat. 12, 85-130. (1974).

[30] H. F. Smith. An Elemntary Proof of Local Solvability in two Dimensions Under Condition (Ψ). Ann. of Math. (2) 136, 335-337.

[31] E. M. Stein. Singular Integrals and Differentiability Properties of Functions. Princeton Uni. Press (1970).

[32] F. Treves. Topological Vector Spaces, Distributions and Kernels. Academic Press (1967).

[33] F. Treves. Introduction to Pseudodifferential and Fourier Integral Operators, I. Plenum Press (1980).

Ringraziamenti

Ringrazio il relatore della mia tesi Alberto Parmeggiani per avermi guidato, consigliato e stimolato in questi anni di Laurea Magistrale, lo ringrazio per tutto il tempo dedicatomi e per la passione che è riuscito a trasmettermi verso le tematiche che ho cercato di analizzare in questo lavoro.
Ringrazio la mia famiglia per essermi stata sempre accanto e avermi sempre supportato. Vorrei inoltre ringraziare Giulia per la pazienza con cui mi è stata vicina in questi anni, per aver sempre capito le mie difficoltà e avermi sostenuto nelle mie scelte.