Developments in Mathematics

David S. Tartakoff

Nonelliptic Partial Differential Equations

Analytic Hypoellipticity and the Courage to Localize High Powers of T

Springer

Nonelliptic Partial Differential Equations

Developments in Mathematics

VOLUME 22

For further volumes:
http://www.springer.com/series/5834

David S. Tartakoff

Nonelliptic Partial Differential Equations

Analytic Hypoellipticity and the Courage to Localize High Powers of T

Springer

David S. Tartakoff
S Morgan St 851
60607-7042 Chicago Illinois
USA
dst@uic.edu

ISSN 1389-2177
ISBN 978-1-4419-9812-5 e-ISBN 978-1-4419-9813-2
DOI 10.1007/978-1-4419-9813-2
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011931713

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Chapter 1
What This Book Is and Is Not

The question of high smoothness of solutions to partial differential equations, in particular the equations studied here and called the "$\bar{\partial}$-Neumann problem" and the complex boundary Laplacian, lie at the interface between real and complex analysis and has deep repercussions in both.

The C^∞ regularity results were established in 1963 by J.J. Kohn [K1] and shortly afterward by others [KN], [Hö1] and only fifteen years later did the local real analytic regularity find resolution, independently by F. Treves [Tr4] and by the present author in [T4], [T5]. G. Métivier generalized the results in 1980 following Treves [Mé2], and then in 1983 J. Sjöstrand [Sj2] used a still different technique, that of the so-called FBI transform, to re-prove these results. These proofs are radically different from one another: Treves' constructs a parametrix for a so-called Grušin operator and then treats the general case as a perturbation, mine explicitly and directly localizes a high power of a vector field and then successively and naturally corrects the commutation errors that arrive in using L^2 estimates, and that of Sjöstrand uses the powerful but somewhat rigid FBI transform, a clever variant of the Fourier transform that permits one to avoid localization directly. See also the work of Okaji [Ok] from 1985. We will comment on Treves' approach in more detail below.

It is the main purpose of this book, however, to familiarize the reader with the technique and constructions that I have developed, which, while utterly elementary in their essence, require a certain amount of time for their exposition, and I felt that a longer format, such as a book, would provide the matrix for this narrative.

All three methods alluded to above have stood the test of time. The one presented here not only is elementary in nature, but also seems to be the most flexible and open to perturbations. And it may be approached through the simple example of sums of squares of real vector fields of the most elementary (nontrivial) sort.

And while the technique is elementary in nature—it uses nothing beyond a good first-year graduate course in analysis—it does not replace that course.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_1,

Thus in order to facilitate readability for those who know, or have heard of, this technique, I have chosen not to start off with the definitions of Lebesgue measure and integration, the Fourier transform, normed vector spaces, Sobolev spaces and the Sobolev embedding theorem, left-invariant vector fields on the Heisenberg group, and some elementary theory of pseudodifferential operators, but have included a kind of "reference section" in the appendix, which contains the necessary definitions and basic results to which the reader may wish to refer from time to time.

However, for those who want to assess right away their preparedness for the main text, here are some of the facts that are developed further in the Appendix. If you are comfortable with these results and are content to proceed without proofs at this point, by all means continue to the next chapter.

- The Fourier transform $\hat{f}(\xi)$ of a function $f(x)$ is an isometry of $L^2(\mathbb{R}^n)$ and takes $\partial f/\partial x_j$ to $(1/i)\xi_j \hat{f}$.
- The Sobolev space H^s consists of all functions (tempered distributions) f for which $(1+|\xi|^2)^{s/2}\hat{f} \in L^2$. Thus if s is a nonnegative integer, $f \in H^s$ if and only if $\partial^\alpha f/\partial x^\alpha \in L^2$ for all multi-indices α with $|\alpha| \le s$.
- (the Sobolev embedding theorem) $f(x)$ is continuous if $f \in H^s$ for some $s > n/2$. It follows that $f \in C^\infty$, provided $f \in H^s\ \forall s$.
- Distributions on $\mathbb{R}^n$, denoted by $\mathcal{D}(\mathbb{R}^n)$, are elements of the topological dual space to $C_0^\infty(\mathbb{R}^n)$. They may be differentiated and multiplied by smooth functions. Every distribution belongs locally to some H^s.
- The vector fields

  $$X_j = \frac{\partial}{\partial x_j} - \frac{y_j}{2}\frac{\partial}{\partial t}, \quad Y_j = \frac{\partial}{\partial y_j} + \frac{x_j}{2}\frac{\partial}{\partial t}$$

  are the so-called left-invariant vector fields on the Heisenberg group. They satisfy

  $$\begin{aligned} &[X_j, Y_k] = \delta_{j,k} T, \\ &[X_j, X_k] = [Y_j, Y_k] = [X_j, T] = [Y_j, T] = 0 \end{aligned}$$

  (We will use nothing else about the Heisenberg group; suffice it to say that these vector fields play the same role vis-à-vis the Heisenberg group (they commute with left translation in the group) that the coordinate partial derivatives do vis-à-vis the usual Euclidean group structure in $\mathbb{R}^n$.) They will provide the simplest model for the vector fields we will study.
- The simplest pseudodifferential operators, as introduced by K.O. Friedrichs in the 1960s, provide the algebraic tool needed to invert many partial differential operators modulo (infinitely, or analytically) smoothing operators. Just as a partial differential operator

  $$P(x, D) = \sum_{|\alpha| \le m} a_\alpha D^\alpha, \quad D = \frac{1}{i}\frac{\partial}{\partial x},$$

may be defined in terms of its "symbol" $p(x,\xi) = \sum_{|\alpha|\le m} a_\alpha \xi^\alpha$,

$$\widehat{P(x,D)u}(\xi) = p(x,\xi)\hat{u}(\xi),$$

so more general "symbols" $p(x,\xi)$, which may be sums of terms homogeneous of decreasing degrees in ξ (or even asymptotic sums of terms of orders going to $-\infty$), define operators by the same formula and obey a calculus similar to that of partial differential operators.

- A partial differential (or pseudodifferential) operator P is called hypoelliptic at x_0 whenever Pu smooth near x_0 (for a function or a distribution u) implies that u is smooth near x_0. The operator P is called analytic hypoelliptic at x_0 whenever Pu real analytic near x_0 (for a function or a distribution u) implies that u is real analytic near x_0.

Chapter 2
Brief Introduction

The techniques and results contained in this monograph arose as, and from, the solution of a long-standing problem on the interface between complex analysis and partial differential equations, namely the (local) analytic hypoellipticity of the "$\overline{\partial}$-Neumann problem" at the boundary of a strictly pseudoconvex domain in $\mathbb{C}^n$. This problem was introduced in its modern form by J.J. Kohn in [K1], where he proved local C^∞ hypoellipticity by way of what amounted, in modern language, to a "subelliptic" estimate for test functions satisfying the so-called $\overline{\partial}$-Neumann boundary conditions ($v \in \mathcal{D}(\overline{\partial}^*)$). This led to the canonical solution of $\overline{\partial}$ on such domains and hence to the solution of the Cousin problem concerning domains of holomorphy. Another approach to the solution of the Cousin problem had been used by Hörmander [Hö4] but is of quite a different character, not involving boundary regularity.

The question of local real analytic regularity of the $\overline{\partial}$-Neumann problem on strictly pseudoconvex domains or slightly more general domains was achieved independently and simultaneously by the present author [T4], [T5] and F. Treves [Tr4]. I shall not comment on Treves' solution until Chap. 14, both because it is extremely different from ours, and, we believe, because it lacks the flexibility and elementary nature of our proof. Indeed, in the succeeding 32 years, a whole host of other problems has proved amenable to our constructions with relatively minor modifications, and I prefer to present those here.

It is this flexibility and enduring success of the present method that have encouraged me to present the story to a wider audience. I will thus explain the origins of the problem as we go along but not dwell on the complex analysis or even on the symplectic geometry involved but rather focus on what has proved to be the most difficult aspect of the work for the reader, namely learning not to become lost in the details.

And details there are. But we believe strongly that once one understands the motivation and meaning of the proof, the details are "mere details", and could be filled in at one's own pace.

It is my fervent hope that the present approach will fulfill this aim.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_2,

Thus the monograph will have an unconventional flow, or at least its flow will differ from that of most “mathematics books”. There will of course be theorems and proofs, but the initial portion will be devoted to a detailed, and extremely intuitive, analysis of the simplest imaginable model that our approach is designed to attack, and even there the result was unknown prior to our work in 1978.

There will be rigorous constructions, but even more space devoted in the first few chapters to the intuitive understanding of *why* these constructions work and, in fact, are the only ones that can work. For the miracle that gradually makes itself visible is that the same basic construction can be easily adapted to numerous more general, and more degenerate, situations.

Only after we feel certain that the serious reader has come to deeply appreciate the value and subtleties of the simplest case will the presentation become more “mathematical” and conventional, with tougher calculations but ones that should not feel more difficult once the first parts have been mastered.

In places, once the new and most difficult material has been introduced and worked through, there will come moments when the proof proceeds precisely as the model case did, where certain concepts and constructions have changed in detail, but not in essential flavor, and in particular have not changed in ways that would affect the argument of the model case. In such situations, we will not repeat all the details of the model case, but will merely substituting the entities that have changed only in ways that do not affect the remaining argument.

We hope and trust that such arguments will not be misunderstood as “hand-waving”; it is also our conviction that at certain points to include every detail would only obfuscate the proof, not elucidate it.

And one more important point: research in this field is ongoing. There are important situations in which analytic hypoellipticity (AHE), even global analytic hypoellipticity (GAHE), fails although many had hoped that it would hold, for example in the case of weakly pseudoconvex domains in complex analysis (cf. [Chr1], [Chr3], [Chr6]), and there are several important open conjectures (the embedding question for strongly pseudoconvex “CR” manifolds of dimension 5, which would follow from a suitable nonlinear AHE result, and the so-called conjecture of Treves, which concerns a certain “Poisson–Treves” stratification of the characteristic manifold, which I shall discuss in due course).

Chapter 3
Overview of Proofs

3.1 A Few Preliminary Definitions

I confess from the outset to a certain prejudice, and we secretly believe that every researcher in this field has one. Mine is that smoothness is good and that more smoothness is better.

Not just good—a differential operator P with the property that whenever the data (the right-hand side of the equation and possibly boundary data) belong to a certain smoothness class, then *any* (distribution) solution has to belong to that class as well (or better, occasionally), has a kind of magic way of conferring on its "offspring" its own characteristics.

Harsh? Rigid? Perhaps, but I believe that studying such disciplined mathematical objects may free one to pursue less clearly defined and restrictive subjects in one's personal life, such as musical aesthetics in our case.

Here, then, are the classes and corresponding types of operators with which we will deal.

The following properties may be *global in* Ω (valid in a given open set Ω), *local at* x_0 (valid in any sufficiently small open set ω containing x_0), *in the sense of germs at* x_0 (valid in some open set containing the point x_0 of interest), or *microlocal at* $(x_0, \xi_0 \neq 0)$ (valid in a sufficiently small open conic neighborhood Γ_{x_0,ξ_0} of (x_0, ξ_0)). A conic neighborhood will mean the Cartesian product of an open set x_0 with an open cone

$$\gamma = \left\{ \xi \in \mathbb{R}^n : \xi \neq 0, \left| \frac{\xi}{|\xi|} - \frac{\xi_0}{|\xi_0|} \right| < \varepsilon \right\}$$

for some (small) ε.

Definition 3.1. A function is called *smooth* if it is infinitely differentiable (C^∞).

Definition 3.2 (HE). A linear partial differential operator is called *hypoelliptic* if any distribution solution must be smooth whenever the data are.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_3,

Definition 3.3. A function g belongs to the *Gevrey class* G^s if it is smooth and satisfies

$$\left|\frac{\partial^k g}{\partial x^k}\right| \leq C_g C_g^k k!^s, \quad k = 1, 2, \ldots .$$

Definition 3.4 (GHE). A linear partial differential operator is called *Gevrey-s hypoelliptic (GHE) (or G^s-hypoelliptic)* if any distribution solution must belong to G^s whenever the data are in G^s.

Definition 3.5. A function is called *real analytic* if it is smooth and belongs to G^1.

Definition 3.6 (AHE). A linear partial differential operator is called *analytic hypoelliptic* (AHE) if any distribution solution must be analytic whenever the data are.

Remark 3.1. Note that the notion of (analytic) hypoellipticity in the sense of germs, say at the origin, may mean that the regularity of the solution may hold only in certain neighborhoods and the proof may not yield local AHE. For example, a proof might apply to neighborhoods with a certain angular symmetry and not to all neighborhoods, no matter how small.

3.2 Elliptic Equations and Boundary Value Problems

Smoothness of solutions to partial differential equations with smooth (analytic) data near a given point is a subject with a long history.

Elliptic equations and systems with suitable coefficients, arguably the "best" PDEs in the sense of my prejudice, have the property that all (distribution) solutions with smooth (real analytic) data near a given point must themselves be smooth (real analytic) near that point.

On the other hand, parabolic equations, such as the heat equation, enjoy the same property but only up to a certain degree of regularity, exhibiting so-called Gevrey-s (G^s) hypoellipticity for $s \geq 2$ but not in the real analytic class (G^1).

Hyperbolic equations, such as the wave equation, are utterly dependent on the way their initial data "propagate," and smooth initial data in any class can "interfere" with themselves (at the vertex of the light cone, for example) and lead even to the nonexistence of solutions or the existence of only very rough ones.

For elliptic equations, such as the Laplace operator, the local C^∞ result is often referred to as Weyl's lemma, although this name can also refer to the real analytic result.

One can also discuss boundary value problems for elliptic operators, and certain homogeneous boundary conditions (such as the Dirichlet or Neumann problem) lead to regularity up to the boundary locally.

We will see below a simple proof of this fact that highlights the difficulties in generalizing the method to other operators.

Of course, it is one thing to prove the local existence of a solution to a partial differential equation (by means of power series in the real analytic category or integration along vector fields using ordinary differential equations in the smooth case, or even by using rather general a priori inequalities) given local data, and it is something else entirely to prove global existence of a solution given global data (generally using a priori inequalities or, in some cases, integration along globally defined vector fields using ODE methods).

Clearly, global existence usually implies local existence, by extending local data to global data using cut-off functions, invoking global existence, and then restricting to the neighborhood in question to generate a local solution. Thus global existence is a more difficult problem than local existence.[1]

On the other hand, local *regularity* of solutions (given *any* solution, which may be a distribution or even a hyperfunction solution, and perhaps defined only near a point, showing that the solution is as regular as the data are near that point) is far more difficult than global regularity, since it requires some way of excluding the influence of more distant information upon the solution. Sometimes the "localization" can be accomplished with smooth cut-off functions φ equal to 1 near the point in question and supported in a small neighborhood ω, but just because one knows that $Lu = f \in C^\infty(\omega)$, one cannot conclude very much about the localized solution φu, since φu is not the solution of the differential equation: all one knows is that

$$L\varphi u = \varphi Lu + [L, \varphi]u = \varphi f + [L, \varphi]u,$$

and the whole story comes down to finding an effective way to handle the commutator term.

It is easy to see that local regularity theorems generate global regularity theorems with no additional work, for if $Lu = f$ globally on some domain Ω with f smooth in all of Ω, then near each point in Ω, f will be smooth; hence, given local regularity, u will be smooth near that point, which is equivalent to saying that the solution u is smooth globally in Ω, i.e., global regularity holds.

However, local regularity is so much more difficult to prove that one usually considers global regularity directly if that is what is of interest.

3.3 The Simplest Subelliptic Case

From the field of several complex variables, one of the early important questions was to identify the "domains of holomorphy" in $\mathbb{C}^n$, namely those domains that are natural domains of existence for holomorphic functions; for any such a domain there exists a function holomorphic there that cannot be extended to any larger domain.

---

[1]The only caveat here is that the extended data, which are usually generated using a cut-off function, will generally need to belong to the class for which global existence is to be proved, and the use of cut-off functions in the real analytic category is not possible without taking great care, as will be described below.

A method to attack this problem was devised by J.J. Kohn. It was a boundary value problem for an elliptic system (proportional to the Laplacian in Euclidean space), known as the "$\overline{\partial}$ -Neumann problem."

This boundary value problem shared some, but not all, properties with coercive (elliptic) boundary value problems. In particular, one could show that this problem was hypoelliptic up to the boundary in the case of so-called strictly pseudoconvex domains (the biholomorphic images of strictly convex domains) in the C^∞ and Gevrey $G^s, s \geq 2$, categories, and even in all $s > 1$ and some quasianalytic classes.

But proving real analytic hypoellipticity (up to the boundary) proved to be a difficult and surprisingly refractory problem, resolved in 1978 independently in [T4], [Tr4].

More recently, these methods have been extended to many weakly pseudoconvex domains and the more abstract case of certain nonelliptic partial differential operators.

The simplest cases to be considered arise in complex analysis under the name of the "Kohn Laplacian" (e.g., for the boundary of the Siegel upper half-space $\Im w > |z|$ in $\mathbb{C}^2$) and in its simplest version may be written as sums of squares of independent vector fields,

$$P^H(x, D) = \sum_{j=1}^{2} \left(X_j^H\right)^2, \tag{3.1}$$

$$X_1^H = \frac{\partial}{\partial x_1} - \frac{x_2}{2}\frac{\partial}{\partial t} = X_1, \quad X_2^H = \frac{\partial}{\partial x_2} + \frac{x_1}{2}\frac{\partial}{\partial t} = X_2, \tag{3.2}$$

or just

$$P = \sum X_j^2$$

for simplicity. This operator has *symplectic* characteristic variety (this means in this setting that the three fields

$$X_1^H,\ X_2^H,\ \text{ and }\ T^H = [X_1^H, X_2^H] \quad \text{span the tangent space } T\mathbb{R}^3 \tag{3.3}$$

near the point in question), and we will be concerned with regularity of solutions near the origin.

This operator P has the coordinate expression

$$\begin{aligned} P &= \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\left(x_1^2 + x_2^2\right)}{4}\frac{\partial^2}{\partial t^2} + \left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right)\frac{\partial}{\partial t} \\ &= \Delta_x + \frac{r^2}{4}\frac{\partial^2}{\partial t^2} + r\frac{\partial}{\partial \theta}\frac{\partial}{\partial t} \end{aligned}$$

with $r^2 = x_1^2 + x_2^2$ and $r\frac{\partial}{\partial\theta} = x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}$.

3.4 Subelliptic Estimates

Establishing subelliptic estimates for the model case and some generalizations is not difficult. For the sum of squares above, it is obvious that we have

$$\sum \|X_j v\|_0^2 \lesssim |(Pv, v)_0| + \|v\|_0^2 \tag{3.4}$$

for v of compact support, where we denote the norm in L^2 by $\|v\|_0$ and the inner product by $(v, w)_0$ or just (v, w). And the notation $A \lesssim B$ means

$$A \leq CB, \tag{3.5}$$

where the constant C will be independent of all variables and functions encountered and might change from line to line within a derivation but could be fixed once and for all.

The estimate (3.4) is actually the *maximal* estimate, not a subelliptic one.

In fact, there are two definitions of subellipticity, closely related to one another. In the form we will first consider them (the "loss of one derivative" case) they are the following:

Definition 3.7 (Norm subellipticity with loss of one derivative).

$$\|v\|_1^2 \lesssim \|Pv\|_0^2 + \|v\|_0^2.$$

Definition 3.8 (Inner product subellipticity with loss of 1/2 derivative).

$$\|v\|_{1/2}^2 \lesssim |(Pv, v)| + \|v\|_0^2,$$

where norms are L^2 Sobolev norms of the indicated order, i.e.,

$$\|v\|_s = \|\hat{v}(\xi)(1 + |\xi|^2)^{s/2}\|_0, \tag{3.6}$$

and v is taken to be smooth and of compact support in a fixed open set. More generally, operators may lose more derivatives, although until recently always with $\varepsilon > 0$:

Definition 3.9 (Norm subellipticity, loss of $2 - 2\varepsilon$ derivatives).

$$\|v\|_{2\varepsilon}^2 \lesssim \|Pv\|_0^2 + \|v\|_0^2. \tag{3.7}$$

Definition 3.10 (Inner product subellipticity, loss of $1 - \varepsilon$ derivatives).

$$\|v\|_{\varepsilon}^2 \lesssim |(Pv, v)| + \|v\|_0^2. \tag{3.8}$$

For model vector fields $X^H, Y^H, T^H = [X^H, Y^H]$ arising as the canonical basis for the so-called left-invariant vector fields on the Heisenberg group (see below), we have the strongest possible (nonelliptic) subelliptic estimate, with $\varepsilon = 1/2$. Starting from

$$\|v\|_{1/2}^2 \lesssim \sum_{j=1}^{2} \|X_j^H v\|_{-1/2}^2 + \|T^H v\|_{-1/2}^2 + \|v\|_0^2, \tag{3.9}$$

where the sum over j is obviously controlled by $|(Pv, v)_0| + \|v\|^2$. For the second term on the right, let Λ denote the pseudodifferential operator with symbol $(1 + |\xi|^2)^{1/2}$. Then

$$\begin{aligned}\|T^H v\|_{-1/2}^2 &= (\Lambda^{-1/2} T^H v, \Lambda^{-1/2} T^H v) = \left(\left(X_1^H X_2^H - X_2^H X_1^H\right) v, \Lambda^{-1} T^H v\right) \\ &= \left(X_2^H v, \left\{\Lambda^{-1} T^H X_1^H + \left[X_1^H, \Lambda^{-1} T^H\right]\right\} v\right) \\ &\quad - \left(X_1^H v, \left\{\Lambda^{-1} T^H X_2^H + [X_2^H, \Lambda^{-1} T^H]\right\} v\right).\end{aligned}$$

These two inner products are similar, and there are two observations to be made:

- The $\|X^H v\|_0^2$ are bounded by $|(Pv, v)_0|$ as desired.
- The brackets and $\Lambda^{-1} T^H$ have order zero and thus are L^2 bounded.

The result, after the use of a "weighted" Schwarz inequality, is that

$$\|T^H v\|_{-1/2}^2 \le \sum \|X_j^H v\|_0^2 + \|v\|_0^2,$$

which implies that

$$\|v\|_{1/2}^2 + \sum \|X_j^H v\|_0^2 \lesssim |(Pv, v)_0| + \|v\|_0^2, \quad v \in C_0^\infty. \tag{3.10}$$

A simple additional argument, namely replacing v by $\Lambda^{1/2} v$, suitably localized in space (since Λ destroys compact support), yields the following result:

Proposition 3.1.

$$\|v\|_1^2 + \sum \|X_j^H v\|_{1/2}^2 + \sum \|X_j^H X_k^H v\|_0^2 \lesssim \|Pv\|_0^2 + \|v\|_0^2, \quad v \in C_0^\infty. \tag{3.11}$$

As for the names "norm subelliptic with loss of one derivative" and "inner-product subelliptic with loss of one-half derivative," both are meant to imply, heuristically at least, that $Pu \in H^s \to u \in H^{s+1}$. While this may appear to be a *gain* of one derivative, compared to the elliptic case ($Pu \in H^s \to u \in H^{s+2}$) it is a weaker result by one derivative, whence the name. And the inner product tag of a "loss of one-half derivative" is a bit harder to justify, although in common usage, since we are using the pairing between $H^{1/2}$ and $H^{-1/2}$, this estimate (3.10) yields

$$\|v\|_{-1/2}^2 \lesssim \|Pv\|_{1/2}^2 + \|v\|_{-1}^2,$$

which clearly represents a loss of one derivative over the elliptic situation.

3.5 Local C^∞ Regularity

To show that all solutions of a partial differential equation satisfying a subelliptic a priori estimate and whose data are smooth must themselves be smooth is a problem that has been studied for a long time, and by now a number of techniques are available.

We identify three notions of regularity of a function, namely *global, local,* and *microlocal*. Their definitions are as follows:

Definition 3.11. The operator P is called *(globally)* $C^\infty(G^s)$ hypoelliptic in Ω if whenever a distribution u satisfies $Pu \in C^\infty(\Omega)\,(G^s(\Omega))$, then $u \in C^\infty(\Omega)\,(G^s(\Omega))$.

Definition 3.12. The operator P is called *(locally)* $C^\infty(G^s)$ hypoelliptic in Ω if each point in Ω has a neighborhood where P is $C^\infty(G^s)$ hypoelliptic.

Here the Gevrey-s class G^s is defined by

Definition 3.13. $w \in G^s$ in U provided every point in U has a neighborhood V and a constant C such that for all α,

$$|D^\alpha w| \leq C^{|\alpha|+1}\alpha!^s,$$

uniformly in V.

In view of Stirling's formula,[2] this is equivalent to the existence of such a C with

$$|D^\alpha w| \leq C^{|\alpha|+1}|\alpha|^{s|\alpha|}.$$

Note that $G^1 = C^\omega$ is the real analytic class.

Again, using the Sobolev embedding theorem,[3] it is not hard to see that these properties are equivalent to the following: for any smooth function $\varphi \equiv 1$ near p but supported in a larger, but still arbitrarily small, neighborhood of p, there exists a constant C with

$$\|\varphi D^\gamma u\|_0 \leq C^{|\gamma|+1}\gamma!^s \quad \forall\gamma$$

or, equivalently, in view of Stirling's formula,

$$\|\varphi D^\gamma u\|_0 \leq C^{|\gamma|+1}|\gamma|^{|\gamma|s} \quad \forall\gamma.$$

We will define microlocal hypoellipticity in various classes a bit later. But microlocal hypoellipticity at all points in the cotangent space will imply local hypoellipticity, and this will imply global hypoellipticity in each of the classes above.

We will work with the local problem and show later how to *microlocalize* the results and the methods.

---

[2] $k^k \leq C^{k+1}k!\ \forall k.$

[3] $H^s \subset C^k$ if $s > k + (n/2)$ and $\|w\|_{C^k} \lesssim \|w\|_s$ in this case.

3.6 Proving C^∞ Regularity

The proof of C^∞ regularity is by now well known, but is not trivial.

First of all, since the space of distributions is the union of all the Sobolev spaces H^s, we will assume that our solution u belongs, say, to H^{s_0} locally or, what is the same thing, that for any function of compact support $\rho(x)$, we have

$$\|\rho u\|_{s_0} < \infty.$$

In an effort to show that ρu also belongs to $H^{s_0+\frac{1}{2}}$, one of the most pleasant proofs, which is linked to our methods in flavor, is to use an "approximate identity," namely convolution with a "bump" function (called a Friedrichs mollifier; cf. [Fr]), unrelated to those to be used later, that renders everything C^∞ and hence allows us to apply the a priori estimate at the H^{s_0} level and later allow the bump function to approach the Dirac delta function suitably and conclude that ρu is in fact in $H^{s_0+\frac{1}{2}}$.

By a subelliptic a priori estimate at the level H^s we mean

$$\|v\|^2_{s+\frac{1}{2}} + \sum \|X_j v\|^2_s \lesssim |(Pv, v)_{H^s}| + \|v\|^2_s \tag{3.12}$$

for $v \in C_0^\infty$.

To obtain this we apply the standard zero-level estimate (3.10) to $\Lambda^s v$, but we need to make sure that we are applying it to functions of compact support.

This is easily done, however: if v has support in a small open set ω, pick a localizing function ρ identically equal to 1 near the closure of ω but supported in a slightly larger open set $\tilde{\omega}$, and observe that the errors committed are of the form $(1-\rho)\Lambda^s v$ in some norm.

But since the supports of v and of $(1-\rho)$ are well separated, such an expression belongs to C^∞ by easy results in pseudodifferential operator theory, and in fact on such v, the operator $(1-\rho)\Lambda^s$ has order $-\infty$. This yields (3.12).

The next step is to define the operator J_ε, following K.O. Friedrichs, by

$$J_\varepsilon w = \chi_\varepsilon * w,$$

where

$$\chi_\varepsilon(x) = \varepsilon^{-n}\chi(x/\varepsilon)$$

for a standard nonnegative C^∞ function χ with support in $(-1, 1)$ and integral equal to 1.

Then for w of small support, $J\varepsilon w$ is in C^∞ and, as $\varepsilon \to 0$, of small support, and $J_\varepsilon w$ will converge to w in any Sobolev space to which w belongs.

But more importantly, if $\{\|J_\varepsilon w\|_s\}$ remains bounded as $\varepsilon \to 0$, then by the Lebesgue monotone convergence theorem (on the Fourier transform side), the $J_\varepsilon w$ actually converge to an element of H^s, which one shows must be w itself.

This is the form in which (3.12) will allow us to show that the solution is in C^∞.

For knowing that $u \in H^{s_0}$, the a priori estimate (3.12) may be applied to ρu with $s = s_0$:

$$
\begin{aligned}
&\|J_\varepsilon \rho u\|^2_{s_0+\frac{1}{2}} + \sum \|X_j J_\varepsilon \rho u\|^2_{s_0} \\
&\quad \lesssim |(P J_\varepsilon \rho u, J_\varepsilon \rho u)_{s_0}| + \|J_\varepsilon \rho u\|^2_{s_0} \\
&\quad \lesssim |(J_\varepsilon \rho P u, J_\varepsilon \rho u)_{s_0}| + \|J_\varepsilon \rho u\|^2_{s_0} + |([P, J_\varepsilon \rho]u, J_\varepsilon \rho u)_{s_0}|.
\end{aligned}
$$

Friedrichs has carried out the estimates on the bracket in his lecture notes [Fr], but it quite clearly suffices to show that $[X_j, J_\varepsilon \rho]$ is bounded in H^{s_0} uniformly in ε or, what is the same thing, that [coeff., J_ε] is of order -1 uniformly in ε.

This is all quite classical, and we refer the reader to [Fr].

Thus we have established

$$
\begin{aligned}
&\|J_\varepsilon \rho u\|^2_{s_0+\frac{1}{2}} + \sum \|X_j J_\varepsilon \rho u\|^2_{s_0} \\
&\quad \lesssim \|J_\varepsilon \rho P u\|^2_{s_0} + \|J_\varepsilon \rho u\|^2_{s_0} + \|[X_j, J_\varepsilon \rho]u\|^2_{s_0} \\
&\quad \lesssim \|J_\varepsilon \rho P u\|^2_{s_0} + \|J_\varepsilon \rho u\|^2_{s_0} + \|J_\varepsilon \rho' u\|^2_{s_0} + \|[J_\varepsilon, X_j]\rho u\|^2_{s_0} \\
&\quad \lesssim \|\rho P u\|^2_{s_0} + \|J_\varepsilon \rho^{(\prime)} u\|^2_{s_0} + \|\rho u\|^2_{s_0},
\end{aligned}
\tag{3.13}
$$

so that indeed, $\|J_\varepsilon \rho u\|^2_{s_0+\frac{1}{2}}$ remains bounded as $\varepsilon \to 0$, and we conclude (using the Lebesgue monotone convergence theorem on the Fourier transform) that $\rho u \in H^{s_0+\frac{1}{2}}$.

If $Pu \in C^\infty = \cap H^s$, then the solution u also belongs to all H^s, hence to C^∞.

Remark 3.2. The same C^∞ regularity result follows from the subelliptic estimate in norm form, (3.11), although the derivation is slightly more complicated, involving double brackets.

3.7 Gevrey Regularity

From the estimate (3.13), or its consequence when $\varepsilon \to 0$, and because of the presence of the term with a derivative on ρ which is in a norm that is better by 1/2 than the norm on the left-hand side, upon iteration until all derivatives on u have been estimated, it is easy to see that in estimating N derivatives on u we encounter $\rho^{(2N)}$ on the right.

If Pu is in G^s with $s = 2$, then there is no obstruction to bounding N derivatives of u by $C^{2N}(2N)!$, i.e., $u \in G^2$, and if Pu belongs to G^s with any larger s, the same estimates will allow us to conclude that the solution belongs to G^s as well. For $s < 2$, the derivatives on ρ will still lead only to G^2 and not better.

More refined analysis is possible for other operators, generally in more variables, where the Gevrey index depends on the direction in which one looks. This is already true for the heat operator, which is analytic hypoelliptic in spatial directions but only G^2 in time, but a much more refined analysis of such operators has been carried out in [ST1] and [ST2].

Now that the solution is kown to be in C^∞, its derivatives may be subjected to the estimate to try to obtain sufficiently uniform estimates.

In the elliptic case, this approach does yield analyticity, as we will now show, but for the subelliptic situation it will not, in general. We start with the easier, elliptic, situation.

3.8 Elliptic Operators

For an elliptic partial differential operator,

$$E = \sum_{j,k} g_{jk}(x) D_j D_k, \tag{3.14}$$

with $\det(g_{jk})$ positive and bounded away from zero, we have the elliptic, or "coercive," a priori estimates

$$\sum_{|\alpha|\le 1} \|D^\alpha v\|_0^2 \lesssim |(Ev, v)_0| + \|v\|_0^2 \tag{3.15}$$

and

$$\sum_{|\alpha|\le 2} \|D^\alpha v\|_0^2 \lesssim \|Ev\|_0^2 + \|v\|_0^2 \tag{3.16}$$

for v of compact support; that is, there is *no loss of derivatives.*

Given $u \in C^\infty$ with $Eu = f \in C^\omega$, the real analytic class, we introduce a nested family of N functions $\{\varphi_j\}$ of compact support, each supported in the set where the next is identically equal to one, $\varphi_1 \equiv 1$ on $\omega \Subset \tilde{\omega}$, and φ_N supported in $\tilde{\omega}$. Derivatives $D^\alpha \varphi_j, \alpha$ a multi-index, will satisfy

$$|D^\alpha \varphi_j| \le CN^{|\alpha|}, \quad |\alpha| \le 2. \tag{3.17}$$

3.8.1 Symmetrization of the Estimates

In the estimate (3.15) above, and also in (3.16), it is useful to symmetrize the location of the cut-off function and other operators.

The following definitions are slightly context-dependent, but the meaning should be clear from the context.

Definition 3.14. For a derivative D_j,

$$^s\|D_j A(v)\|_0 = \|D_j A v\|_0 + \|A D_j v\|_0$$

and, without specifying indices,

$$^{s}\|D^2 A(v)\|_0 = \|D^2 A v\|_0 + \|DADv\|_0 + \|AD^2 v\|_0.$$

Definition 3.15. For a vector field Z and an operator such as a localizing function A,

$$^{s}\|ZAw\|_0 = \|ZAw\|_0 + \|AZw\|_0$$

and, generically,

$$^{s}\|Z^2 Aw\|_0 = \|Z^2 Aw\|_0 + \|ZAZw\|_0 + \|AZ^2 w\|_0.$$

Back to the elliptic case, then, we substitute $v = \varphi_1 D^\beta u$ in (3.15) and commute φ_1 past E. After dropping indices on D, from (3.15) we have

$$\begin{aligned} ^{s}\|D\varphi_1 D^\beta(u)\|_0^2 &\lesssim \|D\varphi_1 D^\beta u\|_0^2 + \|\varphi_1 DD^\beta u\|_0^2 \\ &\lesssim |(E\varphi_1 D^\beta u, \varphi_1 D^\beta u)_0| + {}^{s}\|D\varphi_1^{(\prime)} D^{\beta-1}(u)\|_0^2 \\ &\lesssim |(\varphi_1 D^\beta Eu, \varphi_1 D^\beta u)_0| + |([E, \varphi_1 D^\beta]u, \varphi_1 D^\beta u)_0| \\ &\quad + {}^{s}\|D\varphi_1^{(\prime)} D^{\beta-1}(u)\|_0^2, \end{aligned}$$

so that with integration by parts and a weighted Schwarz inequality,[4] the last expression above satisfies

$$\begin{aligned} \lesssim {}^{s}\|D\varphi_1 D^\beta(u)\|_0^2 &\lesssim \|\varphi_1 D^\beta Eu\|_0^2 + {}^{s}\|D\varphi_1^{(\prime)} D^{\beta-1}(u)\|_0^2 + {}^{s}\|D\varphi_1^{(2)} D^{b-2}(u)\|_0^2 \\ &+ \sum_{0<\tilde\beta\le\beta} \binom{\beta}{\tilde\beta} |(g^{(\tilde\beta)}\varphi_1 D^{\beta+2-\tilde\beta} u, \varphi_1 D^\beta u)_0|. \end{aligned}$$

(Here and elsewhere we may sometimes, and, we admit, abusively, write $\beta - 1$ for some β' with $|\beta'| = |\beta| - 1$ without explicitly mentioning it. Confusing indices and multi-indices in this way will never lead to confusion or to trouble.)

In the last term we integrate one D to the right, in order to produce a gain in $|\beta|$, modulo a possible derivative on φ_1 or another one on one of the coefficients, g, of the elliptic operator E, and also interchange φ_1 and φ_1' in the last term.

The result is

$$\begin{aligned} &\binom{\beta}{\tilde\beta} \left|\left(g^{(\tilde\beta)}\varphi_1 D^{\beta+2-\tilde\beta} u, \varphi_1 D^\beta u\right)_0\right| \\ &\le \binom{\beta}{\tilde\beta} \left|\left(g^{(\tilde\beta)}\varphi_1 D^{\beta+1-\tilde\beta} u, \varphi_1 D^{\beta+1} u\right)_0\right| + \binom{\beta}{\tilde\beta} \left|\left(g^{(\tilde\beta+1)}\varphi_1 D^{\beta+1-\tilde\beta} u, \varphi_1 D^\beta u\right)_0\right| \\ &\quad + \binom{\beta}{\tilde\beta} \left|\left(g^{(\tilde\beta)}\varphi_1 D^{\beta+1-\tilde\beta} u, \varphi_1' D^\beta u\right)_0\right| \end{aligned}$$

---

[4] i.e., for arbitrary $\delta > 0$, $|(v,w)| \le \delta\|v\|^2 + \frac{1}{4\delta}\|w\|^2 = s.c.\|v\|^2 + \ell.c.\|w\|^2$.

$$\leq s.c.^s\|D\varphi_1 D^\beta(u)\|_0^2 + \ell.c.^s\|D\varphi_1' D^{\beta-1}(u)\|_0^2 + C(N^2)^s\|D\varphi_1 D^{\beta-1}(u)\|_0^2$$
$$+\binom{\beta}{\tilde\beta}^2 N^{-2}C^{|\tilde\beta|}(\tilde\beta+1)!^2 \times {}^s\|D\varphi_1 D^{\beta-\tilde\beta}(u)\|_0^2,$$

so that

$$\begin{aligned}&{}^s\|D\varphi_1 D^\beta(u)\|_0\\ &\quad\lesssim \|\varphi_1 D^\beta Eu\|_0 + {}^s\|D\varphi_1' D^{\beta-1}(u)\|_0 + {}^s\|D\varphi_1'' D^{\beta-2}(u)\|_0\\ &\qquad + N^s\|D\varphi_1 D^{\beta-1}(u)\|_0 + \sum_{0<\tilde\beta\leq\beta}\binom{\beta}{\tilde\beta}N^{-1}C^{|\tilde\beta|}(\tilde\beta+1)!^s\|D\varphi_1 D^{\beta-\tilde\beta}(u)\|_0.\end{aligned}$$

To understand this last coefficient, we observe that since $|\beta| \leq N$,

$$\binom{\beta}{\tilde\beta}N^{-1}C^{|\tilde\beta|}(\tilde\beta+1)! \leq C^{\tilde\beta}\binom{\beta}{\tilde\beta}\tilde\beta! = C^{|\tilde\beta|}\frac{\beta!}{(\beta-\tilde\beta)!} \leq C^{\tilde\beta}N^{\tilde\beta}.$$

Thus the estimates may be rephrased:

$$\begin{aligned}{}^s\|D\varphi_1 D^\beta(u)\|_0 \lesssim \|\varphi_1 D^\beta Eu\|_0 + {}^s\|D\varphi_1' D^{\beta-1}(u)\|_0 + {}^s\|D\varphi_1'' D^{\beta-2}(u)\|_0\\ + N\,{}^s\|D\varphi_1 D^{\beta-1}(u)\|_0 + \sup_{0<\tilde\beta\leq\beta} C^{|\tilde\beta|}N^{|\tilde\beta|}\,{}^s\|D\varphi_1 D^{\beta-\tilde\beta}(u)\|_0.\end{aligned} \tag{3.18}$$

Note that of the terms on the right there are:

- a term with Eu, which is known;
- terms where $|\beta|$ has dropped and the drop in $|\beta|$ has appeared as derivatives on φ_1 (at most two of them);
- one term (the fourth) where $|\beta|$ has dropped and instead of a derivative on φ_1 there is a power of N; and
- a term with larger drop in $|\beta|$ and the same kind of compensation with powers of N.

We will not let more derivatives land on φ_1, but introduce φ_2 at this point, then bring φ_1 out of the norm by a constant (N or N^2, depending on how many derivatives have landed on φ_1).

We continue in this way until at most one derivative is left to land on u. We have, starting with $|b| = N$,

$$\|\varphi_1 D^{N+1}u\|_0 \lesssim \sup_{\tilde N\leq N}(CN)^{N-\tilde N}\|D^{\tilde N}Eu\|_{L^2(\tilde\omega)} + (CN)^N\|(D)u\|^2_{L^2(\tilde\omega)},$$

which proves analyticity of u in ω, given that of Eu (and of the coefficients of E in $\tilde\omega$).

3.8.2 Proof viā the Norm Estimate

Rephrasing the above, but this time starting from the norm estimate (3.16);

$$\sum_{|\alpha|\le 2} \|D^\alpha v\|_0^2 \lesssim \|Ev\|_0^2 + \|v\|_0^2,$$

we introduce as before a nested family of N functions $\{\varphi_j\}$ of compact support, each supported in the set where the next is identically equal to one, $\varphi_1 \equiv 1$ on $\omega \Subset \tilde{\omega}$, and φ_N supported in $\tilde{\omega}$, whose derivatives satisfy

$$|D^\alpha \varphi_j| \le CN^{|\alpha|}, \quad |\alpha| \le 2,$$

Again, in the estimate above, symmetrizing the location of the cut-off function as in the definition above; we have

$$\begin{aligned} {}^s\|D^2\varphi_1 D^\beta(u)\|_0 &\equiv \|D^2\varphi_1 D^\beta(u)\|_0 + \|D\varphi_1 DD^\beta(u)\|_0 + \|\varphi_1 D^2 D^\beta(u)\|_0 \\ &\lesssim \|E\varphi_1 D^\beta u\|_0 + {}^s\|D^2\varphi_1^{(\prime)} D^{\beta-1}(u)\|_0 \\ &\lesssim \|\varphi_1 D^\beta Eu\| + \|[E,\varphi_1 D^\beta]u\| + {}^s\|D^2\varphi_1^{(\prime)} D^{\beta-1}(u)\|_0 \\ &\lesssim \|\varphi_1 D^\beta Eu\|_0^2 + \|g(x)\varphi_1^{(\prime)} D^{\beta+1}u\|_0^2 + \|g(x)\varphi_1^{(\prime\prime)} D^\beta u\|_0^2 \\ &\quad + \|\varphi_1[g(x), D^\beta]D^2 u\|_0 + {}^s\|D^2\varphi_1^{(\prime)} D^{\beta-1}(u)\|_0, \end{aligned}$$

where $g(x)$ stands for any of the coefficients of E. Thus

$$\begin{aligned} {}^s\|D^2\varphi_1 D^\beta(u)\|_0 &\lesssim \|\varphi_1 D^\beta Eu\|_0 + {}^s\|D^2\varphi_1' D^{\beta-1}(u)\|_0 + {}^s\|D^2\varphi_1'' D^{\beta-2}(u)\|_0 \\ &\quad + \sum_{0<\tilde{\beta}\le\beta} \binom{\beta}{\tilde{\beta}} \|g^{(\tilde{\beta})}\varphi_1 D^{\beta-\tilde{\beta}} D^2 u\|_0 \\ &\lesssim \|\varphi_1 D^\beta Eu\|_0 + {}^s\|D^2\varphi_1' D^{\beta-1}(u)\|_0 + {}^s\|D^2\varphi_1'' D^{\beta-2}(u)\|_0 \\ &\quad + \sup_{0<\tilde{\beta}\le\beta} C^{|\tilde{\beta}|} \frac{\beta!}{(\beta-\tilde{\beta})!} {}^s\|D^2\varphi_1 D^{\beta-\tilde{\beta}}(u)\|_0, \end{aligned}$$

and from here the estimates proceed exactly as before.

Actually, all of the proofs may be cast in the norm formulation, and in some cases this seems to produce simpler reading.

3.9 Nonelliptic Operators

If, however, the estimate at hand is not elliptic, but rather only subelliptic, it is no longer necessarily true that all solutions need to be locally analytic whenever the data are.

For any given $s > 1$, there is no problem in producing a cut-off function $\varphi \equiv 1$ in ω and supported in $\tilde{\omega}$ with

$$|D^\alpha \varphi| \leq C(C|\alpha|)^{s|\alpha|} \quad \forall \alpha. \tag{3.19}$$

That this is possible amounts to saying that the Gevrey classes strictly larger than the real analytic class are nonquasianalytic, i.e., contain compactly supported functions, and constructing such a φ is simple: since

$$\sum \frac{1}{j^s} < \infty, \tag{3.20}$$

we may take an "infinite convolution" of the bump functions

$$\psi_j(x) = (j^s/d)^n \chi(xj^s) \tag{3.21}$$

for a standard bump $\chi(x)$ of integral one. The ψ_j also have integral equal to one, and a derivative of ψ_j produces a factor j^s, so letting one derivative fall on each of the ψ_j, n derivatives of the convolution would have the bound $\Pi_1^n (Cn)^s = (C_s^k) n!^s$.

Then the above analysis, keeping the same localizing function throughout and using only the $1/2$ norm of the estimate (3.12), will produce

$$|D^k u| \leq \left(C_s^k\right) k!^s, \tag{3.22}$$

provided this is true of Pu as well.

3.9.1 The Baouendi–Goulaouic Example; Sharpness

In 1972 [BG], M.S. Baouendi and C. Goulaouic produced an example of a very simple subelliptic operator that they showed was not analytic hypoelliptic (AHE): there were nonanalytic solutions to the homogeneous equation that had polynomial coefficients. Their operator was

$$P_{BG} = D_x^2 + D_t^2 + x^2 D_s^2. \tag{3.23}$$

As seen above, this operator is C^∞ hypoelliptic and even G^s hypoelliptic for any $s \geq 2$, yet Baouendi and Goulouic produced a homogeneous solution $P_{BG} u = 0$

that was in G^2 but in no G^s with $s < 2$. We will have more to say about such operators below. For the moment, we merely write down a counterexample to AHE for P_{BG}: for any $\varepsilon > 1$, set

$$u_\varepsilon(x,t,s) = \int_0^\infty \exp\left[i\rho^2 s - t\rho - \rho^2 x^2/2 - \rho^\varepsilon\right] d\rho.$$

Then it is not difficult to see that $P_{BG}\, u_\varepsilon = 0$, yet $u_\varepsilon \in G^{2/\varepsilon}$ but is in no smaller Gevrey class. This clearly shows that the operator cannot be G^s hypoelliptic for any $s < 2$. Note the subtle difference between this proof and the construction of Baouendi and Goulalouic, both of which imply non-AHE.

3.10 The Analyticity Problem and Its Solution

3.10.1 Obstructions to Proving Analyticity

Nonelliptic operators P will satisfy a priori estimates weaker than the coercive or elliptic ones but strong enough, nonetheless, for most purposes.

For the sum of squares under consideration, using [RS], one has

$$\|v\|_{1/2}^2 + \sum_{j=1}^{2} \|X_j v\|_0^2 \lesssim |(Pv,v)_0| + \|v\|_0^2, \quad v \in C_0^\infty, \tag{3.24}$$

for all v with support in a fixed bounded open set containing the origin (in (x_1, x_2, t)).

There are, in fact, examples of operators P satisfying the subelliptic estimate above that do not have this last property.

But for those with symplectic characteristic variety, as our P has, it was hoped that the finer regularity results would be true, and this monograph is dedicated to that proof.

3.10.2 Why the Most Naive Approach Fails

Suppose that we know that a function u belongs to C^∞ and that φ is one of the "localizing functions" mentioned above. To obtain sufficient bounds for derivatives of u, given $Pu = f \in G^s, 1 \le s < 2$, one might try to replace the "test function" v in the estimates above by $v = Su = \varphi T^p u$, where $T = \frac{1}{i}\frac{\partial}{\partial t}$, obtaining

$$
\begin{aligned}
&\sum_{j=1}^{2} \|X_j \varphi T^p u\|_0^2 + \|\varphi T^p u\|_{1/2}^2 \\
&\quad = \sum_{j=1}^{2} \|X_j S u\|_0^2 + \|Su\|_{1/2}^2 \lesssim |(PSu, Su)_0| + \|Su\|_0^2 \\
&\quad \lesssim |(SPu, Su)_0| + |([P, S]u, Su)| + \|Su\|_0^2 \\
&\quad = |(Sf, Su)_0| + \sum_{1}^{2} |([X_j^2, S]u, Su)| + \|Su\|_0^2 \\
&\quad = |(Sf, Su)_0| + 2\sum_{j=1}^{2} \Big|([X_j, S]u, X_j^* Su)\Big| + \sum_{j=1}^{2} |([X_j, [X_j, S]]u, Su)| + \|Su\|_0^2.
\end{aligned}
\tag{3.25}
$$

Now the first term on the right is fine: using the Schwarz inequality,

$$
|(Sf, Su)_0| \lesssim \|Sf\|_0^2 + \|Su\|_0^2;
$$

the first of these is known, while the second is the fourth term in the previous line.

That fourth term itself is the same as the second term on the left, but has "gained" half a derivative; this step may be iterated easily.

But it is the errors that come up in treating the other two terms that will bother us. For instance, since in this case $X_j^* = -X_j$,

$$
\begin{aligned}
&\Big|([X_j, S]u, X_j^* Su)\Big| \\
&\quad = |([X_j, S]u, X_j Su)| \lesssim s.c.\|X_j Su\|_0^2 + \ell.c.\|[X_j, S]u\|_0^2 \\
&\quad = s.c.\|X_j Su\|_0^2 + \ell.c.\|\varphi' T^p u\|_0^2 = s.c.\|X_j \varphi T^p u\|_0^2 + \ell.c.\|\varphi' T^p u\|_0^2
\end{aligned}
$$

where the first term on the right will be absorbed on the left-hand side of (3.25) but not the second term, since φ has been differentiated.

To iterate this estimate, ignoring the double commutator term for the moment, we first need to write the last term in the form present on the left of (3.25). We will think (microlocally, since only T derivatives are in question) of

$$
\|\psi T^\ell v\|_0 \sim \|\psi T^{\ell - 1/2} v\|_{1/2},
$$

i.e., deal with pseudodifferential operators, in particular Λ_t, whose symbol is $(1 + |\tau|^2)^{1/2}$, and write

$$
\begin{aligned}
\|\psi T^\ell v\|_0 &\sim \|\psi \Lambda_t^\ell v\|_0 \sim \|\Lambda_t^{1/2} \psi \Lambda_t^{\ell-1/2} v\|_0 + \|[\Lambda_t^{1/2}, \psi]\Lambda_t^{\ell-1/2} v\|_0 \\
&\sim \|\psi \Lambda_t^{\ell-1/2} v\|_{1/2} + \|[\Lambda_t^{1/2}, \psi]\Lambda_t^{\ell-1/2} v\|_0.
\end{aligned}
$$

Now, we know from the calculus of pseudodifferential operators that the bracket here has a whole asymptotic expansion in decreasing powers of Λ_t and increasing numbers of derivatives on ψ; thus once we learn how to control high derivatives of ψ, we will arrive at an iteration of the form

$$\sum_{j=1}^{2} \|X_j \varphi \Lambda_t^k u\|_0^2 + \|\varphi \Lambda_t^k u\|_{1/2}^2$$
$$\lesssim \|\varphi \Lambda_t^k f\|_0^2 + \|\varphi' \Lambda_t^{k-1/2} u\|_{1/2}^2 + \|\varphi'' \Lambda_t^{k-3/2} u\|_{1/2}^2 + \cdots . \qquad (3.26)$$

All of the terms on the right must be treated, but since the orders are decreasing, the highest-order term presents one derivative on the localizing function for a gain of $1/2$ derivative on u, and ultimately leads to the bound (after iteration)

$$\|\varphi \Lambda_t^k u\|_{1/2}^2 \le CC^k \|\varphi^{(2k)} u\|_{1/2}^2.$$

For suitable localizing functions, this will lead to the second Gevrey class and not better.

This is the general reason why the most straightforward methods used for certain operators satisfying subelliptic estimates do not lead to analytic regularity of solutions (even though, in some very important cases, these operators are analytic hypoelliptic).

(We will not dwell on the double commutator $([X_j, [X_j, \varphi T^p]]u, \varphi T^p u)$ at the moment, since we will need to consider it later anyhow in a more complex context and we prefer to leave the introduction without too many technicalities.)

3.10.3 The Flavor of Our Methods

To prove that locally, for a function u, $Pu = f$ in some smoothness class implies that u is in the same (or possibly a different) smoothness class, for our operator P, which is singular (nonelliptic) only "in the $\frac{\partial}{\partial t}$ direction," in a sense to be made precise below, it suffices to show that high derivatives in t can be controlled suitably.

Denoting by T the operator

$$T = \frac{\partial}{\partial t},$$

and taking a suitable function $\varphi \in C_0^\infty(\omega)$, ω a small open set containing the point p_0 in question, with

$$\varphi \equiv 1 \quad \text{near} \quad p_0,$$

it would suffice to show that

$$\|\varphi T^k u\|_0 \le CC^k k!^s \quad \forall k.$$

Of course, since there will be a trade-off between lowering the order k of differentiation in the vector field T and derivatives landing on the localizing function, we will need to use very special localizing functions, or families of such functions, introduced by Ehrenpreis, that behave "analytically" while still having compact support. For the moment, suffice it to say that this is not an insurmountable problem.

Our localization of T^p will be of the form

$$(T^p)_\varphi = \varphi T^p + \text{ terms where } \varphi \text{ is differentiated}$$

such that

$$(T^p)_\varphi = T^p \qquad \text{in any open set where } \varphi \equiv 1.$$

Thus it will suffice to show that

$$\|(T^k)_\varphi u\|_0 \le C^{k+1}k!$$

in order to conclude that u is analytic in any open set where $\varphi = 1$.

In fact, and this is Ehrenpreis's crucial observation, φ may depend on p as long as the constant C does not, and the different φ are all equal to 1 on a common open set and supported where we have information about the data.

What we actually use is that given two nested open sets $\omega \Subset \tilde{\omega}$, with separation $d = \text{dist}(\omega, \tilde{\omega}^c)$ and a positive integer N, there exists a localizing function $\varphi_N \equiv 1$ on ω and in $C_0^\infty(\tilde{\omega})$ with

$$|D^\alpha \varphi_N| \le (C/d)^{k+1}N^k, \quad \forall |\alpha| = k \le 2N. \tag{3.27}$$

And the construction of φ_N is not difficult: one picks the characteristic function of an intermediate open set ω' and N identical copies of a "bump" function of support proportional to d/N but integral equal to 1 and convolves them all; high derivatives may be distributed over the many copies of the bump function so that only two, at most, need land on any one bump function.

Because of the the width of the support of these bump functions, one derivative is proportional to N, and two derivatives proportional to N^2 in the sup norm. More concretely, we will write

$$\varphi_N = N\psi(Nx/2d) * \cdots * N\psi(Nx/2d) * \chi$$

with N copies of $N\psi$, ψ a nonnegative smooth function of compact support in the unit ball and integral equal to one and χ the characteristic function of the intermediate open set ω.

The simplest subelliptic operator we are able to treat is built out of the left-invariant vector fields on the real Heisenberg group:

$$X_j = \frac{\partial}{\partial x_j} - \frac{y_j}{2}\frac{\partial}{\partial t}, \qquad Y_j = \frac{\partial}{\partial y_j} + \frac{x_j}{2}\frac{\partial}{\partial t}, \qquad T = \frac{\partial}{\partial t}.$$

And the operator in question is

$$P = \sum_1^n X_j^2 + \sum_1^n Y_j^2 = \sum_1^{2n} Z_j^2$$

(where each Z is an X or a Y and one may add first-order terms in the X's and Y's at will).

Since these vector fields satisfy the Hörmander condition that they together with their first brackets span the tangent space of $\mathbb{R}^{2n+1}$, this operator is subelliptic with loss of $1/2$ derivative in the inner product sense, and clearly controls the elliptic terms as well: for $v \in C_0^\infty$,

$$\|v\|_{1/2}^2 + \sum_1^{2n} \|Z_j v\|_0^2 \lesssim |(Pv, v)_0| + \|v\|_0^2,$$

or if we assume for the moment that our solution u is already known to be C^∞, we may substitute $v = (T^p)_\varphi u$ and obtain

$$\|(T^p)_\varphi u\|_{1/2}^2 + \sum_1^{2n} \|Z_j (T^p)_\varphi u\|_0^2 \lesssim |(P(T^p)_\varphi u, (T^p)_\varphi u)_0| + \|(T^p)_\varphi u\|_0^2.$$

As before, we know a lot about $(T^p)_\varphi Pu$ but not much about $P(T^p)_\varphi u$, which will require us to commute P with $(T^p)_\varphi$.

We are incredibly fortunate that the Heisenberg group commutation relations among the Z_j permit us to construct the remaining terms of $(T^p)_\varphi$ so that $(T^p)_\varphi$ will commute well with P.

For starters, $(T^p)_\varphi$ must commute well with the Z_j that make up P.

3.10.4 The Construction of $(T^p)_\varphi$

We have seen that we must control the bracket $[X_j, \varphi T^p + \cdots]$, where the $\ldots$ are to be determined, with the property that each term has at least one derivative on φ.

The first requirement is that

$$[X_j, \varphi T^p + \cdots] = (X_j\varphi)T^p + [X_j, \ldots] \text{ not containing } T^p,$$

or, what is the same thing,

$$[X_j, \ldots] = -(X_j\varphi)\, T^k \quad \text{mod terms with at most } T^{k-1}.$$

If we start with $k = 1$, things simplify. A provisional choice is

$$(T^1)_\varphi = \varphi \circ T - \sum (X_j\varphi) \circ Y_j \tag{3.28}$$

as an operator, i.e., (3.28) means

$$(T^1)_\varphi w = \varphi \circ Tw - (X_j\varphi) \circ Y_j w,$$

for then we see rapidly that (always as operators)

$$[X_k, (T^1)_\varphi] = -\sum_j (X_k X_j \varphi) \circ Y_j.$$

But of course, we will also encounter the bracket with Y_k, and we have introduced nothing to "kill" T when bracketed with Y_k.

But the choice is not hard to find; namely, if we let

$$(T^1)_\varphi = \varphi \circ Tw - \sum_j (X_j\varphi) \circ Y_j + \sum_j (Y_j\varphi) \circ X_j,$$

there is little change in the bracket with X_k, since the $[X_j, X_k]$ are equal to 0, but the additional terms perform the same miracle with Y_k:

$$[X_k, (T^1)_\varphi] = -\sum_j (X_k X_j \varphi) Y_j + \sum_j (X_k Y_j \varphi) X_j$$

and

$$[Y_k, (T^1)_\varphi] = -\sum_j (Y_k X_j \varphi) Y_j + \sum_j (Y_k Y_j \varphi) X_j.$$

We are tempted to say so far, so good, except that the above brackets may be written, generically,

$$[Z, (T^1)_\varphi] = \varphi'' Z,$$

which is not so thrilling: starting from $\|\varphi T u\|^2_{1/2}$ we have arrived at $\|\varphi'' Z u\|_0$.

To put the matter differently, since in the estimates the $1/2$ Sobolev norm carries the same weight as having a Z derivative, just appending additional powers of T (without sophisticated localization) would give, in L^2 norms,

$$\|Z(T^1)_\varphi T^\ell u\|_0 \to \|Z\varphi'' T^\ell u\|_0 = \|Z\varphi'' T T^{\ell-1} u\|_0,$$

where a "gain" of one power of T has "cost" two derivatives on φ, but now the $\varphi'' T$ on the right does not occur in the form of $(T^1)_{\varphi''}$, which helped so much in the first bracket, and as we saw above, without this special form we are led only to the second Gevrey class G^2 and not to G^1.

Since it is the apparent trade of one power of T for two derivatives on φ that leads to apparent defeat, a new interpretation of this "trade-off" was what saved the day.

For if one tries to generalize the construction of $(T^1)_\varphi$ in perhaps the only reasonable way, namely,

$$\begin{aligned}(T^2)_\varphi &= (T^1)_\varphi T + \sum_{i,j} \frac{(X_i X_j \varphi)}{2!} Y_i Y_j - \sum_{i,j} \frac{(X_i Y_j \varphi)}{1!1!} X_j Y_i + \sum_{i,j} \frac{(Y_i Y_j \varphi) X_i X_j}{2!} \\ &= (T^1)_\varphi T + \sum_{i,j} \frac{((-X_i)(-X_j)\varphi)}{2!} Y_i Y_j \\ &\quad + \sum_{i,j} \frac{((-X_i) Y_j \varphi)}{1!1!} X_j Y_i + \sum_{i,j} \frac{(Y_i Y_j \varphi)}{2!} X_i X_j \\ &= \sum_{|\alpha|+|\beta|\le 2} \frac{(-X)^\alpha Y^\beta (\varphi)}{\alpha!\beta!} \circ X^\beta Y^\alpha T^{2-|\alpha|-|\beta|}, \end{aligned} \tag{3.29}$$

we obtain the following very interesting commutation relations:

$$[X_\ell, (T^2)_\varphi] \equiv 0 \quad \mod \frac{Z^3 \varphi}{2!} \circ Z^2$$

and

$$[Y_\ell, (T^2)_\varphi] \equiv (T^1)_{T\varphi} \circ Y_\ell \quad \mod \frac{Z^3 \varphi}{2!} \circ Z^2,$$

where again, each occurrence of Z denotes a Z or a Y (or its negative).

In fact, now the obvious generalization

$$(T^p)_\varphi = \sum_{|\alpha|+|\beta|\le p} \frac{(-X)^\alpha Y^\beta (\varphi)}{\alpha!\beta!} \circ X^\beta Y^\alpha T^{2-|\alpha|-|\beta|} \tag{3.30}$$

yields

$$[X_\ell, (T^p)_\varphi] \equiv 0 \quad \mod \frac{Z^{p+1}(\varphi)}{p!} \circ Z^p \tag{3.31}$$

and

$$[Y_\ell, (T^p)_\varphi] \equiv (T^{p-1})_{T\varphi} \circ Y_\ell \quad \mod \frac{Z^{p+1}(\varphi)}{p!} \circ Z^p, \tag{3.32}$$

as a straightforward shift of index argument shows.

Actually there are choices to be made even here, notably the order of X's and Y's to the right and left of φ, and each choice has advantages and disadvantages, but we will settle on this choice. The fact that one cannot satisfy the bracket with the X_ℓ and the Y_ℓ exactly (modulo pure Z's) in fact led to an impasse that lasted for quite a long time, and the beautiful principal error in brackets with Y_ℓ came as a wonderful surprise and provided the essential clue that this construction might be very powerful indeed. More difficult still will be brackets with functions, as we will see below.

3.11 The Role of Strictness

Given a set of real vector fields X_j, Y_k, with, for example, the property that the X_j commute among themselves and the same is true for the Y_j, $j \leq n$, in $\mathbb{R}^{2n+1}$, with a complementary vector field T, the matrix $c_{i,j}$ given by

$$[X_i, Y_j] \equiv c_{i,j} T \quad \text{mod } \{X_j\}, \{Y_k\}$$

plays a crucial role. It is often known as the Levi matrix (especially in the case of complex vector fields).

When the determinant of c_{ij} is nonzero, the structure is known as "nondegenerate," when it is nonnegative definite, the structure is "pseudoconvex," and when it is positive definite, the structure is known as "strictly pseudoconvex," all terms originating from complex analysis.

The case we are studying initially has the matrix c_{ij} a multiple of the $n \times n$ identity matrix. And the case in which the vector fields are those coming from the Heisenberg group provide the simplest model.

We have already met, in the Baouendi–Goulaouic operator above, a case in which, in some sense, one of the eigenvalues is identically equal to zero, and we found that the operator is not real analytic hypoelliptic.

A more subtle situation occurs when the Levi form is weakly definite, such as would be given by the vector fields in $\mathbb{R}^3$,

$$\tilde{X} = \frac{\partial}{\partial x} - \frac{y^3}{3}\frac{\partial}{\partial t}, \quad \tilde{Y} = \frac{\partial}{\partial y} + \frac{x^3}{3}\frac{\partial}{\partial t}.$$

Here $c_{11} = x^2 + y^2$, which is positive definite except at the origin in (x, y) and nonnegative throughout, but the analysis becomes much subtler.

3.12 Treves' Conjecture

For many years there has been a conjecture, shown consistent in many cases but not proven, due to François Treves. We will not go into the full conjecture here, since it goes deep into the symplectic geometry of the characteristic variety of the operator and this book is about positive results, but we shall give some examples and give the conjecture in easy cases.

The conjecture may be most easily stated in terms of a set of real vector fields $X_j = X_j(x, D) = \sum_{k=1}^n a_{jk} D_k$, $D_k = \frac{1}{i}\frac{\partial}{\partial x_k}$, in $\mathbb{R}^n$, with corresponding symbols $\sigma_j(x, \xi) = \sum_{k=1}^n a_{jk}\xi_k$.

The characteristic set of $\{X_1, \ldots, X_m\}$ is given by

$$\text{Char}(X_1, \ldots, X_m) = \{(x, \xi) : \sigma_j(x, \xi) = 0,\ \xi \neq 0,\ \forall j\}.$$

It is the disjoint union of "strata" Σ_j, each of which is a real analytic manifold on which the symplectic form has constant rank and on which the Poisson brackets up to a length r_j vanish, but at each point of Σ_j there is a bracket of length r_{j+1} that does not vanish.

The actual construction consists in looking at the connected real analytic components of $\mathrm{Char}(X_1, \ldots, X_m)$ on each of which the symplectic form has constant rank.

On each of these components, one looks at the iterated Poisson brackets of the symbols of the $\{X_j\}$ and breaks up the component into a disjoint union of subcomponents on which the length of Poisson brackets required to obtain a nonzero function is constant. Those for which the length is r constitute the stratum of depth r.

Each of the subcomponents is treated in the same way.

In the end, one arrives at a (finite!) collection (because of the Hörmander property of the vector fields) of real analytic connected manifolds on which the symplectic form and the Poisson "depth" are constant.

Treves' conjecture is that an operator $P = \sum X_j^2$ should be analytic hypoelliptic if and only if each stratum is *symplectic*, which means that the symplectic form is nondegenerate on each stratum.

3.12.1 A Particular Case

In a particular case that models the operator $D_1^2 + x_1^2 D_2^2 + x_1^4 D_3^2$, the conjecture becomes a bit simpler. Everything in this section is microlocal, i.e., taken near $(x_0, \xi_0) \in T^* R^3$.

The assumptions for this section are that in $\mathbb{R}^3$,

$$P(x, D) = \sum_{j=1}^{3} X_j^2(x, D),$$

and:

(A1) The operator P satisfies the Hörmander Lie algebra condition and hence is C^∞ hypoelliptic.

(A2) The characteristic set Σ_0 of P,

$$\Sigma_0 = \{(x, \xi) : X_j(x, \xi) = 0 \ \forall j\},$$

is an analytic symplectic submanifold of $T^*\mathbb{R}^3 \setminus 0$.

(A3) The vector fields X_j are linearly independent.

(A4) Define

$$\Sigma_1 = \{(x, \xi) \in T^*\mathbb{R}^3 \setminus 0 \mid (x, \xi) \in \Sigma_0, \{X_i, X_j\}(x, \xi) = 0, \forall i, j\},$$

and in general, let $I = (i_1, \ldots, i_k)$, $i_j \in \{1, 2, 3\}$, for $j = 1, \ldots, k$. Writing $|I| = k$, we denote by X_I the iterated Poisson bracket

$$X_I = \{X_{i_1}, \{X_{i_2}, \ldots ,\{X_{i_{k-1}}, X_{i_k}\} \ldots\}\}$$

of the vector fields X_j, $j = 1, 2, 3$; set

$$\Sigma_h = \{(x, \xi) \in T^*\mathbb{R}^3 \setminus 0 \mid (x, \xi) \in \Sigma_{h-1},\ X_I(x, \xi) = 0$$
$$\text{for every index } I \text{ such that } |I| = h\}.$$

Assume that each of the Σ_j is an analytic submanifold.

Then the conjecture, in this setting, says that P should be analytic hypoelliptic if and only if each of the (nonempty) Σ_j is symplectic.

In particular, the operator

$$P_G = D_{x_1}^2 + x_1^{2k} D_{x_2}^2 \tag{3.33}$$

has $\Sigma_0 = \{\xi_1 = x_1 = 0\}$ (since $\xi \neq 0$ and $\xi_1 = x_1\xi_2 = 0 \implies \xi_1 = x_1 = 0$) and also all $\Sigma_\ell = \Sigma_0$ for $\ell < k$. These are all symplectic, and the operator P_G is known to be analytic hypoelliptic.

But by contrast, the Baouendi–Goulaouic operator (see below),

$$P_{BG} = D_{x_1}^2 + D_{x_2}^2 + x_1^2 D_{x_3}^2,$$

has $\Sigma_0 = \{(x, \xi) : x_1 = \xi_1 = \xi_2 = 0\}$, which is not a symplectic submanifold of $T\mathbb{R}^3$.

Together with A. Bove, we have examined numerous examples in the light of the conjecture, with results consistent with the conjecture, and have a further conjecture on the Gevrey regularity and sharp Gevrey regularity index for similar operators in terms of the first k for which Σ_k is not symplectic and the first ℓ for which $\Sigma_\ell = \emptyset$, but the conjectures remain open in general.

3.13 Counterexamples in the Complex Domain

Before getting too enthusiastic over these methods, we need to remind the reader that in fact there are subelliptic sums of squares of real vector fields that are not analytic hypoelliptic. The Baouendi–Goulaouic example is the simplest of these, and in fact can be cast in the context of $\Box_b$ for a pseudoconvex, but not strictly pseudoconvex, domain.

The following argument, inspired by one due to G.B. Folland (private communication), shows how the sum of real vector fields can arise from complex analysis. In $\mathbb{C}^2$, consider the Heisenberg group vector fields $\{L_H, \overline{L_H}\}$ in $\{z_1, w_1\}$ together with a flat piece

$$L_F = \partial_{z_2}.$$

Then a brief calculation gives

$$\frac{1}{2}\left\{L_H \overline{L_H} + \overline{L_H} L_H\right\} = \partial^2_{z_1 \overline{z_1}} + |z_1|^2 \partial^2_{\Re w_1} + \partial_{\theta_1} \partial_{\Re w_1},$$

where

$$\partial_{\theta_1} = x_1 \partial_{y_1} - y_1 \partial_{x_1}, \quad z_1 = x_1 + i y_1.$$

Thus

$$\begin{aligned} &\frac{1}{2}\left\{L_H \overline{L_H} + L_F \overline{L_F} + \overline{L_H} L_H + \overline{L_F} L_F\right\} \\ &= \partial^2_{z_1 \overline{z_1}} + |z_1|^2 \partial^2_{\Re w_1} + \partial_{\theta_1} \partial_{\Re w_1} + \partial^2_{x_2} + \partial^2_{y_2} \\ &= \partial^2_{x_1} + \partial^2_{y_1} + x_1^2 \partial^2_{\Re w_1} + y_1^2 \partial^2_{\Re w_1} + \partial_{\theta_1} \partial_{\Re w_1} + \partial^2_{x_2} + \partial^2_{y_2}. \end{aligned}$$

But on functions independent of y_1 and then evaluated at $y_1 = 0, \partial_{\theta_1} = 0$, and hence on functions independent of y_1 and y_2 and then evaluated at $y_1 = 0$,

$$\frac{1}{2}\left\{L_H \overline{L_H} + L_F \overline{L_F} + \overline{L_H} L_H + \overline{L_F} L_F\right\} = \partial^2_{x_1} + x_1^2 \partial^2_{\Re w_1} + \partial^2_{x_2},$$

the operator of Baouendi and Goulaouic (in the real domain).

3.14 AHE for $D_{x_1}^2 + x_1^{2k} D_{x_2}^2$ (No $(T^p)_\varphi$ Needed!)

We conclude this chapter with a simple situation in which analytic hypoellipticity has been known for some time, namely that of $D_{x_1}^2 + x_1^{2k} D_{x_2}^2$, which we will write as $P_G = D_x^2 + x^{2k} D_y^2$ for simplicity.

Admitting that it suffices to bound powers of the given vector fields D_x and $x^k D_y$ on a solution u to $P_G u = 0$ (using Cauchy–Kovalevskaya to find an inhomogeneous solution to a given inhomogeneous problem and then taking the difference, using the linearity of P_G) in L^2 norm and that for $x \neq 0$ the operator is elliptic, so localization in y alone is required, we write, with $Z = D_x$ or $x^k D_y$,

$$\begin{aligned}
&\|D_x\varphi(y)D_y^p u\|^2 + \|x^k D_y\varphi(y)D_y^p u\|^2 \\
&\quad \le |(P_G\varphi(y)D_y^p u, \varphi(y)D_y^p u)| = |([P_G, \varphi(y)D_y^p]u, \varphi(y)D_y^p u)| \\
&\quad \le |(x^k\varphi' x^k D_y D_y^p u, \varphi D_y^p u)| + |(x^k D_y x^k \varphi' D_y^p u, \varphi D_y^p u)| \\
&\quad \lesssim s.c.\|x^k D_y \varphi D_y^p u\|^2 + \ell.c.\|x^k \varphi' D_y D_y^{p-1} u\|^2.
\end{aligned}$$

Thus except for the order of φ' and D_y in the last term, a phenomenon we will deal with later, and inverting the order clearly will produce φ'' with the order of D_y decreased by one more etc.; except for this, the order of differentiation has dropped by one and the number of derivatives on φ increased by one. At this point one may bring the localizing function out of the norm and introduce another, which moves from being equal to 1 to being 0 in a band of width C/p, or use special functions, described below, that can absorb up to p derivatives and grow as if they were analytic. In either case, iteration leads to $C^p p^p$ or $C^p p!$, i.e., analyticity (in y).

In x, we are led back to D_y derivatives as well as lots of derivatives on φ (or many φ's with nested supports as mentioned above):

$$\begin{aligned}
&\|D_x\varphi(y)D_x^p u\|^2 + \|x^k D_y\varphi(y)D_x^p u\|^2 \\
&\quad \le |(P_G\varphi(y)D_x^p u, \varphi(y)D_x^p u)| = |([P_G, \varphi(y)D_x^p]u, \varphi(y)D_x^p u)| \\
&\quad = |([x^{2k}D_y^2, D_x^p\varphi]u, \varphi D_x^p u)| \\
&\quad \lesssim 2kp|(x^{k-1}D_x\varphi D_x^{p-2}D_y u, x^k D_y \varphi D_x^p u)| \\
&\qquad +|(x^k D_x \varphi' D_x^{p-1}u, x^k D_y\varphi D_x^p u)| + l.o.t. \\
&\quad \lesssim \rightarrow \cdots \rightarrow s.c.(\text{LHS}) + \ell.c.\ C^p p^{p/2}\|\varphi D_y^{p/2}u\|^2 + \ell.c.\ C^p\|\varphi^{(p)}u\|^2.
\end{aligned}$$

But we have seen that y-derivatives grow analytically, whence

$$\|D_x\varphi(y)D_x^p u\|^2 + \|x^k D_y\varphi(y)D_x^p u\|^2 \lesssim C^{p+1}p^p \lesssim C^{p+1}p! \quad \square$$

Chapter 4
Full Proof for the Heisenberg Group

4.1 The Model Operator

The simplest case in which the operator $(T^p)_\varphi$ provided a breakthrough came from the so-called left invariant vector fields on the Heisenberg group.

One need not know anything about the Heisenberg group to write down these vector fields, as we have done above: for $j \leq n$,

$$X_j = \frac{\partial}{\partial x_j} - \frac{y_j}{2}\frac{\partial}{\partial t}, \qquad Y_j = \frac{\partial}{\partial y_j} + \frac{x_j}{2}\frac{\partial}{\partial t}, \qquad T = \frac{\partial}{\partial t}.$$

Actually, via the change of variables $\tilde{x}_j = x_j$, $\tilde{y}_j = y_j$, $\tilde{t} = t - \frac{1}{2}\sum x_j y_j$, these fields are transformed into the fields

$$X_j = \frac{\partial}{\partial x_j} - y\frac{\partial}{\partial t}, \qquad Y_j = \frac{\partial}{\partial y_j}, \qquad T = \frac{\partial}{\partial t}; \tag{4.1}$$

We shall use the fields in this form from now on. The operator in question is

$$P = \sum_1^n X_j^2 + \sum_1^n Y_j^2 = \sum_1^{2n} Z_j^2,$$

where each Z_j is either X_j or Y_j, and we could as well write

$$-P = \sum_1^n X_j^* X_j + \sum_1^n Y_j^* Y_j = \sum_1^{2n} Z_j^* Z_j,$$

since here $Z_j^* = -Z_j$, and in general, for real vector fields, they will differ by a smooth function.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_4,

As we have seen, this operator is subelliptic with loss of $1/2$ derivative in the inner product sense, and clearly controls the "elliptic" terms as well: for $v \in C_0^\infty$,

$$\|v\|_{1/2}^2 + \sum_1^{2n} \|Z_j v\|_0^2 \lesssim |(Pv, v)_0| + \|v\|_0^2. \tag{4.2}$$

4.1.1 T Derivatives

We assume for the moment that our solution u is already known to be C^∞, and taking any localizing function φ, we may take $v = (T^p)_\varphi u$ as above:

$$(T^p)_\varphi = \sum_{|\alpha|+|\beta|\le p} \frac{(-X)^\alpha Y^\beta(\varphi)}{\alpha!\beta!} \circ X^\beta Y^\alpha T^{2-|\alpha|-|\beta|}$$

$$\|(T^p)_\varphi u\|_{1/2}^2 + \sum_1^{2n} \|Z_j (T^p)_\varphi u\|_0^2 \lesssim |(P(T^p)_\varphi u, (T^p)_\varphi u)_0| + \|(T^p)_\varphi u\|_0^2.$$

A frequent complication comes from the position of a Z in the norm, and things become a bit more symmetric and simpler below if we introduce additional terms on the left, to make a more symmetric formulation with regard to the placement of the Z_j.

Namely, we will add $\sum_1^{2n} \|(T^p)_\varphi Z_j u\|_0^2$, which differs from the terms already there by a simple bracket, discussed in (3.31) and (3.32) above. Defining a more symmetric "norm" leads to the following definition.

Definition 4.1.

$$^s\|Z(T^p)_\varphi u\|_0 = \sum_j \|Z_j (T^p)_\varphi u\|_0 + \sum_j \|(T^p)_\varphi Z_j u\|_0.$$

We have, suppressing summations and indices,

$$\|(T^p)_\varphi u\|_{1/2}^2 + {}^s\|Z(T^p)_\varphi u\|_0^2 \lesssim |(P(T^p)_\varphi u, (T^p)_\varphi u)_0| + \|(T^p)_\varphi u\|_0^2. \tag{4.3}$$

Now we must commute P with $(T^p)_\varphi$ to obtain $(T^p)_\varphi Pu$, which is known, but the error we commit in doing so will be denoted be $|E|$, where, writing $\mathcal{A} = (T^p)_\varphi$ for the moment and $\mathcal{A}' = (T^{p-1})_{\varphi'}$,

$$\begin{aligned} E &= (P\mathcal{A}u, \mathcal{A}u)_0 - (\mathcal{A}Pu, \mathcal{A}u)_0 = ([P, \mathcal{A}]u, \mathcal{A}u)_0 \\ &= (Z[Z, \mathcal{A}]u, \mathcal{A}u)_0 + ([Z, \mathcal{A}]Zu, \mathcal{A}u)_0 \\ &\sim (\mathcal{A}'Zu, Z\mathcal{A}u)_0 + (\mathcal{A}'Zu, Z\mathcal{A}u)_0 \\ &\quad + \underline{C^p}\left(\frac{\varphi^{(p+1)}}{p!} Z^p u, Z\mathcal{A}u\right)_0 + \underline{C^p}\left(\frac{\varphi^{(p+1)}}{p!} Z^{p+1} u, \mathcal{A}u\right)_0, \end{aligned}$$

since here $Z = -Z^*$, and hence

$$|E| \lesssim s.c.^s \|Z(T^p)_\varphi u\|_0^2 + \ell.c.^s \|Z(T^{p-1})_{\varphi'} u\|_0^2 + \ell.c.\,\underline{C^p}\,{}^s\left\| Z\frac{\varphi^{(p+1)}}{p!} Z^p u\right\|_0^2.$$

In all, then,

$$\begin{aligned}&\|(T^p)_\varphi u\|_{1/2} + {}^s\|Z(T^p)_\varphi u\|_0\\ &\quad\lesssim \|(T^p)_\varphi Pu\|_0 + {}^s\|Z(T^{p-1})_{\varphi'} u\|_0 + \|(T^p)_\varphi u\|_0 + \underline{C^p}\,{}^s\left\| Z\frac{\varphi^{(p+1)}}{p!} Z^p u\right\|_0.\end{aligned} \tag{4.4}$$

We want to iterate this estimate, with p decreasing each time until $p = 0$. If we assume $Pu = 0$ for simplicity only in this iteration, we obtain

$$\begin{aligned}&\|(T^p)_\varphi u\|_{1/2} + {}^s\|Z(T^p)_\varphi u\|_0\\ &\quad\lesssim C^p\{{}^s\|Z\varphi^{(p)} u\|_0 + \|\varphi^{(p)} u\|_0\} + \sum_{\ell\le p} \underline{C^p}\,{}^s\left\| Z\frac{\varphi^{(p+1)}}{\ell!} Z^\ell u\right\|.\end{aligned} \tag{4.5}$$

4.1.2 Z Derivatives

In the previous section we reduced the estimation of T derivatives to the consideration of pure Z derivatives. As we will see, these will lead *back* to T derivatives, but only half as many, permitting us to finish the proof.

Let ψ have compact support, and consider what happens when we bracket a Z with a large number of Z's:

$$[Z, Z^q] = \underline{q}\, T Z^{q'}, \qquad q' \le q - 1, \tag{4.6}$$

meaning that there will occur *at most* q terms of the form $TZ^{q'}$ with $q' \le q - 1$, since the isolated Z might be, for example, X, and all the other Z's might be Y's or, at the opposite extreme, they could all be X's, in which case there would be no terms containing T.

The a priori estimate (3.24), with $v = \psi Z^q u$, then yields

$$\begin{aligned}\|\psi Z^q u\|_{1/2}^2 + \sum_1^{2n} \|Z_j \psi Z^q u\|_0^2 &\lesssim |(PZ^q\psi u, \psi Z^q u)_0| + \|\psi Z^q u\|_0^2\\ &= |(\psi Z^q Pu, \psi Z^q u)_0| + \|\psi Z^q u\|_0^2 + E,\end{aligned} \tag{4.7}$$

since $Z^* = -Z$. Writing $\mathcal{B} = \psi Z^q,\ \mathcal{B}_1 = \psi' Z^{q-1}$, and $\mathcal{B}_2 = \underline{q} T \psi' Z^{q-2}$, so that $[Z, \mathcal{B}] = Z\mathcal{B}_1 + Z\mathcal{B}_2$, we obtain

$$\begin{aligned} E &= (P\mathcal{B}u, \mathcal{B}u)_0 - (\mathcal{B}Pu, \mathcal{B}u)_0 = ([P, \mathcal{B}]u, \mathcal{B}u)_0 \\ &= (Z[Z, \mathcal{B}]u, \mathcal{B}u)_0 + ([Z, \mathcal{B}]Zu, \mathcal{B}u)_0 \\ &= ([Z, \mathcal{B}]u, Z\mathcal{B}u)_0 + ([Z, \mathcal{B}]Zu, \mathcal{B}u)_0 \\ &= (\{Z\mathcal{B}_1 + Z\mathcal{B}_2\}u, Z\mathcal{B}u)_0 + (\{\mathcal{B}_1 + \mathcal{B}_2\}Zu, Z\mathcal{B}u)_0. \end{aligned}$$

Again, with the symmetric norm,

$$^s\|Z\mathcal{B}.u\| = \|Z\mathcal{B}.u\|_0 + \|\mathcal{B}.Zu\|_0,$$

which introduces an error with q reduced by one and one derivative on ψ, and nothing else. Thus we have

$$\|\psi Z^q u\|_{1/2} + {}^s\|Z\psi Z^q u\| \lesssim \|\psi Z^{q-1} Pu\|_0 + {}^s\|Z\psi' Z^{q-1} u\| + \underline{q}\, {}^s\|Z\psi Z^{q-2} Tu\|, \tag{4.8}$$

where the second term on the right may include a term in which ψ is not differentiated (a milder term).

So two kinds of things may happen: a Z may differentiate ψ, or two Z's may create a q and a T. We will iterate this estimate until there are no more free Z's. The first time we obtained the factor q, and q is reduced by two; the next time will produce the factor $q - 2$ (or less) etc., but the only effective constant to capture this is q^m, since there are m terms, the largest of which is q. Later the distinction between factorials and powers of p will become important, but not at this moment.

We arrive at

$$^s\|Z\psi Z^q u\|_0 \lesssim \underline{C^q} \sup_{r+2m\le q} q^m \left\{\|\psi^{(r)} T^m Z^{q-2m-r} Pu\|_0 + \|\psi^{(r)} T^m u\|_0\right\}, \tag{4.9}$$

or, if we take $Pu = 0$,

$$^s\|Z\psi Z^q u\|_0 \lesssim \underline{C^q} \sup_{r+2m\le q} \frac{\|\psi^{(r)} T^m u\|_0}{m!}, \tag{4.10}$$

while (4.5) is, again,

$$\|(T^p)_\varphi u\|_{1/2} + {}^s\|Z(T^p)_\varphi u\| \lesssim C^p \sup_{\ell\le p} {}^s\left\| Z \frac{\varphi^{(p+1)}}{\ell!} Z^\ell u \right\|. \tag{4.11}$$

4.2 The End of the Proof for the Heisenberg Group

To conclude the proof (with $Pu = 0$), the above estimates telescope, especially if we make the observations that $(T^p)_\varphi = T^p$ where $\varphi \equiv 1$,

$$\|T^p u\|_{L^2(\{\varphi\equiv 1\})} \leq |||Z(T^p)_\varphi u||| + \|(T^p)_\varphi u\|_{1/2} \lesssim \sup_{q\leq p} \underline{C^p} \;{}^s\left\| Z\frac{\varphi^{(p+1)}}{q!} Z^q u\right\|,$$

$${}^s\|Z\psi Z^q \psi u\|_0 \lesssim \underline{C^q} \sup_{r+2m\leq q} q^m \|\psi^{(r)} T^m u\|_0. \tag{4.12}$$

Note that we cannot allow comparisons between $j!$ and N^j involving only C^j.

But if we are permitted C^N, all comparisons are permitted: taking jth roots and dividing through by j shows at once that

$$N^j \leq C^N j!, \quad j \leq N.$$

Since there are many choices involved here, for example the order of differentiation and localization, we should note that even comparing $T^m \psi^{(r)}$ with expressions in which all the derivatives are on the right is not hard:

$$T^m \psi^{(r)} = \sum_{j\leq m,r} \binom{m}{j}\binom{r}{j} j!\psi^{(r+j)} T^{(m-j)},$$

and hence the right-hand side of (4.12) remains unchanged if the derivatives are placed to the left of ψ, though with a different constant. Thus, symmetrizing things a bit between p and q by associating the factorial with the power of T yields

$$\frac{\|T^p u\|_{L^2(\{\varphi\equiv 1\})}}{p!} \leq C^p \sup_{q\leq p} \frac{|||Z^q \frac{\varphi^{(p+1)}}{p!} u|||}{q!} \leq C^p \sup_{r+2m\leq q\leq p} q^m \frac{\|\frac{\varphi^{(p+r+1)}}{p!} T^m u\|}{q!},$$

or

$$\begin{aligned}\frac{\|T^p u\|_{L^2(\{\varphi\equiv 1\})}}{p!} &\lesssim C^p \sup_{r+2m\leq q\leq p} \frac{q^m m!}{p!q!} \|\varphi^{(p+r+1)}\|_{L^\infty} \frac{\|T^m u\|_{L^2(\operatorname{supp}\varphi)}}{m!} \\ &\lesssim C^p \sup_{r+2m\leq q\leq p} \frac{q^m m! N^{p+r+1}}{p!q!} \frac{\|\varphi^{(p+r+1)}\|_{L^\infty}}{N^{p+r+1}} \frac{\|T^m u\|_{L^2(\operatorname{supp}\varphi)}}{m!} \\ &\lesssim C^p \sup_{r+2m\leq q\leq p} \frac{\|\varphi^{(p+r+1)}\|_{L^\infty}}{N^{p+r+1}} \frac{\|T^m u\|_{L^2(\operatorname{supp}\varphi)}}{m!},\end{aligned} \tag{4.13}$$

since for $p \sim N$, under the conditions of the supremum,

$$\frac{q^m m! N^{p+r+1}}{p! q!} \lesssim C^p. \tag{4.14}$$

The crucial observation here is that the number of T derivatives on the right is at most half the number that we started with. We need to introduce a new localizing function, φ_1, at this point, but it will never need to receive more than $p/2$ derivatives.

The next iteration will start with estimating up to $p/2$ derivatives, and lead eventually to the introduction of a third localizing function, φ_2, which will need to be able handle at most $p/4$ derivatives, etc; in other words, if we start with the aim of estimating $N = N_0$ derivatives in T (and in Z along the way), we will need at most $\log_2 N_0$ localizing functions, the first identically equal to 1 in ω_0, where we hope to prove analyticity of the solution, each identically 1 on the support of the previous one, and all supported in the open set $\widetilde{\omega}$, where we assume that the given data are analytic.

If we denote these by $\varphi_j = \varphi_{[N_j]}$, $j \leq \log_2 N$, and the corresponding open sets by ω_j if we need to refer to them, with distance d_j between the closure of ω_j and the complement of ω_{j+1}, then if we permit at most N_j derivatives on the localizing function φ_j, each derivative will contribute a factor $CN_j/d_j = (CN/2^j)/d_j$, and so we will have a constant *independent of* N such that

$$|D^k \varphi_j| \lesssim \left(\frac{CN/2^j}{d_j}\right)^k, \quad k \leq 2N_j = 2N/2^j.$$

We will write d_0 for the separation between the interior of the first and the exterior of the last of the open sets $\widetilde{\omega} = \omega_{\log_2 N_0}$.

It would be tempting to set $d_j = d_0/2^j$, so that each derivative would contribute the factor CN no matter which band or shell we were in. We shall see, however, that this will not work, and we will have to be more careful in our choice of d_j.

In the last estimate above (4.13) of course we might not want to use the *next* φ, since we want m to be comparable to N_j for the φ_j that we use in order that the $m!$ in the denominator will be strong enough to balance as many as m derivatives on the new localizing function.

But as long as we always do this, the analysis above will apply equally well to any φ_j in place of φ, and N_j in place of N, although the crucial terms in (4.13) will require the bounds (4.14) at level j.

More explicitly, since this will lead us to the condition we need on d_j, let's write down the corresponding estimates and conditions in the particular case that the supremum on the right occurred for $m \sim N_\ell$ having started with $p \sim N_0$. Then we have first of all

$$\frac{\|T^p u\|_{L^2(\{\varphi \equiv 1\})}}{p!} \lesssim C^{N_0} \sup_{r+2m \leq q \leq p} \frac{\|\varphi_0^{(p+r+1)}\|_{L^\infty}}{{N_0}^{p+r+1}} \frac{\|T^m u\|_{L^2(\operatorname{supp} \varphi_0)}}{m!} \tag{4.15}$$

and then

$$\frac{\|T^m u\|_{L^2(\{\varphi_\ell \equiv 1\})}}{m!} \lesssim C^{N_\ell} \sup_{\tilde{r}+2\tilde{m} \le \tilde{q} \le m} \frac{\|\varphi_\ell^{(m+\tilde{r}+1)}\|_{L^\infty}}{{N_\ell}^{m+\tilde{r}+1}} \frac{\|T^{\tilde{m}} u\|_{L^2(\operatorname{supp} \varphi_\ell)}}{\tilde{m}!}, \tag{4.16}$$

since now for $m \sim N_\ell$, under the conditions of the new supremum,

$$\frac{\tilde{q}^{\tilde{m}} \tilde{m}! {N_\ell}^{m+\tilde{r}+1}}{m! \tilde{q}!} \lesssim C^m, \tag{4.17}$$

just as we had $p \sim N$ under the conditions of the first supremum,

$$\frac{q^m m! N^{p+r+1}}{p! q!} \lesssim C^p. \tag{4.18}$$

It is very tempting to think that since the localizing function φ_ℓ behaves in nearly the same way that φ_0 did - derivatives bounded by powers of N_j/d_j, that the method would not lead to the desired bounds, for with d_j taken to be $d_0/2^j$, in the end we would have just powers of CN_0, which would not cancel the N_j which are present in the quotient

$$\frac{\|\varphi_\ell^{(m+\tilde{r}+1)}\|_{L^\infty}}{{N_\ell}^{m+\tilde{r}+1}}. \tag{4.19}$$

But suppose we ignore this seeming hitch, and continue to iterate (4.15) and (4.16). With $d_j = d_0/2^j$, $N_j = N_0/2^j$, we wind up with

$$\frac{\|T^p u\|_{L^2(\{\varphi \equiv 1\})}}{p!} \le \prod_{j=1}^{\log_2 N_0} C^{N_j} (2^j)^{\frac{N_0}{2^j}} \|u\|_{H^1(\tilde{\omega})} \le C^{N_0} \left(\prod_{j=1}^{\log_2 N_0} (2^j)^{\frac{1}{2^j}} \right)^{N_0} \|u\|_{H^1(\tilde{\omega})}, \tag{4.20}$$

since

$$\frac{\|\varphi_\ell^{(m+\tilde{r}+1)}\|_{L^\infty}}{{N_\ell}^{m+\tilde{r}+1}} \sim \frac{\left(\frac{CN_0/2^\ell}{d_0/2^\ell}\right)^{m+\tilde{r}+1}}{(N_0/2^\ell)^{m+\tilde{r}+1}} \sim (C 2^\ell/d_0)^{m+\tilde{r}+1}$$

with $\tilde{r} + m + 1 \le 2N_0/2^\ell$.

But in (4.20) the large parenthesis is bounded by a universal constant; hence the whole right-hand side is bounded by $C^{N_0} \|u\|_{H^1(\tilde{\omega})}$, which leads to

$$\frac{\|T^p u\|_{L^2(\{\varphi \equiv 1\})}}{p!} \le C^{N_0} \|u\|_{H^1(\tilde{\omega})}, \tag{4.21}$$

or

$$\|T^p u\|_{L^2(\{\varphi \equiv 1\})} \le C^{N_0} p! \|u\|_{H^1(\tilde{\omega})}, \tag{4.22}$$

which implies analyticity, since $N_0 \sim p$.

Chapter 5
Coefficients

5.1 How Special Is the Heisenberg Model?

The model operator studied above, a constant-coefficient quadratic expression in the left-invariant vector fields from the Heisenberg group, suggests two kinds of generalizations: variable-coefficient expressions in these vector fields and variable-coefficient expressions in smooth vector fields whose bracket relations and span resemble those of the Heisenberg group vector fields.

Fortunately, in the generality we are going to study first, this second generalization is a non-generalization, thanks to the celebrated theorem of Darboux. The Darboux theorem states that any collection of $2n-2$ smooth vector fields in $\mathbb{R}^{2n-1}$ whose characteristic variety is symplectic (i.e., the exterior derivative of the one-form ω defining the span of the vector fields is non-degenerate) then in suitable coordinates the vector fields may be written as (a smooth, invertible) linear combination of the Heisenberg group vector fields.

It follows that to allow such vector fields is not necessary if one can treat variable-coefficient operators built of the Heisenberg group fields.

And clearly the estimates we have worked with are stable under taking invertible linear combinations of the vector fields.

When one introduces (smooth) coefficients in the operator, the situation becomes more complicated. For now, the operator will be written (at least the principal part)

$$P = \sum_{j,k=1}^{n} a_{j,k} Z_j^* Z_k = -\sum_{j,k=1}^{n} a_{j,k} Z_j Z_k.$$

The vector fields will still be from the Heisenberg group as above, but as noted above, the coordinate change

$$t = \tilde{t} + \frac{1}{2}\sum \tilde{x}_j \tilde{y}_j, \quad x_j = \tilde{x}_j, \quad y_j = \tilde{y}_j$$

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_5,

transforms the vector fields $\{X_j, Y_j, T\}$ to have the form

$$X_j = \frac{\partial}{\partial \tilde{x}_j} - \tilde{y}_j \frac{\partial}{\partial \tilde{t}}, \quad Y = \frac{\partial}{\partial \tilde{y}_j}, \quad [X_j, Y_k] = \delta_{jk} \frac{\partial}{\partial \tilde{t}} = \tilde{T} = T. \tag{5.1}$$

The reason that we do this is that now the fields have a sort of commutation property that the original fields did not have:

$$XY^\gamma \rho = Y^\gamma \rho_{\tilde{x}} - \tilde{y} Y^\gamma \rho_{\tilde{t}}. \tag{5.2}$$

That is, while of course $[X, Y^\gamma] \neq 0$ in general, the *derivatives* from X may be passed onto the function ρ if we may leave an additional function (here $\tilde{y}$) with the function ρ (on its right). We will henceforth drop the tildes.

Thus in addition to the brackets we have considered above, we will need to consider brackets of errors of $(T^p)_\varphi$ with these coefficients. The formulas that follow will not be simple, but I trust you will agree that they cannot be avoided. They will have a logic that will become clear as we work with them.

And it turns out that working with the a priori estimate in norm form, not inner product, simplifies the construction and calculations.

5.2 Rigid Coefficients: *T* Derivatives

We start, however, with the calculations when the coefficients are independent of the variable t, a situation called "rigid" in the literature. Letting the coefficients depend on t complicates the calculations significantly, so we will reserve it for the next section.

For rigid coefficients, we have

$$\begin{aligned}
[(T^p)_\varphi, a]u &= \left[\sum_{|\alpha|+|\beta|\le p} \frac{(-X)^\alpha Y^\beta(\varphi)}{\alpha!\beta!} X^\beta Y^\alpha T^{2-|\alpha|-|\beta|}, a \right] u \\
&= \sum_{\substack{\alpha+\beta\le p \\ \alpha_1\le\alpha, \beta_1\le\beta \\ 1\le|\alpha_1+\beta_1|}} \binom{\alpha}{\alpha_1}\binom{\beta}{\beta_1} (X^{\beta_1} Y^{\alpha_1} a) \circ \\
&\quad \circ \frac{(-X)^\alpha Y^\beta(\varphi)}{\alpha!\beta!} X^{\beta-\beta_1} Y^{\alpha-\alpha_1} T^{p-|\alpha|-|\beta|} u \\
&= \sum_{\substack{\alpha+\beta\le p \\ \alpha_1\le\alpha, \beta_1\le\beta \\ 1\le|\alpha_1+\beta_1|}} \frac{(X^{\beta_1} Y^{\alpha_1} a)}{\alpha_1!\beta_1!} \circ \frac{((-X)^{\alpha_1}(-X)^{\alpha-\alpha_1} Y^{\beta-\beta_1}(Y^{\beta_1}\varphi))}{(\alpha-\alpha_1)!(\beta-\beta_1)!} \\
&\quad \times X^{\beta-\beta_1} Y^{\alpha-\alpha_1} T^{p-|\alpha_1+\beta_1|-|\alpha-\alpha_1+\beta-\beta_1|} u.
\end{aligned} \tag{5.3}$$

We will take α_1 and β_1 as parameters here, and note that this expression is *almost* equal to

$$\sum_{\substack{\alpha+\beta\leq p\\ \alpha_1\leq\alpha,\beta_1\leq\beta\\ 1\leq|\alpha_1+\beta_1|}} \frac{(X^{\beta_1}Y^{\alpha_1}a)}{\alpha_1!\beta_1!}(T^{p-|\alpha_1+\beta_1|})_{(-X)^{\alpha_1}Y^{\beta_1}\varphi}u,$$

which would be lovely if exactly true.

The problem is that the extra $(-X)^{\alpha_1}$ derivatives on φ, α_1 of them, are in the wrong place in (5.3), and moving them directly onto φ would produce brackets between $(-X)$ and Y, yielding powers of T, harmless in themselves, but simultaneously decreasing the value of $\beta-\beta_1$, which is definitely not harmless, for that quantity occurs in three other places in the formula and the balance would be badly thrown off.

But I invite the reader to consider just one of these *apparently* extra $(-X)$ derivatives. The change of variables we have employed above in (5.1) makes two nice things happen that permit us to deal with the poorly placed $(-X)$.

Namely, we split $(-X)$ into two parts: the $\partial/\partial x$ part, which may "slide" directly onto φ, and the $y\partial/\partial t$ part, and here we leave the function y where it is, to the left of everything, but pass the $\partial/\partial t$ derivative onto φ as well, since $\partial/\partial t$ commutes with everything, and in any case, the *vector fields themselves* are rigid even if the coefficients are not, a fact we shall use in the non-rigid-coefficient case in the next section.

So we may write (5.3) as

$$[(T^p)_\varphi, a] = \sum_{\substack{\alpha+\beta\leq p\\ \alpha_1'\leq\alpha_1\leq\alpha,\beta_1\leq\beta\\ 1\leq|\alpha_1+\beta_1|}} \frac{(X^{\beta_1}Y^{\alpha_1}a)}{\alpha_1!\beta_1!}\circ y^{\alpha_1'}\frac{((-X)^{\alpha-\alpha_1}Y^{\beta-\beta_1}(D^{|\alpha_1|}Y^{\beta_1}\varphi))}{(\alpha-\alpha_1)!(\beta-\beta_1)!}$$
$$\times X^{\beta-\beta_1}Y^{\alpha-\alpha_1}T^{p-|\alpha_1+\beta_1|-|\alpha-\alpha_1+\beta-\beta_1|}u, \tag{5.4}$$

or

$$[(T^p)_\varphi, a]u = \sum_{\substack{1\leq|\alpha_1+\beta_1|\leq p\\ \alpha_{1'}\leq\alpha_1}} y^{\alpha_{12}}\frac{(X^{\beta_1}Y^{\alpha_1}a)}{\alpha_1!\beta_1!}(T^{p-|\alpha_1+\beta_1|})_{\varphi^{(\alpha_1+\beta_1)}}u, \tag{5.5}$$

and we notice that the additional derivatives on φ are balanced by the drop in exponent of $(T^p)_\varphi$.

We encourage the reader to absorb this process here, since it will occur in more complicated form when we consider nonrigid coefficients and pseudodifferential coefficients below.

If we permit ourselves to confuse multi-indices and ordinary indices, and up to constants to the power of the index they are comparable, we may write this last expression as

$$[(T^p)_\varphi, a]u \sim \sum_{\substack{1\leq r\leq p\\ s\leq r}} y^s\frac{a^{(r)}}{r!}(T^{p-r})_{\varphi^{(r)}}u. \tag{5.6}$$

The intuitive sense in this expression, which one can certainly not gainsay, is that the degree has decreased by at least one, with the corresponding passing to the localizing function, and some powers of y, and when measured in L^2 norm, the powers of y will come out bounded by 1, and the derivatives on a divided by $r!$ bounded by C^r, so that

$$\left\|[(T^p)_\varphi, a]u\right\|_0 \lesssim \sum_{1\le r\le p} C^r \left\|(T^{p-r})_{\varphi^{(r)}}u\right\|_0. \tag{5.7}$$

5.3 Our Estimates and How We Use Them

To keep things in one place, we recall (5.8), (5.9), the results of brackets with the vector fields:

$$[X_\ell, (T^p)_\varphi] \equiv 0 \mod \frac{Z^{p+1}\varphi}{p!} \circ Z^p \tag{5.8}$$

and

$$[Y_\ell, (T^p)_\varphi] \equiv (T^{p-1})_{\varphi'} \circ Y_\ell \mod \frac{Z^{p+1}\varphi}{p!} \circ Z^p. \tag{5.9}$$

These are clearly the relations that will permit us to iterate the a priori estimate with control on the terms.

But first we will collect the a priori estimates in the forms in which we have found they work most simply: that without inner products at all solves nearly every situation, but eventually we will need the inner product version as well.

In the model case in the previous chapter, we have dealt with the position of the vector fields Z needed for the estimates by symmetrizing the norms, and denoted these by ${}^s\|.\|$. Rather than continue to use these norms now, we will simply not worry about the precise position of the Z's in a norm, since we know how to evaluate the bracket of Z with $(T^p)_\varphi$, and so for now we will use "norms" where it makes no difference: all positions are present. However, and this is important, we will continue to use a norm notation. Thus

$$\left\|Z^2(T^p)_\varphi w\right\|_0 \text{ stands for } \left\|Z^2(T^p)_\varphi w\right\|_0, \ \left\|Z(T^p)_\varphi Zw\right\|_0 \text{ or } \left\|(T^p)_\varphi Z^2 w\right\|_0 \tag{5.10}$$

or a sum of all three, and likewise

$$\left\|Z(T^p)_\varphi w\right\|_0 \text{ stands for } \left\|Z(T^p)_\varphi w\right\|_0 \text{ or } \left\|(T^p)_\varphi Zw\right\|_0, \tag{5.11}$$

and, perhaps most potentially ambiguously,

$$\left\|Z^2\frac{\varphi^{(a)}}{b!}Z^rT^sw\right\|_0 \text{ stands for } \left\|Z^2\frac{\varphi^{(a)}}{b!}Z^rT^sw\right\|_0,$$

$$\left\|Z\frac{\varphi^{(a)}}{b!}ZZ^rT^sw\right\|_0, \text{ or } \left\|\frac{\varphi^{(a)}}{b!}Z^2Z^rT^sw\right\|_0, \tag{5.12}$$

and

$$\left\| Z\frac{\varphi^{(a)}}{b!}Z^r T^s w \right\| \quad \text{stands for} \quad \left\| Z\frac{\varphi^{(a)}}{b!}Z^r T^s w \right\|_0 \quad \text{or} \quad \left\| \frac{\varphi^{(a)}}{b!}ZZ^r T^s w \right\|_0 .$$

The reason we feel this is both effective and safe is that these terms differ by terms of one of these forms *but of lower order* in a way we will quantify, and which will appear on the other side of the inequality anyhow.

In exchange for this flexibility of the position of Z derivatives, we pay a small price, namely in bounds for $Z^2(T^p)_\varphi u$ we must include a term with $p = 0$ on the right, in view of (3.31) and (3.32). Thus the estimates will read

$$\left\|Z^2(T^p)_\varphi u\right\|_0 \lesssim \left\|P(T^p)_\varphi u\right\|_0 + \left\|Z^2(T^{p-1})_{\varphi'} u\right\|_0 + \underline{C^p}\left\| Z\frac{\varphi^{(p+1)}}{p!}Z^p u \right\|_0, \quad (5.13)$$

$$\left\|Z(T^p)_\varphi u\right\|_0^2 \lesssim |(P(T^p)_\varphi u, (T^p)_\varphi u)_0| + \left\|Z(T^{p-1})_{\varphi'} u\right\|_0^2 + \underline{C^p}\left\| \frac{\varphi^{(p+1)}}{p!}Z^p u \right\|_0^2, \quad (5.14)$$

$$\left\|Z^2\varphi Z^q T^s u\right\|_0 \lesssim \|P\varphi Z^q T^s u\|_0 + \left\|Z^2\varphi^{(')} Z^{q-1} T^s u\right\|_0, \quad (5.15)$$

$$\|Z\varphi Z^q T^s u\|_0^2 \lesssim |(P\varphi Z^q T^s u, \varphi Z^q T^s u)_0| + \left\|Z\varphi' Z^q T^s u\right\|_0^2. \quad (5.16)$$

This notation means that we may write, using (5.7), and now with the further understanding that even the expressions involving Z's are not sensitive to the position of the leftmost two Z's, so that $Z^2\varphi w$ and $Z\varphi Z w$, and $\varphi Z^2 w$, for example, could be written in any of these ways.

Proposition 5.1.

$$[P, (T^m)_\varphi]w = [a, (T^m)_\varphi]Z^2 w + a[Z^2, (T^m)_\varphi]w$$

$$= \sum_{1\le r\le m} \frac{a^{(r)}}{r!}(T^{m-r})_{\varphi^{(r)}} Z^2 w + aZ(T^{m-1})_{\varphi'} Z w + \underline{C^m} Z\frac{\varphi^{(m+1)}}{m!}Z^m w.$$

This proposition immediately yields the following:

Proposition 5.2.

$$\left\|[P, (T^m)_\varphi]w\right\|_0 \lesssim \left\|[a, (T^m)_\varphi]Z^2 w\right\|_0 + \left\|[Z^2, (T^m)_\varphi]w\right\|_0$$

$$\lesssim \sum_{1\le r\le m} C^r \left\|Z^2(T^{m-r})_{\varphi^{(r)}} w\right\|_0 + \underline{C^m}\left\| Z\frac{\varphi^{(m+1)}}{m!}Z^m w \right\|_0^2.$$

For later reference, we establish the estimates from using inner product norms, even though for rigid coefficients this will not be needed:

Proposition 5.3.

$$\begin{aligned}|([P,(T^m)_\varphi]u,(T^m)_\varphi u)_0| \le \sum_{1\le r\le m} &\left|\left(\frac{a^{(r)}}{r!}(T^{m-r})_{\varphi^{(r)}}Z^2u,(T^m)_\varphi u\right)_0\right| \\ &+ |(aZ(T^{m-1})_{\varphi'}Zu,(T^m)_\varphi u)_0| \\ &+ \underline{C^m}\left|\left(Z\frac{\varphi^{(m+1)}}{m!}Z^m u,(T^m)_\varphi u\right)_0\right|,\end{aligned}$$

where in the first two terms on the right at most one of the Z's in Z^2 is actually to the left of $(T^p)_\varphi$, and it may be integrated by parts, modulo lower-order terms. Bringing one Z to the left and integrating it by parts may introduce additional, lower-order, errors, but still, in view of the analyticity of the coefficients a, this yields the following:

Proposition 5.4.

$$\begin{aligned}|([P,(T^m)_\varphi]u,(T^m)_\varphi u)_0| \le s.c.\,\|Z(T^m)_\varphi u\|_0^2 + \ell.c.\sum_{1\le r\le m} C^r\,\|(T^{m-r})_{\varphi^{(r)}}Zu\|_0^2 \\ +\ell.c.\,\underline{C^m}\,\left\|Z\frac{\varphi^{(m+1)}}{m!}Z^m u\right\|_0^2.\end{aligned}$$

Eventually, using the norm or the inner product estimates, the exponent m in $(T^m)_\varphi$ will drop to zero, as in the last term above, and thus we need to study pure Z and mixed derivatives.

5.4 Pure Z and Mixed Derivatives

To estimate pure Z derivatives, the situation is simpler except that two Z's may bracket to produce a T, or, of course, a Z may land directly on the localizing function, and upon iteration one or the other of these will happen (or the Z's will land on Pu, of course), until there are no more.

Proposition 5.5.

$$\begin{aligned}[P,\psi Z^qT^s]w &\sim \psi[a,Z^qT^s]Z^2w + a[Z^2,\psi Z^qT^s]w \\ &\sim \sum_{\substack{q'\le q, s'\le s\\ 1\le q'+q'}}\binom{q}{q'}\binom{s}{s'}a^{(q'+s')}Z^2\psi Z^{q-q'}T^{s-s'}w \\ &\quad + \underline{q}aZ^2\psi Z^{q-2}T^{q+1}w + aZ^2\psi'Z^{q-1}T^sw + aZ^2\psi''Z^{q-2}T^sw.\end{aligned}$$

Thus we have the following result.

Proposition 5.6.

$$\begin{aligned}\|[P,\psi Z^q T^s]w\|_0 &\lesssim \left\|\psi[a,Z^qT^s]Z^2w\right\|_0+\left\|[Z^2,\psi Z^qT^s]w\right\|_0\\ &\lesssim \sup_{\substack{q'\le q,s'\le s\\ 1\le s'+q'}} \frac{q!s!}{(q-q')!(s-s')!}C_a^{q'+s'}\left\|Z^2\psi Z^{q-q'}T^{s-s'}w\right\|_0\\ &\quad+\underline{q}\left\|Z^2\psi Z^{q-2}T^{q+1}w\right\|_0+\left\|Z^2\psi' Z^{q-1}T^sw\right\|_0\\ &\quad+\left\|Z^2\psi'' Z^{q-2}T^sw\right\|_0.\end{aligned}$$

So, using Proposition 5.6,

$$\begin{aligned}\left\|Z^2\psi Z^qT^su\right\|_0 &\lesssim \|P\psi Z^qT^su\|_0+\left\|Z^2\psi^{(\prime)}Z^{q-1}T^su\right\|_0\\ &\lesssim \|\psi Z^qT^sPu\|_0+\|[P,\psi Z^qT^s]u\|_0+\left\|Z^2\psi^{(\prime)}Z^{q-1}T^su\right\|_0\\ &\lesssim \|\psi Z^qT^sPu\|_0+\sup_{\substack{q'\le q,s'\le s\\ 1\le s'+q'}}\frac{q!s!}{(q-q')!(s-s')!}C_a^{q'+s'}\\ &\quad\times\left\|Z^2\psi Z^{q-q'}T^{s-s'}u\right\|_0+\underline{q}\left\|Z^2\psi Z^{q-2}T^{q+1}w\right\|_0\\ &\quad+\left\|Z^2\psi' Z^{q-1}T^sw\right\|_0+\left\|Z^2\psi'' Z^{q-2}T^su\right\|_0,\end{aligned}$$

where naturally the constant will depend on the (real analytic) coefficients.

That is, one of several things has happened:

- P and u are together, as Pu, which is known, or
- q' Z's have been gained with a factor of $q!/(q-q')!(\le q^{q'})$, or
- one or two Z derivatives were gained by differentiating ψ, or
- two powers of Z were gained, producing a T *and* a factor of q.

Iteration of this estimate, as long as at least two Z's remain, may replace Z's by half as many (new) T's. When there is at most one Z, we stop.

So we have, back in terms of L^2 norms, since $T=[Z,\bar{Z}]$, using (5.13) and Proposition 5.6 and iterating,

$$\begin{aligned}\left\|Z^2T^pu\right\|_{L^2(\{\varphi\equiv1\})}+\left\|T^{p+1}u\right\|_{L^2(\{\varphi\equiv1\})} &\lesssim \left\|Z^2(T^p)_\varphi u\right\|_0\\ &\lesssim \left\|P(T^p)_\varphi u\right\|_0+\left\|Z^2(T^{p-1})_{\varphi'}u\right\|_0+\underline{C^p}\left\|Z\frac{\varphi^{(p+1)}}{p!}Z^pu\right\|_0,\\ &\lesssim \sum_{0\le r\le p}C^r\{\left\|(T^{p-r})_{\varphi^{(r)}}Pu\right\|_0+\left\|Z\varphi^{(p+1)}Z^ru\right\|_0/r!\}\end{aligned} \tag{5.17}$$

and

$$\left\|Z^2\psi Z^q T^b\right\|_0 \lesssim \sup_{q_1+2q_2=q} C^{(q)} q^{q_2} \left\|Z^{<2}\psi^{(q_1)} T^{b+|q_2|} u\right\|_0 + \sup_{q_1+2q_2\le q} C^{|q_2|} q^{|q_2|} \left\|\psi^{(|q_1|)} Z^{q-q_1-2q_2} T^{b+|q_2|} Pu\right\|_0 .$$

This last estimate may be iterated effectively as long as there remain two Z's, at which point we simply stop.

Consider what happens when $Pu = 0$ (for simplicity). Then, since modulo C^p, $r!$ and p^r are comparable, the circuit reads

$$\left\|Z^{\le 2} T^p u\right\|_{L^2(\{\varphi\equiv 1\})} \le C^p \sup_{p_1+2p_2\le p+1} \frac{p^{p_2}|\varphi^{(p_1+p+1)}|_{L^\infty}}{p^{p_1+2p_2}} \left\|Z^{\le 2} T^{p_2} u\right\|_{L^2(\operatorname{supp}\varphi)},$$

or

$$\left\|Z^{\le 2} T^p u\right\|_{L^2(\{\varphi\equiv 1\})} \le C^p \sup_{p_1+2p_2\le p+1} \frac{|\varphi^{(p_1+p+1)}|_{L^\infty}}{p^{p_1}} \frac{\left\|Z^{\le 2} T^{p_2} u\right\|_{L^2(\operatorname{supp}\varphi)}}{p^{p_2}}.$$

Yet again, more suggestively,

$$\frac{\left\|Z^{\le 2} T^p u\right\|_{L^2(\{\varphi\equiv 1\})}}{p^p} \le C^p \sup_{p_1+2p_2\le p+1} \frac{|\varphi^{(p_1+p+1)}|_{L^\infty}}{p^{p_1+p+1}} \frac{\left\|Z^{\le 2} T^{p_2} u\right\|_{L^2(\operatorname{supp}\varphi)}}{p^{p_2}}. \tag{5.18}$$

Now, φ was chosen with reference to p, our starting point. We will switch to a new localizing function at this point, but in handling the quotient $|\varphi^{(p_1+p+1)}|_{L^\infty}/p^{p_1+p+1}$ well we realize that the next localizing function will have to be chosen not only to be identically equal to one on the support of φ but also linked to the remaining number of derivatives, p_2, which is at most equal to $(p+1)/2$.

This will introduce a new quotient and so on, and the product of them all, clearly at most $\log_2 N$ of them, will need to be bounded by (another) universal constant raised to the power p.

And of course when the supports of the φ_j are lined up, starting with $p = N$, the requirements on the φ_j are:

- they must be nonnegative with integral equal to 1,
- denoting by d_j the width of the band in which φ_j drops from being $\equiv 1$ to 0, roughly $\operatorname{dist}(\operatorname{supp}\varphi_j, (\operatorname{supp}\varphi_{j+1})^c)$, then

$$\sum d_j = \operatorname{dist}(\omega_0, (\omega_1)^c) = d \ll 1,$$

- φ_j may receive as many as $2p_j$ derivatives ($\sum p_j = N, p_j \le (p_{j-1}+1)/2$) each roughly proportional to p_j/d_j in L^∞ norm,

- in a worst-case scenario, iterating (5.18) $\log_2 N$ times, we will find that (5.18) is replaced by

$$\frac{\left\|Z^{\leq 2}T^{p}u\right\|_{L^2(\{\varphi\equiv 1\})}}{p^p} \leq \prod_{j=1}^{\log_2 N} C^{p_j} \sup_{p'_j+2p''_j\leq p_j+1} \frac{\left|\varphi^{(p'_j+p_j+1)}\right|_{L^\infty}}{p_j^{p'_j+p_j+1}} \|u\|_{H^2(\operatorname{supp}\varphi)}.$$

Thus, if we iterate this last estimate, (5.18), $\log_2 N$ times, starting originally with $p = N$ but starting the jth iteration with p_j and φ_j, the ultimate estimate is

$$\frac{\left\|Z^{\leq 2}T^{p}u\right\|_{L^2(\omega_0)}}{p^p} \leq C \left(\prod_{2p_j\leq p_{j-1}+1} C^{p_j} \sup_{r_j\leq 2p_j+1} \frac{\left|\varphi_j^{(r_j)}\right|_{L^\infty}}{p_j^{r_j}} \right) \|u\|_{H^2(\omega_1)}. \quad (5.19)$$

As above, in (4.20), (4.21), and (4.22), we find that we need not choose the p_j and φ_j in a delicate manner. Merely taking $p_j = N/2^j$ and the φ supported in a band of width $d/2^j$ will suffice to demonstrate analyticity, even though each derivative on a localizing function still contributes a factor of N that does not behave well when divided by $p_j^{r_j}$.

The miracle is that in the product, dealing with terms whose powers decrease by a factor of at least two each time, this effect goes away:

$$\left(\prod_{j=1}^{\log_2 N_0} (2^j)^{\frac{1}{2^j}} \right)^{N_0} \leq \left(\prod_{j=1}^{\infty} (2^j)^{\frac{1}{2^j}} \right)^{N_0} \leq C^{N_0}, \quad (5.20)$$

even though

$$\left(\prod_{j=1}^{\log_2 N_0} (n)^{\frac{1}{n}} \right)^{N_0} \quad (5.21)$$

is not bounded by any constant raised to the power N_0 uniformly in N_0, since

$$\log \prod_{j=1}^{\infty} (n)^{\frac{1}{n}} = \sum_{1}^{\infty} \frac{\log n}{n}$$

diverges.

5.5 Formal Observations

Rather than follow the cascade of terms down to the norm of u with only a couple of derivatives, and to give a flavor of what will be required later, we could make just one iteration of the a priori estimate and keep track of the size of the result and make sure that we could iterate the estimate, or if we could not, describe what needed to be done.

In the rigid case this may seem silly, since in the expressions below several combinations of the indices are and will remain equal to zero.

But in the nonrigid case they will need to be considered, and thus the reader is encouraged to read this rephrasing of the rigid proof with an eye to the next section.

Here is the strategy. Given an expression of the sort we encounter, or a sum of them, since after an application of an inequality we have a sum of terms, we will try to systematize what happens to that expression under an application of the inequality.

This will be schematized so that upon iteration, the expression becomes better and better. And it is true that the order "drops," yet in order to be able to continue another time, we need to maintain a pair of Z's. Eventually this will fail to be the case, since the estimate, with $Z^2\varphi Z T^s u$, on the left, for example, will lead to $[P, \varphi Z T^s]u$, which will contain the term $\varphi Z T^{s+1}u$ where we must halt.

Where will we be when this occurs? We have addressed this question in a slightly hand-waving kind of way above. Here we would like to formalize that argument, and keep track of the terms that arise.

We will do many things at once. We will start with a rather general expression, with variable coefficients, denoted by

$$\mathcal{G} = \sum_A G_{A,\psi} = \sum_A F_A (T^m)_{\psi^{(r)}} Z^q T^s u \tag{5.22}$$

with each $A = A_{(q,s,m,r)}$.

To formalize the idea that the order drops, we denote the "order" of this expression by

$$|A| = q + s + m = |G_{A,\psi}| < N, \tag{5.23}$$

and for later use, will simultaneously keep track of an "anisotropic" order of A that "permits" two Z's to generate a T in a special, and natral, fashion. To wit, we define

$$\|A\| = |A| + s = \left\| G_{A,\psi} \right\|, \tag{5.24}$$

which awards T derivatives double weight (though not in $(T^p)_\varphi$.)

The only use to which we will put this double-bar norm is the following: *if* we can keep applying the estimate over and over and find that the order ($|A|$) drops each time while the double-bar order ($\|A\|$) does not increase, then if we were to start with $Z^2 T^a u$ to estimate in L^2 norm in one open set (observe that $(T^a)_\varphi = T^a$ in such an open set if $\varphi \equiv 1$ there, and then proceed with the estimate many times from $Z^2(T^a)_\varphi u$) and reached a term where a had been reduced to zero (say as the "error" $ZZ^{a'}, a' \leq a$) in brackets of $(T^a)_\varphi$ with Z) and continued to apply the estimate until the terms no longer contained two Z's, we would have only $(Z)\varphi^{(r)}T^{\tilde{a}}$ inside the norm, we could conclude that $\tilde{a} \leq a'/2 \leq a/2$ by the property of double-weighting free T derivatives in the overall passage from $Z^2(T^a)_\varphi$ to $(Z)\varphi^{(r)}T^{\tilde{a}}$ (again, T derivatives in $(T^a)_\varphi$ are not weighted doubly even in $\|A\|$).

This means that if we start with nearly pure T derivatives in one open set and come back to nearly pure T derivatives in this way, there will be at most half as many.

Now there is another measure we need to introduce, and one that has a quirky element.

In order not to have to keep track of all the constants that are generated in each term using the estimate, we devise a measure of the size of the term(s) before and after an application of the estimate. And we conclude that the size has not grown, assuming that certain constants are well chosen relative to one another (this is a method that the author has used many times before, but which is hardly original with him).

Thus with $\mathcal{G}$ as immediately above,

$$|||\mathcal{G}|||_N = \sup_{A, \mathrm{supp}\,(\psi)} |F_A|\, K_0^m K_1^q K_2^s N^{|A|+r+m}/m!, \tag{5.25}$$

where K_0, K_1, K_2 will be chosen later, large relative to other constants that enter and subject to

$$K_0 \gg K_1 \gg K_2 \gg 1. \tag{5.26}$$

The flow will be to apply the a priori estimate repeatedly, as long as there are at least two Z's (3.11) or one Z (3.10), and track the behavior of the sizes of the indices A that enter, noting that $|A|$ will decrease and $\|A\|$ will not increase, roughly speaking, and that the norms $|||\mathcal{G}|||_N$ will also not increase. When the constants K_j are subject to (5.26), a transfer of a derivative from $(T^m)_\psi$ to Z (or T) will introduce a factor K_1/K_0 or K_2/K_0, both less than one, and when two Z's bracket to produce a T, the factor will be K_2/K_1^2, also less than one.

As long as $q \geq 2$, we may use the a priori estimate optimally.

Thus we will start with pure powers of T, say $p+1$ of them, and an open set ω, write one T as $[Z, Z]$ and then pass to $(T^p)_\varphi$ with $\varphi \equiv 1$ on a neighborhood of ω. This means estimating $\|T^{p+1}u\|_{L^2(\omega)} \leq \|(T^p)_\varphi Z^2 u\|$, and we will study the effects of the estimate on this initial

$$G_{A_{(2,0,p,0)}} = (T^p)_\varphi Z^2 u, \tag{5.27}$$

for which the norms are

$$|A_{(2,0,p,0)}| = \left\|A_{(2,0,p,0)}\right\| = p+2 \tag{5.28}$$

and

$$|||G_{A_{(2,0,p,0)}}|||_N = K_0^p K_1^2 N^{2p+2}/p!. \tag{5.29}$$

Proposition 5.7. *The first iteration of the a priori estimate will yield terms*

$$[P, (T^p)_\varphi]u = [f, (T^p)_\varphi]Z^2 + f[Z^2, (T^p)_\varphi]$$

with

$$[f, (T^p)_\varphi]Z^2u = \sum_{\substack{1\le r\le p\\ s\le r}} y^s \frac{f^{(r)}}{r!}(T^{p-r})_{\varphi^{(r)}} Z^2u = \sum_{\substack{1\le r\le p\\ s\le r}} y^s \frac{f^{(r)}}{r!} G_{A(2,0,p-r,r)=A^{1,r}} \tag{5.30}$$

with

$$\left\|[f, (T^p)_\varphi]Z^2u\right\|_0 \le \sum_{\substack{1\le r\le p\\ s\le r}} \left\|y^s \frac{f^{(r)}}{r!} G_{A^{1,r}}u\right\|_0 \le \sum_{\substack{1\le r\le p\\ s\le r}} C_f^r \left\|(T^{p-r})_{\varphi^{(r)}}u\right\|_0 \tag{5.31}$$

and

$$\begin{aligned} f[(T^p)_\varphi, Z^2]u &\equiv f(T^{p-1})_{\varphi'} Z^2u + f\frac{\varphi^{(p+1)}}{p!} \circ Z^{p+1}u \\ &= fG_{A^2_{(2,0,p-1,1)}} + fG_{A^3_{(p+1,0,0,p+1)}} \end{aligned} \tag{5.32}$$

with

$$\begin{aligned} \left\|f[(T^p)_\varphi, Z^2]u\right\|_0 &\le \left\|fG_{A^2_{(2,0,p-1,1)}}u\right\|_0 + \left\|fG_{A^3_{(p+1,0,0,p+1)}}u\right\|_0 \\ &\le C_f\left\|(T^{p-1})_{\varphi'}Z^2u\right\|_0 + C_f\left\|\frac{\varphi^{(p+1)}}{p!} \circ Z^{p+1}u\right\|_0, \end{aligned} \tag{5.33}$$

for which (with $r \ge 1$)

$$\begin{aligned} |A^{1,r}| &= \left\|A^{1,r}\right\| = 2 + p - r, \\ |A^2_{(2,0,p-1,1)}| &= \left\|A^2_{(2,0,p-1,1)}\right\| = p + 1, \\ |A^3_{(p+1,0,0,p+1)}| &= \left\|A^3_{(p+1,0,0,p+1)}\right\| = p + 1, \\ |||s\frac{f^{(r)}}{r!}G_{A^{1,r}}|||_N &= C_f^r K_0^{p-r} K_1^2 N^{p+2+p-r}/(p-r)!, \\ |||fG_{A^2_{(2,0,p-1,1)}}|||_N &= C_f K_0^{p-1} K_1^2 N^{p+2+p-1}/(p-1)!, \\ |||fG_{A^3_{(p+1,0,0,p+1)}}|||_N &= C_f K_1^{p+1} N^{p+1+p+1}/p!. \end{aligned}$$

For $s \le N$, we have

$$N^{s-1}/(s-1)! \le N^s/s!. \tag{5.34}$$

Thus all of the norms of A have dropped, and the $|||.|||_N$ norms have not increased, assuming that the relations between the K_j above are observed and all K are taken large relative to C_f.

Subsequent iterations will have small changes, but these illustrate the reason that the formal norms are defined as they are. The above relations will be replaced by those in the following proposition.

Proposition 5.8. *In subsequent iterations of the a priori estimate we have, instead of (5.31) and (5.33),*

$$[P,(T^p)_\varphi]u = [f,(T^p)_\varphi]Z^2 + f[Z^2,(T^p)_\varphi]$$

with

$$\begin{aligned}[f,(T^{p-r})_{\varphi^{(r)}}]Z^2u &= \sum_{\substack{1\le r'\le p-r\\ s\le r'}} y^s \frac{f^{(r')}}{r'!}(T^{p-r-r'})_{\varphi^{(r+r')}}Z^2u\\ &= \sum_{\substack{1\le r\le p-r\\ s\le r'}} y^s \frac{f^{(r')}}{r'!}G_{A_{(2,0,p-r-r',r+r')}=A^{1,r+r'}},\end{aligned} \tag{5.35}$$

for which

$$\begin{aligned}\left\|[f,(T^{p-r})_{\varphi^{(r)}}]Z^2u\right\|_0 &\le \sum_{\substack{1\le r'\le p-r\\ s\le r'}} \left\| y^s\frac{f^{(r')}}{r'!}G_{A^{1,r+r'}}u\right\|_0\\ &\le \sum_{\substack{1\le r'\le p-r\\ s\le r'}} C_f^{r'}\left\|(T^{p-r-r'})_{\varphi^{(r+r')}}Z^2u\right\|_0\end{aligned} \tag{5.36}$$

and

$$\begin{aligned}f[(T^{p-r})_{\varphi^{(r)}},Z^2]u &\equiv f(T^{p-r-1})_{\varphi^{(r+1)}}Z^2u + f\underline{C^{p-r}}\frac{\varphi^{(p+1)}}{(p-r)!}Z^{p-r+1}u\\ &= fG_{A^2_{(2,0,p-r-1,r+1)}} + fG_{A^3_{(p-r+1,0,0,p+1)}}\end{aligned} \tag{5.37}$$

with

$$\begin{aligned}&\left\|f[(T^{p-r})_{\varphi^{(r)}},Z^2]u\right\|_0\\ &\le \left\|fG_{A^2_{(2,0,p-r-1,r+1)}}u\right\|_0 + \left\|f\underline{C^{p-r}}G_{A^3_{(p-r+1,0,0,p+1)}}u\right\|_0\\ &\le C_f\left\|(T^{p-r-1})_{\varphi^{(r+1)}}Z^2u\right\|_0 + C_fC^{p-r}\left\|\frac{\varphi^{(p+1)}}{(p-r)!}Z^{p-r+1}u\right\|_0,\end{aligned} \tag{5.38}$$

for which (with $r \geq 1$)

$$
\begin{aligned}
|A^{1,r+r'}| &= \|A^{1,r+r'}\| = 2 + p - r - r', \\
\left|A^2_{(2,0,p-r-1,r+1)}\right| &= \left\|A^2_{(2,0,p-r-1,r+1)}\right\| = p - r + 1, \\
\left|A^3_{(p-r+1,0,0,p+1)}\right| &= \left\|A^3_{(p+1,0,0,p+1)}\right\| = p - r + 1, \\
|||s\frac{f^{(r')}}{r'!} G_{A^{1,r+r'}}|||_N &= C_f^r K_0^{p-r-r'} K_1^2 N^{p+2+p-r-r'}/(p - r + -r')!, \\
|||f G_{A^2_{(2,0,p-r-1,r+1)}}|||_N &= C_f K_0^{p-1} K_1^2 N^{p+2+p-r-1}/(p - r - 1)!, \\
|||f G_{A^3_{(p-r+1,0,0,p+1)}}|||_N &= C_f K_1^{p+1} N^{p+1+p-r+1}/(p - r - 1)!.
\end{aligned}
$$

Again we will use (5.34) to verify that these triple-bar norms have not increased, and it is evident that the single and double-bar norms of A have behaved as before, with $|A|$ decreasing and $\|A\|$ not increasing.

We should verify the fate of terms such as the last in (3.32) above when subjected to the a priori estimate.

As long as $q \geq 2$, we may treat the $G_{A^3_{(p-r+1,0,0,p+1)}}$ term again, and generically so:

Proposition 5.9.

$$
\begin{aligned}
&[P, \varphi^{(p+1)} T^s Z^{p-r-1}]u \\
&\quad = [f Z^2, \varphi^{(p+1)} T^s Z^{p-r-1}] = f \varphi^{(p+3)} T^s Z^{p-r+1} u + 2 f \varphi^{(p+2)} T^s Z^{p-r} u \\
&\qquad + \underline{(p - r - 1)}\, f \varphi^{(p+1)} T^{s+1} Z^{p-r-1} u \\
&\qquad + \sum_{\substack{s' \leq s, r' \leq p-r-1 \\ s'+r' \geq 1}} \binom{s}{s'} \binom{p-r-1}{r'} f^{(s'+r')} \varphi^{(p+1)} T^{s-s'} Z^{p-r-r'-1} u \\
&\quad = f G_{A^1_{(p-r-1,s,0,p+3)}} + 2 f G_{A^2_{(p-r,s,0,p+2)}} + (p - r - 1) f G_{A^3_{(p-r-1,s,0,p+1)}} \\
&\qquad + \sum_{\substack{s' \leq s, r' \leq p-r-1 \\ s'+r' \geq 1}} \frac{s!(p - r - 1)!}{(s - s')!(p - r - r' - 1)!} \frac{f^{(s'+r')}}{r'!s'!} G_{A^4_{(p-r-r'-1,s-s',0,p+1)}}.
\end{aligned}
\tag{5.39}
$$

Thus

$$
\begin{aligned}
\left\|\varphi^{(p+1)} T^s Z^{p-r+1} u\right\|_0 &\leq \left\|P \varphi^{(p+1)} T^s Z^{p-r-1} u\right\|_0 + \left\|\varphi^{(p+1)} T^s Z^{p-r} u\right\|_0 \\
&\leq \left\|\varphi^{(p+1)} T^s Z^{p-r-1} P u\right\|_0 + \left\|[P, \varphi^{(p+1)} T^s Z^{p-r-1}]u\right\|_0 \\
&\quad + \left\|\varphi^{(p+1)} T^s Z^{p-r} u\right\|_0
\end{aligned}
$$

and

$$\begin{aligned}
&\left\|[P, \varphi^{(p+1)} T^s Z^{p-r-1}]u\right\|_0 \\
&\quad = \left\|[f Z^2, \varphi^{(p+1)} T^s Z^{p-r-1}]u\right\|_0 = \left\| f G_{A^1_{(p-r-1,s,0,p+3)}} u \right\|_0 \\
&\qquad + \left\| 2 f G_{A^2_{(p-r,s,0,p+2)}} u \right\|_0 + \left\| (p-r-1) f G_{A^3_{(p-r-1,s+1,0,p+1)}} u \right\|_0 \qquad (5.40)\\
&\qquad + \left\| \sum_{\substack{s' \le s, r' \le p-r-1 \\ s'+r' \ge 1}} \frac{s!(p-r-1)!}{(s-s')!(p-r-r'-1)!} \frac{f^{(s'+r')}}{r'!s'!} G_{A^4_{(p-r-r'+1,s-s',0,p+1)}} u \right\|_0 \\
&\quad \le C_f \left\|\varphi^{(p+3)} T^s Z^{p-r-1} u\right\|_0 + 2C_f \left\|\varphi^{(p+2)} T^s Z^{p-r} u\right\|_0 \\
&\qquad + \underline{(p-r-1)} C_f \left\|\varphi^{(p+1)} T^{s+1} Z^{p-r-1} u\right\|_0 \\
&\qquad + \sum_{\substack{s' \le s, r' \le p-r-1 \\ s'+r' \ge 1}} \frac{s!(p-r-1)!}{(s-s')!(p-r-r'-1)!} C_f^{s'+r'} \left\|\varphi^{(p+1)} T^{s-s'} Z^{p-r-r'+1} u\right\|_0 .
\end{aligned}$$

(5.41)

The norms behave as expected: with $A = (p-r+1, s, 0, p+1)$, $|A| = p - r + 1 + s$, $\|A\| = p - r + 1 + 2s$, the $|.|$ norms decrease, while the $\|.\|$ norms do not increase:

$$\begin{aligned}
&|A^1| = p - r - 1 + s < |A|, \quad \left\|A^1\right\| = p - r - 1 + 2s < \|A\|, \\
&|A^2| = p - r + s < |A|, \quad \left\|A^2\right\| = p - r + 2s < \|A\|, \\
&|A^3| = p - r - 1 + s + 1 < |A|, \quad \left\|A^3\right\| = p - r - 1 + 2(s+1) = \|A\|,
\end{aligned}$$

and

$$|A^4| = p - r - r' + 1 + s - s' < |A|, \quad \left\|A^4\right\| = p - r - r' + 1 + 2(s - s') = \|A\|,$$

under the condition that $r' + s' > 0$, while dividing by $(p-r)!$, which is how these occur, and noting that we started with $|||G|||_N = K_0^p K_1^2 N^{2p+2}/p!$, we have

$$|||f G_{A^1_{(p-r-1,s,0,p+3)}}|||_N/(p-r)! = C_f K_1^{p-r-1} K_2^s N^{p-r-1+s+p+3}/(p-r)!,$$

which is bounded by $K_0^p K_1^2 K_2^s N^{2p+2+s}/p!$, since $N^{p-r}/(p-r)! \le N^p/p!$, for the usual choice of the K_j;

$$|||2 f G_{A^2_{(p-r,s,0,p+2)}}|||_N/(p-r)! = C_f K_1^{p-r} K_2^s N^{p-r+s+p+3}/(p-r)!$$

which is bounded by $K_0^p K_1^2 K_2^s N^{2p+2+s}/p!$ as well, as just above,

$$\left|\!\left|\!\left|(p-r-1)fG_{A^3_{(p-r-1,s+1,0,p+1)}}\right|\!\right|\!\right|_N \Big/ (p-r)!$$
$$= C_f(p-r-1)K_1^{p-r-1}K_2^s N^{p-r-1+s+1+p+1}/(p-r)!,$$

which is also bounded by $K_0^p K_1^2 K_2^s N^{2p+2+s}/p!$ under the conventions on the K_j; and finally,

$$\sup_{\substack{s'\le s, r'\le p-r-1\\ s'+r'\ge 1}} \left|\!\left|\!\left|\frac{s!(p-r-1)!}{(s-s')!(p-r-r'-1)!}C_f^{s'+r'}\varphi^{(p+1)}T^{s-s'}Z^{p-r-r'+1}\right|\!\right|\!\right|_N \Big/ (p-r)!$$
$$= \sup \frac{s!(p-r-1)!}{(s-s')!(p-r-r'-1)!}C_f^{s'+r'}K_1^{p-r-r'+1}K_2^{s-s'}N^{2p-r-r'+2+s-s'}\Big/(p-r)!$$
$$\le K_0^p K_1^2 K_2^s N^{2p+2+s}/p!$$

for suitable K_j for the reasons discussed above.

All of this continues as long as we have two Z derivatives. For rigid coefficients, brackets of $P = fZ^2$ with $(T^m)_\varphi$ always retain at least two Z's, but in the case of nonrigid coefficients, can generate new Y's, as in Proposition 5.13 below, which we merely touch upon here in order to show how a single Z can also be handled, a situation that does not arise with rigid coefficients. For a subsequent bracket of Z^2 with a single Y could produce ZT, the feared case of a single Z.

For rigid coefficients, only when m becomes zero can all but one of the Z's be lost: the Z's from P can differentiate φ, leaving fewer Z's, or one of them can bracket with a Z from Z^q to give a T and a large coefficient proportional to q; in any case, all roads lead to norms of $F\varphi^{(\tilde r)}(Z)T^{\tilde m}$, i.e., with at most one Z, and at this point we build a new $(T^{\tilde m})_{\tilde\varphi}$ with a new localizing function $\tilde\varphi$, and then bring $F\varphi^{(\tilde r)}$ out of the L^2 norm.

But if we examine the restrictions imposed by $\|A\|$, together with the $|\!|\!|\cdot|\!|\!|$ norm behavior (nonincreasing), we find that in all cases, $2\tilde m + 1 \le p+1, \tilde r \le 2p+1$, and

$$|\!|\!|F\varphi^{(\tilde r)}(Z)T^{\tilde m}|\!|\!|_N = \sup|F|K_0^0K_1K_2^{\tilde m}N^{1+\tilde m+\tilde r}$$
$$\le K_0^pK_1^1K_2^0N^{1+2p}/p! = |\!|\!|(Z)(T^p)_\varphi|\!|\!|_N.$$

This means that $\sup|F| \le K_0^pK_2^{-\tilde m}N^{2p-\tilde m-\tilde r}/p!$ and hence that

$$\left\|Z^2T^pu\right\|_{L^2(\varphi\equiv 1)} \le \sup_{\substack{2\tilde m\le p\\ \tilde r\le 2p+1}} \left\|F\varphi^{(\tilde r)}(Z)T^{\tilde m}u\right\|_0$$
$$\le \sup_{\substack{2\tilde m\le p\\ \tilde r\le 2p+1}} K_0^pK_2^{-\tilde m}\frac{|\varphi^{(\tilde r)}|}{N^{\tilde r}}N^{2p-\tilde m}u\left\|(Z)T^{\tilde m}u\right\|_{L^2(\operatorname{supp}\varphi)}\Big/p!$$

We have proved, since $N^p/p! \le C^p$ for $p \sim N$, the following:

Proposition 5.10. *For rigid coefficients in P, starting with pure T derivatives in the form $G = T^{p+1}u$ in an open set ω leads immediately to two terms $G = T^p Z^2 u$ in ω and then to $G = (T^p)_\varphi Z^2 u$ with $\varphi \equiv 1$ near ω, and then, after at most p iterations of the a priori estimate to terms without $(T_\varphi^{\tilde{p}})u$ and without increasing $|||G|||_N$ but with p dropping to zero and $\|A\|$ at most p but $|A| \le p/2$ and $q < 2$ and thus*

$$\frac{\left\|Z^{\le 2}T^p u\right\|_{L^2(\omega)}}{N^p} \le \sup_{\substack{2\tilde{m}\le p\\ \tilde{r}\le 2p+1}} C^p \frac{|\varphi^{(\tilde{r})}|}{N^{\tilde{r}}} \frac{\left\|(Z)T^{\tilde{m}}u\right\|_{L^2(\operatorname{supp}\varphi)}}{N^{\tilde{m}}}$$

This last line is equivalent to (5.18) above, and the argument concludes exactly as it does there.

5.6 Nonrigid Coefficients

When the coefficients depend on t, the analysis becomes more complex (and tedious). We have

$$\begin{aligned}
[(T^p)_\varphi, a]ZZu &= \left[\sum_{|\alpha+\beta|\le p} \frac{((-X)^\alpha Y^\beta \varphi)\circ X^\beta Y^\alpha T^{p-|\alpha+\beta|}}{\alpha!\beta!}, a\right] ZZu \\
&= \sum_{\substack{|\alpha+\beta|\le p\\ \alpha_1\le\alpha,\beta_1\le\beta\\ 1\le j+|\alpha_1+\beta_1|}} \binom{\alpha}{\alpha_1}\binom{\beta}{\beta_1}\binom{p-\alpha-\beta}{j}(Y^{\alpha_1}X^{\beta_1}T^j a) \\
&\qquad \circ \frac{((-X)^\alpha Y^\beta \varphi)\circ X^{\beta-\beta_1}Y^{\alpha-\alpha_1}T^{p-|\alpha+\beta|-j}}{\alpha!\beta!} ZZu \\
&= \sum_{\substack{|\alpha+\beta|\le p\\ \alpha_1\le\alpha,\beta_1\le\beta\\ 1\le j+|\alpha_1+\beta_1|}} \binom{p-\alpha-\beta}{j}\frac{(Y^{\alpha_1}X^{\beta_1}T^j a)}{\alpha_1!\beta_1!} \\
&\qquad \times \frac{((-X)^\alpha Y^\beta \varphi)\circ X^{\beta-\beta_1}Y^{\alpha-\alpha_1}T^{p-|\alpha+\beta|-j}}{(\alpha-\alpha_1)!(\beta-\beta_1)!} ZZu.
\end{aligned}$$

Naturally the binomial coefficient contains a $j!$ in the denominator, which is appropriate to offset and control the T derivatives of the coefficient a, which were, however, missing in the rigid case. But the rest of the binomial coefficient *also* contains j, and in the wrong form for the definition of $(T^p)_\varphi$, whose balance is quite precise and essential.

But it turns out that there is a way to introduce parameters and rewrite the binomial coefficient to yield a sum of balanced operators $(T^{\tilde{p}})_{\tilde{\varphi}}$ involving these parameters in a harmless (if complicated) way.

With $A = |\alpha_1 + \beta_1|$ and $B = |(\alpha - \alpha_1) + (\beta - \beta_1)|$, we have [Fel]

$$\binom{p-(A+B)}{j} = \sum_{\substack{t \le j \\ t \le |B|}} \binom{p-A-t}{j-t} (-1)^t \binom{B}{t} \tag{5.42}$$

and then

$$\binom{B}{t} = \sum_{\substack{|\alpha_2+\beta_2|=t \\ \alpha_2 \le \alpha-\alpha_1 \\ \beta_2 \le \beta_1}} \frac{(\alpha-\alpha_1)!(\beta-\beta_1)!}{t!((\alpha-\alpha_1)-\alpha_2)!((\beta-\beta_1)-\beta_2)!}. \tag{5.43}$$

Before we use the property (5.2) of the new vector fields to "commute" the "extra" Y^{β_1} applied to $T^{\beta_1}\varphi$, may now be effectively moved to φ as in (5.2): letting

$$\tilde{\alpha} = \alpha - \alpha_1 - \alpha_2, \qquad \tilde{\beta} = \beta - \beta_1 - \beta_2,$$

as we must, since the new denominators become the internal indices, we obtain

$$\begin{aligned}
&[(T^p)_\varphi, a]ZZu \\
&= \sum_{\substack{|\alpha+\beta| \le p \\ \alpha_1 \le \alpha, \beta_1 \le \beta \\ 1 \le j + |\alpha_1+\beta_1|}} \sum_{\substack{t \le j \\ t \le |B|}} \binom{p-|\alpha_1+\beta_1|-t}{j-t} (-1)^t \frac{(Y^{\alpha_1}X^{\beta_1}T^j a)}{\alpha_1!\beta_1!t!} \\
&\quad \circ \sum_{\substack{|\alpha_2+\beta_2|=t \\ \alpha_2 \le \alpha-\alpha_1 \\ \beta_2 \le \beta_1}} \frac{((-X)^{\alpha-\tilde{\alpha}}(-X)^{\tilde{\alpha}}Y^{\tilde{\beta}}(Y^{\beta-\tilde{\beta}}\varphi)) \circ X^{\beta_2}X^{\tilde{\beta}}Y^{\tilde{\alpha}}Y^{\alpha_2}T^{p-|\alpha+\beta|-j}}{\tilde{\alpha}!\tilde{\beta}!} ZZu \\
&= \sum_{\substack{\alpha+\beta \le p \\ \alpha_1 \le \alpha, \beta_1 \le \beta \\ 1 \le j + |\alpha_1+\beta_1|}} \sum_{\substack{t \le j \\ t \le |B|}} \binom{p-|\alpha_1+\beta_1|-t}{j-t} \frac{j!}{t!} (-1)^t \frac{(Y^{\alpha_1}X^{\beta_1}T^j a)}{\alpha_1!\beta_1!j!} \\
&\quad \circ \sum_{\substack{|\alpha_2+\beta_2|=t \\ \alpha_2 \le \alpha-\alpha_1 \\ \beta_2 \le \beta_1}} \frac{((-X)^{\alpha_{12}}(-X)^{\tilde{\alpha}}Y^{\tilde{\beta}}(Y^{\beta_{12}}\varphi)) \circ X^{\beta_2}X^{\tilde{\beta}}Y^{\tilde{\alpha}}Y^{\alpha_2}T^{p-(j+|\alpha_{12}+\beta_{12}|)-|\tilde{\alpha}+\tilde{\beta}|}}{\tilde{\alpha}!\tilde{\beta}!} ZZu
\end{aligned}$$

with $\alpha_{12} = \alpha_1 + \alpha_2$ and $\beta_{12} = \beta_1 + \beta_2$.

Proposition 5.11. *The binomial coefficient will be estimated by*

$$\binom{p-|\alpha_1+\beta_1|-t}{j-t} \frac{j!}{t!} (-1)^t = C_{p,\alpha_1,\beta_1,j} \lesssim C^j p^{j-t}, \tag{5.44}$$

completing the expression of the bracket:

Proposition 5.12.

$$[(T^p)_\varphi, a]ZZu = \sum_{\substack{\alpha+\beta\le p\\ \alpha_1\le\alpha,\beta_1\le\beta\\ 1\le p-j-|\alpha+\beta|}} \sum_{\substack{|\alpha_2+\beta_2|=t\le j\\ \alpha_2\le\alpha-\alpha_1\\ \beta_2\le\beta-\beta_1}} C_{p,\alpha_1,\beta_1,j}\frac{(Y^{\alpha_1}X^{\beta_1}T^j a)}{\alpha_1!\beta_1!j!}$$
$$\circ\frac{((-X)^{\alpha_{12}}X^{\tilde\alpha}Y^{\tilde\beta}Y^{\beta_{12}}\varphi)\circ X^{\beta_2}X^{\tilde\beta}Y^{\tilde\alpha}Y^{\alpha_2}T^{p-j-|\alpha_{12}+\beta_{12}|-|\tilde\alpha+\tilde\beta|}}{\tilde\alpha!\tilde\beta!}ZZu. \tag{5.45}$$

Except for the *location* of $(-X)^{\alpha_{12}}$ and X^{β_2}, we may understand these indices as follows:

- We will think of the last line as $(T^{p-(j+|\alpha_1+\beta_1+\alpha_2+\beta_2|)})_{(-X)^{\alpha_{12}}Y^{\beta_{12}}\varphi}X^{\beta_2}Y^{\alpha_2}$, noting that the $(-X)^{\alpha_{12}}$ actually occur to the left of the derivatives on φ. They will be commuted to the right, onto $Y^{\beta_{12}}\varphi$. It is the special form of the vector fields that will make this an effective move, since the vector fields that balance free vector fields will not be disturbed and there will appear harmless powers of y to the extreme left.
- The drop of j in the exponent of T is balanced by p^j in (5.44).
- φ has received $\gamma = \alpha_{12}+\beta_{12} = \alpha_1+\beta_1+\alpha_2+\beta_2$ additional derivatives, balancing the drop in the exponent of $(T^p)_\varphi$.
- There are $\alpha_2+\beta_2$ "new" Z's, balanced by $p^{-t} = p^{-(|\alpha_2+\beta_2|)}$.
- The "new" X^{β_2} to the right of φ, when commuted to the left so that they are composed with the whole $(T^{p-(j+|\alpha_1+\beta_1+\alpha_2+\beta_2|)})_{(-X)^{\alpha_{12}}Y^{\beta_{12}}\varphi}$ on the left, may, in the commuting, land on the derivatives on φ, and because of the very special form of the vector fields, may have their derivatives slide onto φ while leaving harmless powers of y as coefficients to the left. Thus instead of X^{β_2} to the left there may be fewer, the others landing as derivatives on φ. But keep in mind that with each X or new derivative on φ comes an inverse power of p.

With these observations in mind, we may write the following result.

Proposition 5.13.

$$\begin{aligned}&[(T^p)_\varphi, a]ZZu\\ &\quad=\sum C_{p,\alpha_1,\beta_1,j}\ (y)^{\beta_2''+\alpha_{12}}\circ\frac{(Y^{\alpha_1}X^{\beta_1}T^j a)}{\alpha_1!\beta_1!j!}\\ &\qquad\circ X^{\beta_2'}\frac{((-X)^{\tilde\alpha}Y^{\tilde\beta}\varphi^{(\beta_2''+\beta_{12}+\alpha_{12})})\circ X^{\tilde\beta}Y^{\tilde\alpha}T^{(p-j-|\alpha_{12}+\beta_{12}|)-|\tilde\alpha+\tilde\beta|}}{\tilde\alpha!\tilde\beta!}Y^{\alpha_2}ZZu.\\ &\quad=\sum C_{p,\alpha_1,\beta_1,j}\,(y)^{\beta_2''+\alpha_{12}}\frac{(Y^{\alpha_1}X^{\beta_1}T^j a)}{\alpha_1!\beta_1!j!}X^{\beta_2'}\left(T^{p-j-|\alpha_{12}+\beta_{12}|}\right)_{\varphi^{(\beta_2''+\beta_{12}+\alpha_{12})}}Y^{\alpha_2}Z^2u,\end{aligned} \tag{5.46}$$

and $C_{p,\alpha_1,\beta_1,j} \lesssim C^j p^{j-t}$, *and where the summation is over all indices with* $t = \beta_{12} + \alpha_{12}, \beta_{12} = \beta_1 + \beta_2, \alpha_{12} = \alpha_1 + \alpha_2, \beta_2 = \beta_2' + \beta_2''$ *and* $j + |\alpha_{12} + \beta_{12}| \le p$. *Here* (y) *denotes either* y *or* 1.

Note that the powers of y will just come out of the norm, and we may assume that $|y| \le 1$.

Again, we introduce notation for variable-coefficient sums of the general expressions that will arise. With

$$G_{A,\psi} = C_A (T^m)_{\psi^{(r)}} T^s Z^q$$

and $A = A_{(q,s,m,r)}$, we assign single- and double-bar norms that reflect the distinction between Z's and T's, assigning free T's double weight: let

$$|A_{(q,s,m,r)}| \equiv |G_{A_{(q,s,m,r)},\psi}| = m + q + s$$

and

$$\left\| A_{(q,s,m,r)} \right\| \equiv \left\| G_{A_{(q,s,m,r)},\psi} \right\| = |A_{(q,s,m,r)}| + q = m + 2q + s.$$

To a variable coefficient sum of such expressions,

$$\mathcal{G} = \sum F_A G_{A;\psi} \text{ with } |A| \le N, \tag{5.47}$$

we set, again,

$$|||\mathcal{G}|||_N = \sup_{A,\, \mathrm{supp}\,(\psi)} |F_A|\, C_A K_0^m K_1^q K_2^q N^{|A|+r+m}/m!, \tag{5.48}$$

where K_0, K_1, K_2 will be chosen later, large relative to other constants that enter and subject to

$$K_0 \gg K_1 \gg K_2 \gg 1. \tag{5.49}$$

We shall use the notation $F_{(k)}^{(j)}$ for a function or sum of functions such that in the support of all localizing functions, and for all σ,

$$|D^\sigma F_{(k)}^{(j)}| \le C^{|\sigma|+j+k+1}(|\sigma| + k)!, \tag{5.50}$$

with the constant C universal for the present problem. For example, $F_{(k)}^{(j)}$ could stand for any $(j + k)$th derivative of any coefficient divided by $j!$

Definition 5.1. An operator $G_{A;\psi}$ will be called *admissible* if it contains two Z's. $G_{A;\psi}$ will be called *simple* if it is not admissible and $m = 0$, i.e., if it of the form

$$\psi^{(r)} T^s (Z) = \text{ coeff } \psi^{(r)} T^s Z \quad \text{or} \quad \psi^{(r)} T^s.$$

A variable-coefficient sum of such operators $\mathcal{G}$ will be said to have one of these properties if it is true of each member of the sum.

We have actually proved the following propositions, which show the effect of commuting $(T^m)_\psi$, or more generally $G_{A;\psi}$, with a function or a vector field.

Proposition 5.14 (Bracket with a function). *Let $G^1_{A;\psi} = (T^m)_{\psi^{(r)}} T^s Z^q$, and let $f(x,y,t)$ be a real analytic function in the support of ψ (which will be smooth and of compact support). Then, noting that ρ may equal zero, so that this expression contains the information about the commutator,*

$$G^1_{A;\psi}\, f Z^2 = \sum_{|\rho|\geq 0} f_{(\rho)} G^1_{A_{[-\rho]};\psi} Z^2 = \tilde{\mathcal{G}}^1_{A_{[0]};\psi} Z^2$$

with

$$f_{(0)} = f, \quad A_{[0]} = A, \quad |D^\sigma f_{(\rho)}| \leq C^{|\sigma+\rho|+1}|\sigma|!, \quad \forall\sigma,$$

and

$$|A_{[-\rho]}| \leq |A| - |\rho|, \quad \left\|A_{[-\rho]}\right\| \leq \|A\| - \rho,$$

and

$$|||\tilde{\mathcal{G}}^1_{A_{[0]};\psi} Z^2|||_N \leq C_f\, |||G^1_{A;\psi} Z^2|||_N .$$

We are using the notation $A_{[-\rho]}$ to denote a tuple whose $|\cdot|$ norm is $|A| - \rho$. And note that there may be more Z's in $G^1_{A_{(-\rho)}}$ than there were in the initial $G^1_{A;\psi}$ due to Proposition 5.13.

In the bracket with Z's, there is no difference from the previous section, since the vector fields are the same. Thus we have the following:

Proposition 5.15 (Bracket with Z's). *Let $A = (q+2, s, m, r)$. Then*

$$a[(T^m)_{\psi^{(r)}} T^s Z^q, Z^2] = \sum_{|\rho|\geq 1} a f_{(\rho)} G^2_{A_{[-\rho]};\psi} = \tilde{\mathcal{G}}^2_{A_{[-1]};\psi}$$

with

$$|D^\sigma f_{(\rho)}| \leq C^{|\rho+\sigma|}|\sigma|!\ \forall\alpha, \quad |A_{[-\rho]}| \leq |A| - |\rho|, \quad \left\|A_{[-\rho]}\right\| \leq \|A\|,$$

and

$$|||\tilde{\mathcal{G}}^2_{A_{[-1]};\psi}|||_N \leq |||G^2_{A;\psi}|||_N .$$

Thus the single-bar norm drops, while the double-bar norm does not rise and the triple-bar norms are well controlled.

Note that $\tilde{\mathcal{G}}^2_{A_{[-1]};\psi}$ will always contain at least one Z: if $q > 1$, then the result contains two Z's, and is thus admissible; if $q = 1$, this Z may take up one of the two Z's to the right to produce a T, leaving one Z, and if $q = 0$, then bracketing with Z^2 leaves Z^2, and hence is admissible. We never use up all Z's in this way.

Corollary 5.1. *Let* $G^3_{A;\psi} = (T^m)_{\psi^{(r)}} T^s Z^q$. *Then*

$$\left\| P G^3_{A;\psi} u \right\|_0 \leq \left\| G^3_{A;\psi} P u \right\|_0 + \sup_{|\rho|\geq 0} C_f^{|\rho|+1} \rho! \left\| G^3_{A_{[1-\rho]},\psi} u \right\|_0$$
$$= \left\| G^3_{A;\psi} P u \right\|_0 + \sup_{|\rho|\geq 0} C_f^{|\rho|+1} \rho! \left\| G^3_{A_{[1-\rho]},\psi} u \right\|_0$$

and

$$|||C_f^{|\rho|+1} \rho! G^3_{A_{[1-\rho]},\psi}|||_N \leq |||Z^2 G^3_{A;\psi}|||_N, \quad |A_{[1-\rho]}| < |A| + 2,$$

and

$$\left\| A_{[1-\rho]} \right\| \leq \|A\| + 2.$$

That is, starting with Z^2 and $G^3_{A;\psi} = (T^m)_{\psi^{(r)}} T^s Z^q$ together, and applying the estimate, the right-hand side will be bounded as shown and the formal $|||.|||_N$ norm is controlled from where we started. The single-bar norm of A has dropped, and the double-bar norm of A has not risen.

However, a new feature enters with nonrigid coefficients: the bracket with a coefficient, as in Proposition 5.13, may contain new Z's (i.e., new Y's), which, while seemingly a good thing, may lead to inadmissible terms on the next iteration, without having reduced m to zero, as the schema

$$(T^m)_\psi T^s Z^2 \to [(T^m)_{\psi^{(r)}}, f] T^s Z^2 \to f'(T^{m-1})_{\psi^{(r+1)}} T^s Y Z^2/m$$
$$\to f'(T^{m-1})_{\psi^{(r+1)}} T^{s+1} Z/m \tag{5.51}$$

shows. In the rigid case this did not occur: m would have become zero, the two Z's from P never bracket one another, and the phenomenon of new Y from a bracket with a coefficient does not occur. Thus there will always remain two Z's, and when ($m = 0$ and) at most two Z's we introduced a new localizing function.

In the nonrigid case, we could invoke the inner-product form of the a priori estimate where one Z suffices:

$$\|Zv\|_0^2 + \|v\|_{1/2}^2 \lesssim |(Pv, v)_{L^2}| + \|v\|_0^2 \,|\cdot| \tag{5.52}$$

But since in the next chapter we will deal with coefficients that are not only nonrigid but even pseudodifferential, we feel encouraged here to use an extremely simple pseudodifferential operator and take advantage of the full power of the norm estimate (3.11):

$$\left\| Z^2 v \right\|_0^2 + \|Zv\|_{1/2}^2 + \|v\|_1^2 \lesssim \|Pv\|_0^2 + \|v\|_0^2, \qquad v \in C_0^\infty. \tag{5.53}$$

Then starting as just above but taking $q = 0$ for simplicity, from $(T^m)_\psi Z^2 = G_{A_1}$, which has $A_1 = (2, 0, m, 0)$ and hence $|A_1| = 2 + m$, $\|A_1\| = 2 + m$, and $|||G_{A_1}|||_N = K_0^m K_1^2 K_2^0 N^{2+r+2m}/m!$, we arrive (5.51) at $G_{A_2} = (T^{m-1})_{\psi(r+1)} TZ/m$, for which

$$|A_2| = 1 + 1 + m - 1 = m + 1,$$
$$\|A_2\| = 1 + 2 + m - 1 = m + 2,$$

and

$$|||G_{A_2}|||_N = K_0^{m-1} K_1 K_2 N^{1+1+r+1+2(m-1)}/m(m-1)! = K_0^{m-1} K_1 K_2 N^{2m+r+1}/m!,$$

all norms have dropped or remained constant.

In order to apply (3.11) optimally, we use the operator defined by $\widehat{\Lambda_t^{1/2} w} = (1 + |\tau|^2)^{1/4}\, \widehat{w}(\xi, \tau)$:

$$\begin{aligned}\left\| ZT(T^{m-1})_{\psi(r+1)} u \right\|_0 &= \left\| Z\Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} u \right\|_{1/2} \\ &\quad + \left\| Z^2\Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} u \right\|_0 + \left\| \Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} u \right\|_1 \\ &\lesssim \left\| P\Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} u \right\|_0 + \left\| \Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} u \right\|_0 \\ &\le \left\| \Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} Pu \right\|_0 + \|Eu\|_0 + \left\| \Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} u \right\|_0,\end{aligned}$$

where $E = [\Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)}, aZ^2]u$ is much as before with the added $\Lambda_t^{1/2}$ (since the vector fields Z are rigid, i.e., $[Z, \Lambda_t] = 0$) except for the term in which the bracket $[\Lambda_t^{1/2}, a]Z^2$ enters.

However, $[\Lambda_t^{1/2}, a]$ is a bounded operator on L^2, and so

$$\begin{aligned}\|Eu\|_0 &\sim \left\| [\Lambda_t^{1/2}, a](T^{m-1})_{\psi(r+1)} Z^2 u \right\|_0 \lesssim \left\| (T^{m-1})_{\psi(r+1)} Z^2 u \right\|_0 \\ &\le \varepsilon \left\| \Lambda_t^{1/2}(T^{m-1})_{\psi(r+1)} Z^2 u \right\|_0\end{aligned}$$

when supports are small; hence this term will be absorbed on the left. The others satisfy the norm requirements.

We conclude that terms with one Z are not a problem either.

And finally, we never arrive at terms with no Z's unless m drops to zero. Terms with $m = 0$ are called *simple*.

Upon iteration, we have proved, the following result.

Proposition 5.16 (Applying the estimate). *Let $H_{A;\Psi}$ be admissible and $Pu = f$ in V. Then*

$$\|H_{A;\Psi} u\|_0^2 \le C^N \left(\left\| G_{A_{[-2]};\Psi} f \right\|_0^2 + \sup \left\| G^{sim}_{A';\Psi} u \right\|_0^2 \right),$$

where the supremum is over all simple $G^{sim}_{A';\Psi}$ *whose norms satisfy*

$$|||G^{sim}_{A';\Psi}|||_N \le |||H_{A;\Psi}|||_N\,, \quad |A'| \le |A| - 1, \quad \textit{and} \; \left\|A'\right\| \le \|A\|\,.$$

The simple terms have also been treated, with the same control over norms, since for them when $q \le 1$ we just stop and introduce a new localizing function in view of (5.18) above and the sentences following it. This concludes the proof of analyticity with the same choice of localizing functions.

Chapter 6
Pseudodifferential Problems

6.1 Generalization to Pseudodifferential Operators

In 1980, Guy Métivier generalized the above results, still with the same essential geometric conditions, i.e., in which the operator degenerated to exactly order two on the characteristic manifold, for instance, but in higher codimension included the case of pseudodifferential operators, or, what amounted to the same thing, operators such as we have studied above but with pseudodifferential *coefficients*. We discuss this generalization now in our formulation, which the reader will realize is a direct generalization of the first few chapters of the present book.

The operator to be studied is

$$P(x, D) = \sum_{|I|\leq 2} C_I(x, D) Z^I, \tag{6.1}$$

where the Z are the usual X or Y and the $C(x, D)$ are "classical" pseudodifferential operators of order zero. The replacement for the "sum of squares" hypothesis is that

$$\sum_{|I|=2} C_I(x_0, \xi_0)\zeta^I \neq 0, \quad \zeta \in \mathbb{R}^{2\nu}, \tag{6.2}$$

and subellipticity becomes that the operator $\sum_{|I|=2} C_I(x_0, \xi_0) B^I$ has no kernel in $\mathcal{S}$, where

$$B_j = \frac{\partial}{\partial x_j} - x_{\nu+j}, \quad B_{\nu+j} = \frac{\partial}{\partial x_{\nu+j}},$$

the vector fields Z_j after Fourier transform in the t variable at the covariable $\xi = (0, \ldots, 0, 1)$.

Rather than develop the whole theory of pseudodifferential operators from scratch, we refer the reader to standard sources and merely cite the results we will need, none of which are at all deep.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_6,

If we freeze the coefficients, and write

$$P_{(x_0,\xi_0)}(x,D)=\sum_{|I|\le 2} C_I(x_0,\xi_0)X^I,$$

the hypothesis (6.2) implies the following estimate of Grušin type: $\forall v\in C_0^\infty$,

$$\|Z^2v\|_0\lesssim \|P_{(x_0,\xi_0)}(x,D)v\|_0+\|v\|_0,$$

and then, first bounding $\|Zv\|_0$ via the Schwarz inequality, and then $\|v\|_1^2\sim \|Zv\|_0+\|Tv\|_0\sim \|Zv\|_0+\|Z^2v\|_0$, and finally $\|Zv\|_{1/2}^2\lesssim \|Z^2v\|_0^2+\|v\|_1^2$, we obtain

$$\|Z^2v\|_0+\|Zv\|_{1/2}+\|v\|_1\lesssim \|P_{(x_0,\xi_0)}(x,D)v\|_0+\|v\|_0. \tag{6.3}$$

We may pass from here to the analogous estimate for $P_{(x,\xi_0)}(x,D)$ for v of small support, since modulo errors with zero symbol (hence of arbitrarily low order, which do not disturb analyticity) they are small multiples of the left-hand side, although we would have had to introduce fractional powers of the Laplacian in t.

But the support in ξ is not so easily controlled and requires a nested system of cones with suitable families of cut-off functions, just as the spatial (local) proof above required a nested system of open sets with suitable families of cut-off functions.

In order to localize also in the covariables ξ, we proceed analogously: for any N, and open sets $\omega\Subset\tilde\omega$ and $\Gamma\subset\tilde\Gamma$ with $\Gamma,\tilde\Gamma$ both open cones, there exist $\varphi_N(x)$ as before, with $\varphi_N(x)\in C_0^\infty(\tilde\omega)\equiv 1$ on ω and

$$|D^\alpha\varphi_N|\le C(CN)^{|\alpha|},\quad |\alpha|\le 2N,$$

and $\psi_N(\xi)\equiv 1$ on Γ and $\psi_N(\xi)\equiv 0$ in the complement of $\tilde\Gamma$, with similar growth of derivatives, but conically,

$$|D^\alpha\psi_N|\le C(CN/|\xi|)^{|\alpha|},\quad |\alpha|\le 2N,$$

and $(1-\lambda_N(\xi))\in C_0^\infty(\{|\xi|<1/2\})$ with

$$|D^\alpha\lambda_N(\xi)|\le C(CN)^{|\alpha|},\quad |\alpha|\le 2N.$$

Then we set

$$\rho_N(x,\xi)=\varphi_N(x)\psi_N(\xi)\lambda_N(\xi),$$

and use the symbols ψ and λ to denote both the "symbols" and also the operators $\psi(D_\xi),\lambda_N(D_\xi)$.

As above, it will be necessary to nest many of these open sets. Given an overall N, we will seek to bound derivatives of our solution u of order N, in L^2 norm, by $C(CN)^N$ in the form

$$\|D^\alpha\rho_N u\|\lesssim (CN)^{|\alpha|},\quad |\alpha|\le 2N.$$

The nesting will be done so that with a sequence of $N_i = N/2^i$, and open sets $\{\omega_j\}_{j=1,\ldots,\log_2 N}$, with $p_0 \in \omega = \omega_1 \Subset \cdots \Subset \omega_{\log_2 N} = \tilde{\omega}$ and the separation between the support of ω_j and $(\omega_{j+1})^c = d_j = d_0/2^j$ and a similar nesting of the Γ_j so that for $|\xi| = 1$, the separation of the supports is proportional to $1/2^j$ as well, and the same for the nested sets, which serve to exclude the origin in ξ-space.

Note that λ_N is of compact support, not conic support, but this localization away from the origin will matter the least, since if a derivative ever lands on λ_N, the resulting function will be supported in a region where $|\xi|$ is bounded (by 1), and the solution will certainly be analytic there. (That is, $\lambda_N'(\xi)\hat{u}(\xi)$ will have support in $|\xi| \leq 1$, and hence $\lambda_N'(D)u$ will be real analytic.)

6.2 The Microlocal $(T^p)_{\varphi\psi\lambda}$

Definition 6.1. We write $(x_0, \xi_0) \notin WF_A(v),\ v \in \mathcal{D}'$, if there exist an open cone Γ_0 in $\mathbb{R}^n \setminus 0$, a neighborhood V_0 of x_0, and a constant C such that for every N, there exists $v_N = v$ in $V_0,\ v_n \in \mathcal{E}'$, with

$$|\widehat{v_N}(\xi)| \leq C^N \left(1 + \frac{|\xi|}{N}\right)^{-N}, \quad \xi \in \Gamma_0.$$

Definition 6.2. We define

$$(T^p)_{\varphi\psi\lambda}(x, D) = \sum_{|\alpha+\beta|\leq p} \frac{\mathrm{ad}^{\alpha}_{(-X)}\, \mathrm{ad}^{\beta}_{Y}(\varphi\psi\lambda)}{\alpha!\beta!}(x, D)(X)^{\beta} Y^{\alpha} T^{p-|\alpha+\beta|}\, w,$$

where

$$\mathrm{ad}^{\alpha}_{X} = \mathrm{ad}_{X_1^{\alpha_1}} \cdots \mathrm{ad}_{X_n^{\alpha_n}}$$

and

$$\mathrm{ad}^{k}_{X}(w) = \underbrace{[X, [X, \ldots, [X, w] \ldots]]}_{k}.$$

Then we have the usual bracket relationships

$$[X, (T^p)_{\varphi\psi\lambda}] \equiv 0$$

and

$$[Y, (T^p)_{\varphi\psi\lambda}] \equiv (T^{p-1})_{\mathrm{ad}_T(\varphi\psi\lambda)} \circ Y$$

modulo C^p terms of the form

$$\mathrm{ad}^{p+1}_{Z}(\varphi\psi\lambda)/p! \circ Z^p.$$

6.3 Brackets with Coefficients

Brackets of $(T^p)_{\varphi\psi\lambda}$ with the vector fields X and Y behave beautifully. That is how the (micro-)localization of T^p was built. But brackets with coefficients were already complicated when the coefficients were functions; now they are pseudodifferential operators (of order zero).

We will disregard the tail of asymptotic series. The subject of analytic pseudo differential operators is well studied, and the new part of our proof is made entirely explicit in the families of cut-off functions in the dual variables.

For the bracket of a coefficient with $(T^p)_{\varphi\psi\lambda}$, we recall the expression we obtained when the coefficients were just functions and recall where each part of the expression came from. In Proposition 5.13 we had

$$[(T^p)_\varphi, a] = \sum C_{p,\alpha_1,\beta_1,j}(y)^{\beta_2''+\alpha_{12}} \frac{(Y^{\alpha_1} X^{\beta_1} T^j a)}{\alpha_1!\beta_1!j!}$$
$$\circ X^{\beta_2'} \left(T^{p-j-|\alpha_{12}+\beta_{12}|}\right)_{\varphi^{(\beta_2''+\beta_{12}+\alpha_{12})}} Y^{\alpha_2}$$
$$= \sum_{\substack{t+j+|\gamma|\le p\\ t=t_1+t_2'+t_2''}} F_{(0)}^{(\alpha_1+\beta_1+j)}(y)^{t_2''+\gamma} \frac{X^{t_{2'}}}{p^{t_{2'}}} \frac{p^j}{p^{t_2''}} (T^{p-j-|\gamma|})_{\varphi^{(|\gamma|+t_2'')}} \frac{Y^{t_1}}{p^{t_1}} ZZu,$$

$C_{p,\alpha_1,\beta_1,j} \lesssim C^j p^{j-t}$,

recalling the convention

$$|D^\sigma F_{(k)}^{(j)}| \le C^{|\sigma|+j+k+1}(|\sigma|+k)!.$$

When a is a zero-order pseudodifferential operator $C_I(x, D)$, $|I| = 2$:

- The derivatives on a are from

$$\mathrm{ad}_{(-X)}^{\alpha_1}\, \mathrm{ad}_Y^{\beta_1}\, \mathrm{ad}_T^j(a)$$

  and will remain in this form,

$$\mathrm{ad}_{(-X)}^{\alpha_1}\, \mathrm{ad}_Y^{\beta_1}\, \mathrm{ad}_T^j(C_I),$$

  together with the factorials. These will occur in the coefficients $F_{(0)}^{(\alpha_1+\beta_1+j)}$ but will be zero-order pseudodifferential operators, simple to calculate but bounded in L^2.
- Derivatives on the localizing function come from $\mathrm{ad}_W(\varphi)$ for some vector field $W = X, Y, T$ or $\partial/\partial x$, and the notation may still be used.
- Additionally, the special form of our vector fields, which permitted us to pass the function y from one side of (derivatives on) φ to the other, will behave a bit differently here. While $\varphi y = y\varphi$ for any function φ, i.e., $\mathrm{ad}_y(\varphi) = [y, \varphi] = 0$, it will no longer be true that $[\mathrm{ad}_y, \varphi\psi\lambda] = 0$. But it is not hard to see that

$$[\mathrm{ad}_y, \varphi\psi\lambda] = \varphi(\psi\lambda)_\eta,$$

where η is dual to y (i.e., the η derivative of the symbol of $\psi\lambda$ interpreted as an operator). That means that

$$\mathrm{ad}_{y\frac{\partial}{\partial t}}\left(\mathrm{ad}_{(-X)}^{\tilde{\alpha}}\,\mathrm{ad}_{Y}^{\tilde{\beta}}(\varphi\psi\lambda)\right)=\mathrm{ad}_{(-X)}^{\tilde{\alpha}}\,\mathrm{ad}_{Y}^{\tilde{\beta}}\left(\varphi(\psi\lambda)_{\eta}\frac{\partial}{\partial t}\right)+y\,\mathrm{ad}_{(-X)}^{\tilde{\alpha}}\,\mathrm{ad}_{Y}^{\tilde{\beta}}(\varphi_t\psi\lambda).$$

This means that the nonrigid bracket formula (5.13) may be imported into the pseudodifferential-coefficient case, except that wherever we had written (y) we should now have two terms, (y) together with the t derivative on φ (as in (5.13)) plus a term without (y) but with $\varphi\psi\lambda$ replaced by

$$\varphi(\psi\lambda)_{\eta}\frac{\partial}{\partial t}=\varphi\partial_{\eta}(\sigma(\psi\lambda))\partial_t^u,$$

i.e., ∂_t^u "sees" only u:

$$\begin{aligned}&[(T^p)_{\varphi\psi\lambda},C_I(x,y,t;D)]ZZu\\&=\sum_{\substack{t+j+|\gamma|\le p\\ t=\sum t_j}}F_{(0)}^{(\alpha_1+\beta_1+j)}(y)^{t_3+\gamma}\frac{X^{t_2}}{p^{t_2}}\frac{p^j}{p^{t_3+t_4}}(T^{p-j-|\gamma|})_{(\varphi\psi\lambda)_{(t_4)}^{(|\gamma|+t_3)}\partial_t^{t_4}}\frac{Y^{t_1}}{p^{t_1}}ZZu.\end{aligned}\tag{6.4}$$

Denoting by W a Z or spatial derivative, we have set

$$(\varphi\psi\lambda)_{(t_4)}^{(|\gamma|+t_3)}=W_t^{|\gamma|+t_3}D_{\eta}^{t_4}(\varphi\psi\lambda),\tag{6.5}$$

and here the "coefficient" $F_{(0)}^{(\alpha_1+\beta_1+j)}$ now denotes, analogously,

$$F_{(0)}^{(\alpha_1+\beta_1+j)}=C_{p,\alpha_1,\beta_1,j}\frac{\mathrm{ad}_Y^{\alpha_1}\,\mathrm{ad}_X^{\beta_1}\,\mathrm{ad}_T^{j}(C_I)}{\alpha_1!\beta_1!j!}$$

with, as before,

$$C_{p,\alpha_1,\beta_1,j}\lesssim C^j p^{j-t}.$$

Of course, this is only part of the bracket of a typical high-order derivative with $P=(\sum_{|I|\le 2})C_I Z^I$, which we have abbreviated as $P=C_I Z^2$. The fuller expression is

$$\begin{aligned}&[(T^p)_{\varphi\psi\lambda}T^sZ^q,C_I(x,D)Z^2]\\&\quad=[(T^p)_{\varphi\psi\lambda},C_I]T^sZ^{q+2}+(T^p)_{\varphi\psi\lambda}[T^sZ^q,C_I]Z^2\\&\qquad+C_I[(T^p)_{\varphi\psi\lambda},Z^2]T^sZ^q+C_I(T^p)_{\varphi\psi\lambda}[T^sZ^q,Z^2].\end{aligned}$$

In taking the norm of the last two terms, the operator $C_I(x, D)$ is certainly bounded in L^2 and will be brought out of the norm majorized by a (universal) constant. The first term on the right we have just expanded above in (6.4), so the only new term is the second on the right. But we always have

$$[T^s Z^q, C_I] = \sum_{\substack{s' \leq s \\ q' \leq q \\ q'+s' \geq 1}} \binom{s}{s'} \binom{q}{q'} \mathrm{ad}_T^{q'} \, \mathrm{ad}_Z^{q'} (C_I) T^{s-s'} Z^{q-q'}.$$

Finally, the order relations described above persist, and as long as we stay with *admissible* terms, we may iterate the a priori inequality, and when $(T^p)_{\psi\psi\lambda}$ has been exhausted, we observe that all terms contain only Z derivatives; when they bracket to produce new T's, there will be at most $p/2$ of them, half the original number, and we change the choice of ψ, ψ, and λ, perhaps differently in each term, chosen to correspond to the number of free derivatives in the term in question.

The nesting of the spatial open sets is as before; the conical ones and the λ's are similarly handled given the growth of their derivatives.

Chapter 7
General Sums of Squares of Real Vector Fields

7.1 A Little History

The Laplacian

$$\Delta = \sum_{1}^{n} \frac{\partial^2}{\partial x_j^2}$$

and the partial Laplacian

$$\Delta' = \sum_{1}^{n'<n} \frac{\partial^2}{\partial x_j^2}$$

are the simplest examples of sums of squares of (real) vector fields, and their regularity properties couldn't be more different: Δ is C^∞, Gevrey, and real analytic hypoelliptic, while Δ' is none of these.

It will be interesting to differentiate between full and partial hypoellipticity below, but first, I want to demonstrate, by way of examples studied by Baouendi, Oleinik–Radkevich, and Grushin, how the form of the vector fields seems to make an enormous difference.

The vector fields studied above, from the Heisenberg group, have the form

$$X_H = \frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial t}$$

and

$$Y_H = \frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial t}.$$

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_7, © Springer Science+Business Media, LLC 2011

Yet if we define the Oleinik-type vector fields

$$X_{O1} = \frac{\partial}{\partial x}, \quad X_{O2} = -\frac{y}{2}\frac{\partial}{\partial t},$$
$$Y_{O1} = \frac{\partial}{\partial y}, \quad Y_{O2} = \frac{x}{2}\frac{\partial}{\partial t},$$

so that $X_H = X_{O1} + X_{O2}$ and $Y_H = Y_{O1} + Y_{O2}$, we have seen above, after a *nontrivial* amount of work, that

$$P_H = X_H^2 + Y_H^2 = \left(\frac{\partial}{\partial x} - \frac{y}{2}\frac{\partial}{\partial t}\right)^2 + \left(\frac{\partial}{\partial y} + \frac{x}{2}\frac{\partial}{\partial t}\right)^2$$

is HE and AHE (and GHE in all Gevrey classes G^s with $1 \leq s$).

However, the proof that the very similar-looking operator

$$P_O = X_{O1}^2 + X_{O2}^2 + Y_{O1}^2 + Y_{O2}^2 = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{y}{2}\frac{\partial}{\partial t}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 + \left(\frac{x}{2}\frac{\partial}{\partial t}\right)^2$$

has the same properties requires only a few lines.

7.2 Proof for a Sum of Monomials

To illustrate the simplicity of the proof for monomials, consider the still simpler operator

$$P_G = D_x^2 + x^2 D_t^2$$

in $\mathbb{R}^2$. We have written D_w here for $\partial/\partial w$, $w = x, t$. Then it suffices to show that

$$\|D_t^k u\| \leq C C^k k!$$

Let $Z = D_x$ or $x D_t$. The obvious a priori estimate is

$$\|Z^2 v\| + \|Zv\|_{1/2} + \|v\|_1 \lesssim \|P_G v\| + \|v\|_0, \qquad v \in C_0^\infty,$$

and localization is needed only in the t variable, since for $x \neq 0$, P_G is elliptic (so any product localizing function, $\varphi(t)\psi(x)$, if differentiated, whenever ψ' enters we may stop since we are in the elliptic region where AHE is known).

Take $u \in C^\infty$ with $P_G u = f \in C^\omega$, and set $v = \varphi(t) D_t^k$ in the estimate

$$\|Z^2 \varphi D_t^k u\| + \|Z \varphi D_t^k u\|_{1/2} + \|\varphi D_t^k u\|_1 \lesssim \|P_G \varphi D_t^k u\| + \|\varphi D_t^k u\|_0.$$

Since $\varphi D_t^k P_G u$ is known and

$$[P_G, \varphi] D_t^k u \sim 2x^2 \varphi' D_t^{k+1} u + x^2 \varphi'' D_t^k u$$

(since $[D_x, \varphi] = 0$), perhaps a more useful version of the estimate would read

$$\begin{aligned} J_{k,\varphi} &\equiv \|D_x^2 \varphi D_t^k u\|_0 + \|x^2 D_t^2 \varphi D_t^k u\|_0 + \|x^2 D_t \varphi D_t^{k+1} u\|_0 + \|x^2 \varphi D_t^{k+2} u\|_0 \\ &\quad + \|Z \varphi D_t^k u\|_{1/2} + \|\varphi D_t^k u\|_1 \lesssim \|P_G \varphi D_t^k u\| + J_{k-1,\varphi^{(')}} \\ &\lesssim \|P_G \varphi D_t^k P_G u\| + J_{k-1,\varphi^{(')}} + J_{k-2,\varphi^{('')}}. \end{aligned}$$

Here $\varphi^{(')}$ (or $\varphi^{('')}$) designates the localizing function φ with *as many as* one (or two) derivatives, but perhaps fewer.

The essential property reflected here, and a by-product of the fact that the given vector fields are monomials, is that whenever such a vector field differentiates φ, there is a factor of x that together with D_t makes a new Z. Hence the induction may proceed at once, without the need to correct high powers (or any powers) of $T = D_t$.

For this last estimate may simply be iterated, with Ehrenpreis-type analytic families of localizing functions or Hörmander's version, where once φ receives a few derivatives it may be replaced by another, with slightly larger support, and derivatives estimated by powers of N, (the overall bound on the number of derivatives being estimated), though uniformly in N.

At any event, when the number of derivatives still in play decreases, a factor of CN to the corresponding power may enter, and when the number is reduced to essentially zero, a factor of $(CN)^N$.

Any reader who has persevered to this point will have no trouble seeing that analyticity of the solution, or AHE, is a stone's throw away with no careful localization of powers of D_t.

Thus, we reiterate, a sum of squares of monomials seems to have a vastly simpler analysis than a sum of squares of composite vector fields, even though the structure of the brackets of the vector fields seems similar.

Microlocally it is still possible that these operators P_H and P_O are essentially similar, but this author remains skeptical. The difficulty seems to be that the characteristic varieties $\Gamma_H = \{\xi - \frac{1}{2} y\tau = \eta + \frac{1}{2} x\tau = 0, \tau \neq 0\}$ and $\Gamma_O = \{\xi = \eta = x = y = 0, \tau \neq 0\}$, while both symplectic, seem substantially different, being, for example, of different codimensions.

7.3 Partial Regularity

In the case of monomials it is relatively easy to prove some partial hypoellipticity results. Here we give the simplest case, and then a more general result for rather general operators of Oleinik type.

Recently there has been study of Gevrey hypoellipticity when s takes on values other than $1/m$. Christ [Chr2] has very recently obtained sharp isotropic results in rational Gevrey classes that are better than those predicted by the subellipticity index. In this section and those that follow we improve on those regularity results by considering nonisotropic classes and show that all of these results are accessible with L^2 methods alone.

We consider here the particular, though apparently fairly typical, example

$$P = \frac{\partial^2}{\partial x^2} + \left(x\frac{\partial}{\partial t}\right)^2 + \left(x^2\frac{\partial}{\partial s}\right)^2 = X_1^2 + X_2^2 + X_3^2. \tag{7.1}$$

Theorem 7.1. *The operator P is G^d hypoelliptic for all $d \geq 3/2$.*

Remark. This theorem is due to M. Christ. However, our proof is both elementary and allows more precise results about partial regularity, where derivatives in different directions grow at different rates.

Theorem 7.2 (Bove–Tartakoff, 1996). *The operator P is G^{d_1,d_2,d_3} hypoelliptic for $d_1 \geq 7/6$, $d_2 \geq 1$, and $d_3 \geq 3/2$, and microlocally as well.*

Proof. We shall use the a priori estimate, for $v \in C^\infty$, any vector field W, and the localizing function $\varphi = \varphi_N$, as above:

$$\sum_j \|X_j\varphi W^p v\|_{L^2}^2 + \sum_j \|\varphi X_j W^p v\|_{L^2}^2 + \|\varphi W^p v\|_{1/3}^2$$
$$\leq C\left\{|(P\varphi W^p v, \varphi W^p v)_{L^2}| + \sum_j \|X_j(\varphi)W^p v\|_{L^2}^2 + \|\varphi W^{p-1}v\|_{L^2}^2\right\}, \quad v \in C^\infty. \tag{7.2}$$

(Note that we distinguish the different derivatives on φ, since the X_j carry coefficients that are powers of x, a fact that will be crucial in the sequel.)

We have used several facts in writing down this estimate: the form of P clearly allows us to bound the basic vector fields X_j on v, and thus the $1/3$ norm, since second brackets suffice to span the tangent space, and we have used a symmetric form on the left, with the X_j either before or after the localizing functions, since both will occur as errors. It follows that if $Pw = f \in G^s$ for some $s \geq 3$, the same is true of w locally.

We shall obtain bounds of the solution locally (in L^2 norms) of the type

$$||\varphi D_x^{\alpha_1} D_t^{\alpha_2} D_s^{\alpha_3} u||_{L^2} \leq CC^{|\alpha|}\alpha_1!^{d_1}\alpha_2!^{d_2}\alpha_3!^{d_3}, \quad |\alpha| \leq N,$$

and to do this it suffices, by integration by parts and a simple inductive hypothesis, to treat pure powers of D_x, D_t, and D_s.

We start with $W = D_t$, since this turns out to be the simplest. Then

$$[P, \varphi_N D_t^p] = D_x(\varphi_N)_x D_t^p + (\varphi_N)_x D_x D_t^p + x D_t x(\varphi_N)_t D_t^p + x^2(\varphi_N)_{tt} D_t^p + x^2 D_s x^2(\varphi_N)_s D_t^p + x^4(\varphi_N)_{ss} D_t^p.$$

Now, the (t, s) derivatives on φ_N are serious (recall that x derivatives leave us in the elliptic region where the result is known), but even these derivatives will not be harmful to a proof even of analyticity (in the variable t) *if* there is a corresponding gain in powers of D_t without losing the two good X's.

These two good X's may either come from P (one X comes from P in the first, third, and fifth terms on the right here, though in the first two terms the presence of an x derivative on φ_N lands us in the elliptic region) or are *created* by combining powers of the variable x with copies of D_t, namely once in the third and fifth terms and twice in the fourth and sixth, thus reducing the power p on D_t. Note that these (one or two) powers of D_t must still be commuted past φ_N to be in a useful position, and they may land on the localizing function. If they do, we must try to bring another D_t out, etc. But to simplify this exposition, we include only the principal term, ignoring these second-order brackets. Once we have X's on the left, one of these X's will be moved to the right (by integration by parts). We obtain

$$\begin{aligned}\left(\left[P, \varphi_N D_t^p\right] u, \varphi_N D_t^p u\right) = &\left(D_x(\varphi_N)_x D_t^p u, \varphi_N D_t^p u\right)_{L^2} \\ &+ \left((\varphi_N)_x D_x D_t^p u, \varphi_N D_t^p u\right)_{L^2} \\ &+ 2\left(x D_t(\varphi_N)_t D_t^{p-1} u, x D_t \varphi_N D_t^p u\right)_{L^2} \\ &+ \left(x D_t(\varphi_N)_{tt} D_t^{p-2} u, x D_t \varphi_N D_t^p u\right)_{L^2} \\ &+ 2\left(x^2 D_t(\varphi_N)_s D_t^{p-1} u, x^2 D_s \varphi_N D_t^p u\right)_{L^2} \\ &+ \left(x^2 D_t(\varphi_N)_{ss} D_t^{p-2} u, x^2 D_t \varphi_N D_t^p u\right)_{L^2},\end{aligned}$$

so that (recalling the norms are squared) for any $\varepsilon > 0$,

$$\begin{aligned}\left|\left([P, \varphi_N D_t^p] u, \varphi_N D_t^p u\right)_{L^2}\right| \le\; &C\|\varphi_N D_t^p u\|_{L^2}^2 + C C^{2|\alpha|}(2|\alpha|)! \\ &+ \varepsilon\|x D_t \varphi_N D_t^p u\|_{L^2}^2 + C_\varepsilon\left\{\|x D_t(\varphi_N)_t D_t^{p-1} u\|_{L^2}^2\right. \\ &\left.+ \|x D_t(\varphi_N)_{tt} D_t^{p-2} u\|_{L^2}^2\right\} \\ &+ C_\varepsilon\left\{\|x^2 D_t(\varphi_N)_s D_t^{p-1} u\|_{L^2}^2\right. \\ &\left.+ \|x^2 D_t(\varphi_N)_{ss} D_t^{p-2} u\|_{L^2}^2\right\}.\end{aligned}$$

After replacing the L^2 norm first on the right by again ε times the $1/3$ norm modulo a large constant times the -1 norm and using this to absorb one power of D_t, this leads, upon iteration p times, starting from (7.2) with $W = D_t$, to the bounds

$$\sum_j \|X_j \varphi_N D_t^p u\|_{L^2}^2 + \sum_j \|\varphi_N X_j D_t^p v\|_{L^2}^2 + \|\varphi_N D_t^p u\|_{1/3}^2$$
$$\leq C C_{u,Pu}^{2|\alpha|}(2|\alpha|)! + C^{2|\alpha|+2}(2|\alpha|)!\|u\|_{L^2}^2$$
$$= C C_{u,Pu}^{2|\alpha|}(2|\alpha|)! + C_u^{2|\alpha|+2}(2|\alpha|)!,$$

which yields analytic growth in the D_t derivatives.

Next we tackle D_s derivatives, which are not nearly as simple:

$$[P, \varphi_N D_s^p] = D_x(\varphi_N)_x D_s^p + (\varphi_N)_x D_x D_s^p + 2x D_t x(\varphi_N)_t D_s^p$$
$$+ x^2(\varphi_N)_{tt} D_s^p + 2x^2 D_s x^2(\varphi_N)_s D_s^p + x^4(\varphi_N)_{ss} D_s^p.$$

Again, the x derivatives on φ_N are not serious, and the last two terms may be treated precisely as we did with high powers of D_t: combining x^2 with D_s "creates" a "good" derivative (i.e., an X_j) that may be integrated to the right in the inner product. This exchange, a derivative on φ_N for a gain in p, i.e., a gain in power of D_s derivatives, is what has led to optimal (i.e., analytic) regularity all along.

But it is the third and fourth terms that are more troublesome, and where the new features arise. For combining only one power of x with D_s will not generate a "good" vector field sufficient to balance a derivative (D_t) falling on φ_N. Two powers of x are required. However, a little patience will produce two, even in this "worst-case scenario." For if we are repeatedly so unlucky (and all cases do occur) as to bracket with xD_t (and we ignore other contributions), we find that after *two* such brackets we may decrease p by one.

But something even better happens. Consider: we start with $X\varphi_N D_s^p u$, and after the first use of the basic a priori estimate we have come up with $x(\varphi_N)_t D_s^p u$. As with the proof in the general subelliptic case above, we do have a gain of $1/3$ derivative to make use of. That is, for the next iteration we write

$$\|x(\varphi_N)_t D_s^p u\|_{L^2} = \|\Lambda^{-1/3} x(\varphi_N)_t D_s^p u\|_{1/3}^2,$$

and treat $\Lambda^{-1/3} x(\varphi_N)_t D_s^p u$ as the new version of $\varphi_N D_s^p u$ whose X_j derivatives as well as $1/3$ norm are bounded in the estimate. (To justify treating $\Lambda^{-1/3} x(\varphi_N)_t D_s^p u$ as if it had compact support, one introduces a new function $\Psi \equiv 1$ near the support of all φ just to the right of $\Lambda^{-1/3}$. It is harmless there and when commuted past $\Lambda^{-1/3}$ for support considerations the error committed is of a full order lower.) In the next iteration, we will be led to analogous "errors," such as $\Lambda^{-1/3} x^2(\varphi_N)_{tt} D_s^p u$ with one more power of x and another (D_t) derivative of φ_N.

But now something very exciting happens. We *may* treat the x^2 together with D_s as a *good* vector field, namely as one of the X_j's. After these two iterations we have gone from $\|X_j \varphi_N D_s^p u\|_{L^2}$ to $\|X_j \Lambda^{-1/3} (\varphi_N)_{tt} D_s^{p-1} u\|_{L^2}$. After *three* of *these* iterations we will obtain

$$\begin{aligned}\|X_j \Lambda^{-1} (\varphi_N)_{tttttt} D_s^{p-3} u\|_{L^2} &\sim \|X_j (\Lambda^{-1} D_s)(\varphi_N)_{tttttt} D_s^{p-4} u\|_{L^2} \\ &\sim \|(X_j)(\varphi_N)_{tttttt} D_s^{p-4} u\|_{L^2},\end{aligned}$$

which exhibits a trade of *four* D_s derivatives for *six* derivatives on the localizing function φ_N.

This is the "trade-off" that leads to the Gevrey class $G^{3/2}$.

Finally, we turn to x derivatives, where the computations are simpler, but seem to depend on the preceeding ones: this time, $W = D_x$ and the a priori estimate reads essentially

$$\begin{aligned}&\sum_j \|X_j \varphi_N D_x^p u\|_{L^2}^2 + \sum_j \|\varphi_N X_j D_x^p u\|_{L^2}^2 + \|\varphi_N D_x^p u\|_{1/3}^2 \\ &\quad \leq C \left\{ \left| \left(\varphi_N D_x^p P u, \varphi_N D_x^p u\right)_{L^2} \right| + \sum_{j=1}^{3} \left| \left([\varphi_N D_x^p, X_j^2] u, \varphi_N D_x^p u \right)_{L^2} \right| \right. \\ &\qquad \left. + \sum_j \|\varphi_N^{(\prime)} D_x^p u\|_{L^2}^2 \right\}.\end{aligned}$$

In the bracket (second term on the right) the problem is now not keeping good vector fields: the D_x themselves are good. The difficult term that arises is that in which $x^2 D_s [D_x^p, x^2 D_s]$ (or $x D_t [D_x^p, x D_t]$ or double brackets) enters. All terms contribute $\underline{p}$ terms where one of the D_x differentiates x. Thus the error that arises is of the form $\underline{p} x D_s D_x^{p-1}$ (and $\underline{p} D_t D_x^{p-1}$) after the other X_j has been moved to the second member in the inner product. From the double brackets we may also have $\underline{p(p-1)} D_s D_x^{p-2}$, etc. That is, starting with $X \varphi_N D_x^p$ we now have essentially $\underline{p} X x \varphi_N D_s D_x^{p-2}$ or $\underline{p^2} X \varphi_N D_s D_x^{p-3}$, and similar terms with D_t.

Now the former, $\underline{p} X x \varphi_N D_s D_x^{p-2}$, turns out to be the worst in terms of managing growth of derivatives: there has been a "gain" of two D_x's, but a factor of p and a new D_s derivative. When this continues, what started as $X \varphi_N D_x^p$ becomes, after the next iteration, $\underline{p(p-1)} x^2 \varphi_N D_s^2 D_x^{p-3}$ or $\underline{p} \varphi_N D_s D_x^{p-2}$ or $\underline{p^3} X \varphi_N x D_s^2 D_x^{p-5}$, etc., where again, in the first term we may call $x^2 D_s$ an X. That is, this term is of the form $\underline{p^2} X \varphi_N D_s D_x^{p-3}$. *This*, iterated $p/3$ times, will lead to terms such as $\underline{p^{2p/3}} D_s^{p/3}$. Since we know that derivatives in s of the solution grow like the Gevrey class $G^{3/2}$, we will get

$$\|X \varphi_N D_x^p u\|_{L^2} \leq C^p p^{2p/3} p!^{(3/2)(1/3)} \leq C^p p!^{7/6}.$$

This is the origin of the $G^{7/6}$ behavior in x. Finally, the microlocalization poses no additional problems.

7.4 Other Special Cases, Leading to a General Conjecture

An analysis of the above proof, which certainly includes the statement of $G^{3/2}$ hypoellipticity, shows that it applies equally well to the somewhat more general operator

$$P = \frac{\partial^2}{\partial x^2} + \left(x^{p-1}\frac{\partial}{\partial t}\right)^2 + \left(x^{q-1}\frac{\partial}{\partial s}\right)^2 \tag{7.3}$$

for $1 \le p \le q$.

Theorem 7.3. *The operator P in (7.3) is G^d hypoelliptic for all $d \ge q/p$.*

Remark. This theorem is also due to M. Christ. However, our proof is both elementary and allows more precise results about partial regularity, where derivatives in different directions grow at different rates.

Theorem 7.4 (Bove–Tartakoff, 1996). *The operator P in (7.3) is G^{d_1,d_2,d_3} hypoelliptic for any $d_1 \ge 1 + 1/p - 1/q,\ d_2 \ge 1$, and $d_3 \ge q/p$.*

Theorem 7.5. *When $p = q = 1$ this yields analytic hypoellipticity, but in all other cases yields new examples of Gevrey classes of solutions to subelliptic partial differential equations.*

Theorem 7.6 (Bove–Tartakoff, 1996). *The Baouendi–Goulaouic example,*

$$P = D_x^2 + D_t^2 + x^2 D_s^2,$$

is G^{d_1,d_2,d_3} hypoelliptic for any $d_1 \ge 3/2,\ d_2 \ge 1$, and $d_3 \ge 2$ but not hypoelliptic in any smoother Gevrey class.

Proof. The first part is a special case of the above result. The sharpness comes by studying the function

$$u_\varepsilon(x,t,s) = \int_0^\infty \exp[i\rho^2 s - t\rho - \rho^2 x^2/2 - \rho^\varepsilon]\, d\rho$$

for $\varepsilon > 1$, which solves $Pu_\varepsilon = 0$, yet brief calculations show that u_ε satisfies

$$|\partial_t^k u_\varepsilon(0)| = \left|\int_0^\infty e^{-\rho^\varepsilon}\rho^k d\rho\right| \sim C^k k!^{1/\varepsilon},$$
$$|\partial_s^k u_\varepsilon(0)| = \left|\int_0^\infty e^{-\rho^\varepsilon}\rho^{2k} d\rho\right| \sim C^k k!^{2/\varepsilon},$$

and

$$|\partial_x^{2k} u_\varepsilon(0)| = \left|\int_0^\infty e^{-\rho^\varepsilon}\rho^{2k} k! d\rho\right| \sim C^k k!^{1+2\varepsilon} \sim C^k (2k)!^{1/2+1\varepsilon},$$

showing that for any $\varepsilon > 1,\ u_\varepsilon \in G^{1/2+1/\varepsilon,1/\varepsilon,2/\varepsilon}$ and no better.

More general situations clearly pose no additional difficulties, such as those studied (isotropically) in [Chr2]:

$$P = \frac{\partial^2}{\partial x^2} + a_1(x,t,s)\left(x^{p-1}\frac{\partial}{\partial t}\right)^2 + a_2(x,t,s)\left(x^{q-1}\frac{\partial}{\partial s}\right)^2, \tag{7.4}$$

where both $a_1(x,t,s)$ and $a_2(x,t,s)$ are strictly positive and belong to the Gevrey classes under consideration.

7.5 The General Conjecture and Result

The general conjecture suggested by these results, and voiced by Treves, concerns the iterated brackets of a set of real vector fields in the operator

$$P = \sum_{i=1}^{r} X_i^2,$$

where the X_j satisfy the Hörmander condition that their iterated brackets span the tangent space. Define $A_1 = \{X_1,\ldots,X_N\}$, $A_2 = A_1 \cup \{[X_i,X_j],\ i \neq j\},\ldots,A_N = \bigcup_{i=1}^{N-1} A_i \cup \{[X_{i_1},[X_{i_2},[\ldots,[X_{i_{N-1}},X_{i_N}]\ldots]\ ]\ ]\}$. We agree to call A_j, $j = 1,\ldots,N$, the "layers" of the Lie algebra of the tangent vector fields.

Now set $\Sigma_1 = \{(x,\xi)|X_j(x,\xi) = 0, j = 1,\ldots,r\}$, the multiple characteristic set of the operator P, $\Sigma_2 = \{(x,\xi)|Y(x,\xi)| = 0\forall Y \in A_2\}$, i.e., the characteristic set of the second layer of the Lie algebra. Proceeding in this way, we get to defining $\Sigma_N = \{(x,\xi)|Y(x,\xi)| = 0\forall Y \in A_N\}$. By the Hörmander assumption, Σ_N coincides with the null section of the cotangent bundle.

Of course the Σ_j are a decreasing finite sequence of subsets of the cotangent. We need not assume that they are manifolds; if they are not, just think of them as stratified varieties, whose strata are smooth manifolds.

Tentative statement 1: If all the Σ_j's are symplectic—i.e., every layer of each Σ_j is symplectic—then P is analytic hypoelliptic.

Tentative statement 2: Assume that Σ_ℓ is the first of the Σ_j that is not symplectic, i.e., there is at least one stratum with an involutive leaf. Then the operator P is G^s hypoelliptic if $s \geq N/\ell$.

The proofs above yield the second statement in the cases studied as well as in others under consideration by E. Bernardi, A. Bove, and the author. The first conjecture remains open.

Chapter 8
The $\overline{\partial}$-Neumann Problem and the Boundary Laplacian on Strictly Pseudoconvex Domains

One of the most urgent motivations for these analytic regularity issues came from the $\overline{\partial}$-Neumann problem, which we will define below.

This problem shares a great many features with a related operator, the so-called complex boundary Laplacian or Kohn Laplacian, which we also develop here, and which may be thought of roughly as a nonelliptic "Laplacian" living in the boundary of the original domain but involving only those holomorphic and antiholomorphic vector fields that are *tangent* to the boundary, and then the whole operator is restricted to the boundary.

The resulting operator, denoted by $\Box_b$, is precisely of the form studied in the previous chapters, since, using the Darboux theorem, the real and imaginary parts of those tangential vector fields have the same span as do the Heisenberg group vector fields, and the operator to be studied is a variable-coefficient, second-order operator constructed from real vector fields as studied in the previous chapters, and satisfying the a priori maximal and subelliptic estimates.

For the $\overline{\partial}$-Neumann problem, instead of dealing with a nonelliptic partial differential operator in a local manner, one deals with a "noncoercive" boundary value problem (the so-called $\overline{\partial}$-Neumann boundary condition) for an elliptic operator (the Laplacian in the Euclidean case).

The only real work for this case amounts to formulating the $\overline{\partial}$-Neumann problem in such a way that the previous techniques carry over nearly unchanged (in tangential directions), and handling the boundary conditions in an invariant manner (see below).

Real analytic regularity of these problems in a strictly pseudoconvex (or nondegenerate) setting was proved first globally by Tartakoff [T2], Derridj and Tartakoff [DT7], and Komatsu [Kom] in 1976, with local Gevrey results and even quasianalytic results at about the same time [T10], [DT8]. The crucial observation was that when the Levi form is nondegenerate, one may take a real vector field T, complementary to the holomorphic tangent space (to the boundary) $\mathcal{H}$, localize it to obtain ψT with ψ arbitrary, and modify *this* by a section Y_ψ of $\mathcal{H}$ so that

$$[\psi T + Y_\psi, \mathcal{H}] \subset \mathcal{H}.$$

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoelliptіcity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_8,

Local analyticity was proved in 1978, again for strictly pseudoconvex domains (or nondegenerate ones with maximal estimates) by Tartakoff and Treves independently [T4], [Tr4] using, respectively, purely L^2 methods and Fourier integral operators.

The results of Diederich and Fornaess in [DF], however, led one to conjecture that on weakly pseudoconvex domains one should at least have *global* analytic hypoellipticity. Indeed, S.-C. Chen has recently proved global analyticity for certain domains, including complete Reinhardt domains [Che1], [Che2].

8.1 Statement of the Theorems

The $\overline{\partial}$-Neumann problem is by now well known; we formulate it here for completeness.

We consider an open submanifold $\Omega \Subset \Omega'$, Ω' a real analytic complex Hermitian manifold, such that its boundary $b\Omega$ is a real analytic submanifold of Ω'. We denote by $C^\infty_{p,q}(\overline{\Omega})$ the space of (p,q)-forms which belong to $C^\infty(\overline{\Omega})$. The operator $\overline{\partial}$: $C^\infty_{p,q}(\overline{\Omega}) \to C^\infty_{p,q+1}(\overline{\Omega})$ has a formal adjoint $\overline{\partial}^*$, and $\mathcal{D}^{p,q}$ will denote the subspace

$$\mathcal{D}^{p,q} = \{v \in C^\infty_{p,q}(\overline{\Omega}) : (\overline{\partial}u, v)_\Omega = (u, \overline{\partial}^* v)_\Omega \text{ for all } u \in C^\infty_{p,q-1}(\overline{\Omega})\}.$$

The $\overline{\partial}$-Neumann problem consists in the following: given a (p,q) form α on $\overline{\Omega}$, find a (p,q) form u on $\overline{\Omega}$ satisfying the $\overline{\partial}$-Neumann boundary conditions $u \in \mathcal{D}^{p,q}$ and $\overline{\partial}u \in \mathcal{D}^{p,q+1}$ and such that

$$Q(u,w) \equiv (\overline{\partial}u, \overline{\partial}w)_{L^2(\Omega)} + (\overline{\partial}^* u, \overline{\partial}^* w)_{L^2(\Omega)} = (\alpha, w)_{L^2(\Omega)}, \quad \forall w \in \mathcal{D}^{p,q}$$

(so that $\Box u = (\overline{\partial}\,\overline{\partial}^* + \overline{\partial}^*\overline{\partial})u = \alpha$ in Ω).

It follows from standard arguments that if u is, in addition, say, in $C^1(\overline{\Omega})$ and $\overline{\partial}\alpha = 0$, then also $\overline{\partial}u = 0$ and $\overline{\partial}(\overline{\partial}^* u) = \alpha$, with $\overline{\partial}^* u$ orthogonal to the kernel of $\overline{\partial}$. This is the so-called Kohn solution or canonical solution to $\overline{\partial}$.

Since we are interested in the real analytic regularity of solutions to this problem, we always assume that the boundary is real analytic and we merely refer to the previous chapters or [K1], [D3], [T1], and [T3] for the C^∞ and Gevrey regularity of u up to the boundary.

Next, we describe the analogous problem for the operator $\overline{\partial}_b$ and its Laplacian $\Box_b$, living on the boundary of Ω, or, more generally, on a $(2n-1)$ real dimensional smooth Cauchy–Riemann (or C-R) manifold Γ, i.e., the complexified tangent space $\mathbb{C}T\Gamma = T^{1,0} \oplus T^{0,1} \oplus N$, where the subbundle $T^{0,1}$ ($=$ the conjugate of $T^{1,0}$) is assumed integrable and to have trivial intersection with $T^{1,0}$ and N has real dimension 1. If θ is a real nonvanishing one-form that annihilates $T^{1,0}$ (and hence its conjugate $T^{0,1}$), it determines the Levi form $\mathbf{L}_\theta$ defined on $T^{1,0}$ by

$$\mathbf{L}_\theta(V, W) = -i(d\theta)(v, \overline{W}).$$

(Any other choice θ' must be of the form $f\theta$, and so $\mathbf{L}_{f\theta} = f\mathbf{L}_\theta$.)

The manifold is called (strongly) pseudoconvex if the form $\mathbf{L}_\theta$ is (strictly) definite. We assume that we are given a Hermitian metric on Γ, i.e., a Riemannian metric on Γ extended to $\mathbb{C}T\Gamma$ in such a way that $T^{1,0} \perp T^{0,1}$. Let $E = (T^{1,0} \oplus T^{0,1})^\perp$. Than $\Lambda^{p,0}$ denotes the p-forms in $\mathbb{C}T\Gamma$ that annihilate $E \oplus T^{0,1}$ and $\Lambda^{p,q}$ the subbundle of $\Lambda^{p+q}(\mathbb{C}T\Gamma)$ generated by $\Lambda^{p,0} \wedge \Lambda^{0,q}$, $\Lambda^{0,q}$ the conjugate of $\Lambda^{q,0}$. Then $\overline{\partial}_b : C^\infty(\Lambda p, q) \to C^\infty(\Lambda p, q+1)$ is defined by

$$\overline{\partial}_b = \pi_{p,q+1} d,$$

where $\pi_{p,q+1}$ is the orthogonal projection of Λ^{p+q+1} onto $\Lambda^{p,q+1}$. It follows that $(\overline{\partial}_b)^2 = 0$. We denote by $\overline{\partial}_b^*$ the adjoint of $\overline{\partial}_b$ and define

$$\Box_b = \overline{\partial}_b^* \overline{\partial}_b + \overline{\partial}_b \overline{\partial}_b^*.$$

We refer to [FK] for full details.

By $\mathcal{A}(\overline{\Omega} \cap V)$ we denote the space of all forms that are real analytic up to the boundary of Ω in V.

Theorem 8.1. *In a neighborhood V of $p_0 \in \partial\Omega$, let Ω be strictly pseudoconvex with real analytic boundary. Then the $\overline{\partial}$-Neumann problem is analytic hypoelliptic up to the boundary near* 0; *that is, if u is a solution to the $\overline{\partial}$-Neumann problem and if $\alpha \in \mathcal{A}(\overline{\Omega} \cap V)$, then $u \in \mathcal{A}(\overline{\Omega} \cap V)$.*

Theorem 8.2. *Let V be a neighborhood of a point p_0 in a real analytic, strictly pseudoconvex C-R manifold Γ. Let u solve $\Box_b u = \alpha$ in V with α real analytic in V. Then u is real analytic in V.*

Remark 8.1. The analytic regularity result for $\Box_b$ on a strictly pseudoconvex domain follows from the results we have already proven, in previous chapters, and we shall not pursue the proof of Theorem 8.2 further.

8.2 Notation and a Priori Estimates

It is convenient to define $T = \partial/\partial t$ at this time; in our setting obviously we have $[T, L_j] = 0$.

(For the proof, however, $[T, L_j] = 0 \mod \{\{L_j\}, \{\overline{L_j}\}\}$ would suffice.)

Our notation for products of noncommuting vector fields is

$$W^{I_a} = W_{I_1} W_{I_2}, \cdots, W_{I_a} \quad \text{with} \quad |I_a| = a,$$

while D shall stand for any first derivative (always taken from a fixed, finite set of vector fields) and D_{tg} will stand for any D that is tangent to Γ. Standard arguments, such as the Sobolev embedding theorem cited earlier (see, for example, [T2]), show that to prove u real analytic in $\overline{\Omega} \cap V$, it suffices to show (for Theorem 8.1) that *for all* $V' \Subset V$,

$$\|(D)D_{tg}^{\beta}u\|^2_{L^2(\overline{\Omega}\cap V')} \le C_{V'}^{|\beta|+1}|\beta|! \qquad \forall \beta, \tag{8.1}$$

where here and in the sequel, enclosing an operator in parentheses means that it may or not be present. Thus in this case we may be measuring all tangential derivatives of $|\beta|$ and one nontangential one. The above estimate states the classical fact that it suffices, using the differential equation, to estimate high order tangential derivatives in L^2 norm, since the boundary is noncharacteristic for *any* elliptic operator, and $\Box$ is always elliptic. For Theorem 8.1, it suffices to estimate *for all* $V' \Subset V$,

$$\|D^{\beta}u\|^2_{L^2(V')} \le C_{V'}^{|\beta|+1}|\beta|! \qquad \forall \beta. \tag{8.2}$$

We shall sometimes work in a "special boundary chart" in the sense of Ash [Ash] (see also [FK] for complete details). This means that we take as an orthonormal frame for $T^{1,0}(\Omega')$ (complex) vector fields $\tilde{L}_1, \dots, \tilde{L}_{n-1}, \tilde{L}_n$ with $\tilde{L}_1, \dots, \tilde{L}_{n-1}$ tangent to Γ and $\tilde{L}_n r = 1$ on $\Gamma, r = \Im w - g(z, \overline{z})$. In terms of these vector fields the operator $\Box$ has the following simpler form: on a (p, q) form $\varphi = \sum \varphi_{IJ}\lambda^I\overline{\lambda}^q$,

$$\Box\varphi = (h^2 L_n\overline{L}_n + \sum_{j<n} \tilde{L}_j\overline{\tilde{L}}_j) \text{ Id } \varphi + \text{ lower order terms,}$$

while the $\overline{\partial}$-Neumann boundary conditions become that on Γ,

$$(i)\ \varphi_{IJ} = 0 \text{ if } n \in J \qquad \text{and} \qquad (ii)\ (\overline{\tilde{L}}_n + A)\varphi_{IJ} = 0 \text{ if } n \notin J, \tag{8.3}$$

where A is an analytic matrix that reflects the fact that $\overline{\partial}$ is not a homogeneous operator in general on the *coefficients* of forms since the frame will not always be $\overline{\partial}$ closed. These vector fields $\tilde{L}_j$ are just linear combinations of the L_j introduced above (on the boundary).

8.2.1 Maximal Estimate

Kohn has shown that for pseudoconvex domains in $\mathbb{C}^n$ one always has for $\varphi \in \mathcal{D}^{p,q}$ the estimate

$$\sum_{I,K}{}' \Big\| \sum_{j<n} L_j\varphi_{I,jK}\Big\|^2_{L^2(\Omega)} + \sum_{j\le n}\sum_{I,M}{}' \|\overline{L}_j\varphi_{I,M}\|^2_{L^2(\Omega)} + \sum_{I,n\in M}{}' \|\varphi_{I,M}\|^2_{H^1(\Omega)}$$
$$\le C\left(Q(\varphi,\varphi) + \|\varphi\|^2 L^2(\Omega)\right),$$

where $\sum_{I,K}{}'$ denotes summation over all strictly increasing indices of length p, K of length q. Here $\varphi_{I,jK}$ denotes 0 if j is in K and $\varphi_{I,\langle jK\rangle}$ times the sign of the permutation taking jK into the increasing q-tuple $\langle jK\rangle$ if j is not in K. From now on the norms will be taken over Ω without further mention, and the norm $\|\cdot\|_1$ denotes the Sobolev 1-norm measuring all first derivatives in L^2 norm.

Derridj has shown in [D2] (see also [S1]) that when the boundary is locally defined by $\Im w = h(|z|^2)$, one always has a stronger estimate, known as the "maximal" estimate:

$$\sum_{I,K}{}'\sum_{j<n}\|L_j\varphi_{I,K}\|^2_{L^2(\Omega)}+\sum_{j\le n}\sum_{I,M}{}'\|\overline{L}_j\varphi_{I,M}\|^2_{L^2(\Omega)}+\sum_{I,n\in M}{}'\|\varphi_{I,M}\|^2_{H^1(\Omega)}$$
$$\le C\left(Q(\varphi,\varphi)+\|\varphi\|^2_{L^2(\Omega)}\right),\qquad \varphi\in\mathcal{D}^{p,q}. \tag{8.4}$$

In fact, since $h(|z|^2)$ will not be constant, and is real analytic, we also have a *subelliptic* estimate:

$$\sum_{I,K}{}'\sum_{j<n}\|L_j\varphi_{I,K}\|^2_{L^2(\Omega)}+\sum_{j\le n}\sum_{I,M}{}'\|\overline{L}_j\varphi_{I,M}\|^2_{L^2(\Omega)}+\sum_{I,n\in M}{}'\|\varphi_{I,M}\|^2_{H^1(\Omega)}$$
$$+\sum_{I,M}{}'\|\varphi_{I,M}\|^2_{\varepsilon(\Omega)}\le C\left(Q(\varphi,\varphi)+\|\varphi\|^2_{L^2(\Omega)}\right),\qquad \varphi\in\mathcal{D}^{p,q}, \tag{8.5}$$

for some ε and all $\varphi\in\mathcal{D}^{p,q}$ supported in a fixed neighborhood of p_0 (cf. [K2]). Here the norm $\|\cdot\|_{\varepsilon(\Omega)}$ denotes the Sobolev ε-norm in Ω; that is, in local coordinates, we may define a *tangential Laplacian* Λ^{tg} with symbol $(1+|\xi'|^2)^{1/2}$, $\xi'=(\xi_1,\dots,\xi_{2n-1})$, and then define

$$\|w\|^2_{\varepsilon(\Omega)}=\|\Lambda^{\varepsilon}_{tg}w\|^2_{L^2}+\|\Lambda^{\varepsilon-1}_{tg}D_nw\|^2_{L^2}.$$

This definition is of course not invariant, but different choices of local coordinates and normal derivatives D_n give equivalent a priori estimates in view of the noncharacteristic nature of the boundary.

Actually, we do not use the subelliptic estimate. What we do use is a *compactness* estimate (which follows easily from the subelliptic one): for arbitrary $\mathcal{K}$ there exists a constant $C_{\mathcal{K}}$ such that

$$\sum_{I,K}{}'\sum_{j<n}\|L_j\varphi_{I,K}\|^2_{L^2(\Omega)}+\sum_{j\le n}\sum_{I,M}{}'\|\overline{L}_j\varphi_{I,M}\|^2_{L^2(\Omega)}+\sum_{I,n\in M}{}'\|\varphi_{I,M}\|^2_{H^1(\Omega)}$$
$$+\mathcal{K}\sum_{I,M}{}'\|\varphi_{I,M}\|^2_{0(\Omega)}\le CQ(\varphi,\varphi)+C_{\mathcal{K}}\|\varphi\|^2_{-1(\Omega)},\qquad \varphi\in\mathcal{D}^{p,q}. \tag{8.6}$$

We shall take $D_n=\mathcal{V}$ on the boundary to be JT, J the complex structure tensor from $\mathbb{C}^n$. Then as in [T4], [T5], we write $X_1,\dots,X_{2n-2}$ for the real and imaginary parts of the $L_j,\ j < n$, so far defined only on Γ, and extend the X_j and T, as well as $\mathcal{V}$, in such a way that *near* Γ,

$$[X_j, \mathcal{V}] = [X_j, T] = [X_j, \mathcal{V}] = 0 \qquad (\text{and so } [[X_j, X_k], \mathcal{V}] = 0), \tag{8.7}$$

and also so that it continues to be the case that $\mathcal{V} = JT$. In fact, we start by using the Cauchy–Kovalevskaya theorem to extend $\mathcal{V}$ in such a way that $[\mathcal{V}, J\mathcal{V}] = 0$ near Γ. Then extending the X_j and T via $[Xj, \mathcal{V}] = 0$ and $T = q\mathcal{V}$, it follows that $[X_j, T]$ is zero near Γ, since $[\mathcal{V}, [X_j, T]] = 0$ by the Jacobi identity.

For later use, we also list the maximal estimate for $\Box_b$ in the quadratic form sense, proved in this context by Grigis–Sjöstrand [GS]: for arbitrary $\mathcal{K}$ there exists a constant $C_{\mathcal{K}}$ such that for all $\varphi \in \mathcal{D}^{p,q}$,

$$\sum_{I,K}{}' \|X_j\varphi\|^2_{L^2} + \mathcal{K}\sum_{I,K}{}' \|\varphi_{I,K}\|^2_{L^2} \le CQ_b(\varphi,\varphi) + C_{\mathcal{K}}\|\varphi\|^2_{-1}, \tag{8.8}$$

where

$$Q_b(v, w) = (\bar\partial_b v, \bar\partial_v w)_{L^2} + (\bar\partial_b^* v, \bar\partial_b^* w)_{L^2}. \tag{8.9}$$

To prove Theorem 8.1, we shall insert $\varphi = \psi Y^{I_a} u \in \mathcal{D}^{p,q}$ into (8.6) with ψ of compact support, with each Y denoting T or an X. While u satisfies the $\bar\partial$-Neumann boundary conditions, $\psi Y^{I_a} u$ will satisfy in general only the first of them, even though $[\bar{L}_n, X \text{ or } T] = 0$ on the boundary. It will, however, be very useful to use localizing functions ψ such that $\bar{L}_n\psi = 0$ *on the boundary*. But this is easy: we may specify ψ to be arbitrary on the boundary and let the equation

$$\bar{L}_n\psi = 0 \qquad \text{on } \Gamma \tag{8.10}$$

determine the normal derivative of ψ on the boundary. Then any extension off of Γ with this normal derivative will do: it may easily be taken to have compact support in the whole space, and, as we shall see later, satisfy rather stringent conditions on how its high derivatives grow. Actually, for later reference, we indicate here how to prescribe cut-off functions like ψ above but taht have $\bar{L}_n\psi = 0$ to high order on Γ: since $\bar{L}_n r = 1$ on the boundary, if we initially extend an arbitrary ψ to be *independent* of $s = \Im w$, and with this ψ let

$$\psi_{(m)} = \sum_{j\le m} (-ir)^j T^j \psi/j!, \tag{8.11}$$

it follows that indeed $\bar{L}_n\psi_{(m)} = 0$ to high order on Γ, while $L_n\psi_{(m)} = (T\psi)_{(m-1)} + (T\psi)_{(m)}$. Note that ψ may be arbitrary on Γ and we may multiply $\psi_{(m)}$ by a function equal to one near Γ but vanishing outside a small neighborhood of Γ. Usually we shall drop the subscript on $\psi_{(m)}$.

While it is quite clear that the estimation of $\psi Y^{I_a} u$ for arbitrary I_a will be the most difficult when $Y^{I_a} = T^a$, estimating these derivatives, as we shall see, will lead us to estimate arbitrary $Y^{I_a} u$, so it is plainly best to do them all at the same time.

For this problem, smooth functions satisfying these boundary conditions do *not* satisfy a "coercive" or "elliptic" a priori estimate as would the Dirichlet or Neumann boundary conditions, but rather, for many domains of interest including the "strictly pseudoconvex" domains (the biholomorphic images of strictly convex domains) a subelliptic one reminiscent of those we have studied and, in fact, our starting point in the creation of $(T^p)_\varphi$:

$$|||v|||^2_{\frac{1}{2},t} \lesssim |(\Box v, v)_{L^2(\Omega)}| + \|v\|_0^2, \qquad v \in \mathcal{D}(\overline{\partial}^*),$$

which in terms of complex analysis may be written

$$|||v|||^2_{\frac{1}{2},t} \lesssim Q(v, v)_{L^2(\Omega)} + \|v\|_0^2,$$

where we use the notation

$$\Box = \overline{\partial}\overline{\partial}^* + \overline{\partial}^*\overline{\partial}$$

and

$$Q(v, w) = (\overline{\partial} v, \overline{\partial} w) + (\overline{\partial}^* v, \overline{\partial}^* w).$$

The boundary condition $v \in \mathcal{D}(\overline{\partial}^*)$ is self-explanatory, where $\overline{\partial}^*$ is the L^2 adjoint of $\overline{\partial}^*$. The remarkable thing about this formulation of the $\overline{\partial}$-problem is that a solution $u \in \mathcal{D}(\overline{\partial}^*)$ to

$$Q(u, v) = (f, v) \qquad \forall v \in \mathcal{D}(\overline{\partial}^*)$$

with $\overline{\partial} f = 0$ automatically satisfies $\Box u = f$ (taking all $v \in C_0^\infty(\Omega)$), and consequently, integrating by parts for arbitrary $v \in \mathcal{D}(\overline{\partial}^*)$, one finds that $\overline{\partial} u = 0$ in Ω (and is smooth up to the boundary if f is) and thus that

$$\overline{\partial}(\overline{\partial}^* u) = f.$$

This solution, $w = \overline{\partial}^* u$ to $\overline{\partial} w = f$ given $\overline{\partial} f = 0$, is called the "canonical" or "Kohn" solution, and the operator $f \to w = Nf$ the *Neumann* operator.

This estimate is on forms, for convenience taken to be of type $(0, 1)$, satisfying the $\overline{\partial}$-Neumann boundary conditions. The norm $|||v|||^2_{\frac{1}{2},t}$ here stands for

$$|||v|||^2_{\frac{1}{2},t} = \|\Lambda_{tg}^{1/2} v\|^2_{L^2(\Omega)} + \|\Lambda_{tg}^{-1/2} D_n v\|^2_{L^2(\Omega)}$$

in local coordinates, where Λ_{tg} is the pseudodifferential operator with symbol $(1 + |\xi'|^2)^{1/2}$ and D_n is differentiation normal to the boundary of Ω.

In working with this estimate, both in order to prove C^∞ regularity and higher regularity, one needs to put the boundary conditions in "translation-invariant" form, so that one may split tangential operations and normal ones. This can be done, using the "special frames" of J.J. Kohn, so that component-wise, tangential differentiation preserves the boundary conditions. In the simplest (flat) case, admittedly not strictly pseudoconvex, the boundary conditions (in $\mathbb{C}^n$, $\Im z_n \geq 0$) are that

- The nth component, v_n, vanishes on the boundary of Ω and the free boundary condition (automatically satisfied by a solution) that permits us to solve $\overline{\partial}$ itself is that
- the (complex) normal derivative of the remaining components vanishes on the boundary: $\partial w/\partial \overline{z}_n \omega_j = 0$ on $\partial\Omega$, $j < n$.

The proofs in previous chapters carry through, if one is careful to treat tangential differentiation first and then use the philosophy that tangential information plus use of the equation will provide normal information as well, provided the boundary is noncharacteristic, which is certainly the case here, since $\Box$ is elliptic.

Thus, for instance, we mollify tangentially, choose our normal derivative in such a way that it commutes with the tangential operators, etc.

When it comes to defining $(T^p)_\varphi$, or $(T^p)_{\varphi\psi\lambda}$, we do so tangentially only, but in a way that is "constant near the boundary" so that $[D_n, (T^p)_{\varphi\psi\lambda}] = 0$, for instance.

Finally, general strictly pseudoconvex domains do not involve the Heisenberg group vector fields, but still have the same nondegeneracy condition, referred to as "symplecticity": in the special complex local frame, there are complex vector fields $\{L_1, \ldots, L_n\}$ with L_j tangential to the boundary of Ω for $j < n$ (i.e., they annihilate the defining function of the boundary), and also $T = L_n - \overline{L_n}$, with the $\{L_j, \overline{L_k}\}$ independent and the crucial matrix c_{jk} given by

$$[L_j, \overline{L_k}] \equiv c_{jk} T \qquad \mod \{L_\ell, \overline{L_\ell}\}$$

with

$$\det c_{jk} \neq 0.$$

The matrix c_{jk} is called the Levi matrix.

Under these conditions, the celebrated theorem of Darboux ensures the existence of real coordinates such that the real span of the $\{L_j, \overline{L_k}\}$, $j, k < n$, on the boundary is the same as that of the Heisenberg group vector fields commuting with T.

8.3 The Heat Equation for $\Box_b$

Together with Nancy Stanton, we have studied the partial regularity of the heat equation for $\Box_b$ on strictly pseudoconvex domains in [ST1], [ST2].

While the main thrust of the longer of those two references [ST2] was the establishment of a heat kernel, we also looked at Gevrey/analytic regularity.

Perhaps because the problem split so cleanly into time and space, it came as no surprise that we were able to show that while mixed derivatives and purely time derivatives grew in G^2, purely spatial derivatives of the solution grew analytically.

In fact, although we did not state this explicitly, mixed derivatives would grow as a mixture of the two: locally and uniformly in ℓ, κ,

$$|D_t^\ell D_x^\kappa u| \le C^{\ell+|\kappa|}\ell!(\kappa!)^2.$$

The proof is a direct application of the present techniques.

8.4 Weakly Pseudoconvex Domains

For less-pseudoconvex domains, for example with some zero eigenvalues of the Levi matrix at certain points, the local and global analytic regularity is less well understood, though subellipticity still forces local C^∞ and Gevrey regularity with the critical Gevrey index tied to the degree of subellipticity.

8.5 Global Regularity

The global situation is quite different: while the global regularity results (even real analytic ones) are far simpler to prove than local ones, owing to the possibility of far less subtle localization, there are a weakly pseudoconvex domain (the "worm domain") with smooth boundary on which the $\bar{\partial}$-Neumann problem is not globally C^∞-hypoelliptic (in $\mathbb{C}^2$) and also a bounded domain in $\mathbb{C}^2$ with real analytic boundary on which this problem is not real analytic hypoelliptic (cf. [Chr6], [Chr3]). By contrast, the global analytic regularity for strongly pseudoconvex domains was proved by Boutet de Monvel [B2] in Stein manifolds in a very brief paper using an ingenious argument.

Chapter 9
Symmetric Degeneracies

9.1 Weakly Pseudoconvex Domains

As a first step in treating weakly pseudoconvex domains, in 1988 [DT1] Derridj and Tartakoff treated certain kinds of spherically symmetric degeneracies. That is, the domain was given, for real analytic $h(s)$, by

$$\Im w > h(|z'|^2), \quad h(s) > 0,\ s \neq 0,\ h(0) = 0,$$

and both $\Box_b$ and the $\bar{\partial}$-Neumann problems were shown to be locally real analytic hypoelliptic at the origin in $\mathbb{C}^n$, where $z = (z_1, \ldots, z_{n-1})$.

Since $h(s)$ is real analytic, its zero at the origin is of finite order, and it is not difficult to show that one has a subelliptic estimate for both problems [Hö1], [RS] and also a "maximal" estimate in which all of the tangential holomorphic and antiholomorphic vector fields are bounded optimally. That is, with

$$g(z, \bar{z}) = h(|z|^2),$$

$$L_j = \frac{\partial}{\partial z_j} + i g_{z_j} \frac{\partial}{\partial t} = \frac{\partial}{\partial z_j} + i \bar{z}_j h'(|z|^2) \frac{\partial}{\partial t}$$

working for now with $\Box_b$, and the frame $\{L_1, \ldots, L_{n-1}\}$ and its conjugate, $\{\overline{L_1}, \ldots, \overline{L_{n-1}}\}$, (with $[L_j, L_k] = 0$,) and the compact maximal estimate, expressed in terms of the components $\{v_j\}_{j<n}$ of a form of type $(0, 1)$, resembles one we have seen before: for all K, there exists a constant C_K such that

$$\sum_1^{n-1} \left\{ \|L_j v_k\|_0^2 + \|\overline{L_j} v_k\|_0^2 + K \|v_j\|_0^2 \right\} \lesssim Q_b(v, v) + C_K \|v\|_{-1}^2 \tag{9.1}$$

with the usual definition, as above, of Q_b.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_9,

Now, since $h(s)$ is real analytic, it satisfies a relationship of the form

$$h'(s) + sh''(s) = \frac{\mu}{2}\, h'(s)$$

(suppose, in the power series for h at the origin, the first nonzero coefficient is a_m; then

$$h'(s) + sh''(s) = \sum_m^{\infty}(1 + j)a_j s^j = \frac{\mu}{2}\, h'\,(s))\,.$$

Now as before, what will matter is to counteract the coefficient of

$$T = \frac{\partial}{\partial t}$$

in the bracket

$$[L_j, \varphi T] \equiv (L_j\varphi)T \mod \{L_k, \overline{L_k}\}$$

and, if possible, to do so with a linear combination of the $\overline{L_m}$. Then we should be able to add the conjugate of that correction, which will work for brackets with $\overline{L_j}$, and all errors in brackets with $\Box_b$ will be in "directions" optimally controlled by the estimate above.

Fortunately, we do not have far to look.

A first observation is that any localizing function φ need depend only on t, since if we write a general localizing function as a product

$$\varphi(z, \bar{z}, t) = \varphi_1(z, \bar{z})\, \varphi_2(t),$$

and a derivative lands on φ_1, the support of that derivative of φ_2 will be in a region where $|z|$ is bounded away from zero and the Levi form is invertible, a case already well studied in [T4] and [Tr4]; hence the solution is known to be analytic there.

Once the localizing function may be viewed as a function of t alone, the above reads

$$[L_j, \varphi T] = (ig_j\varphi_t)T,$$

since $[L_j, T] = 0$.

Now the vector field

$$M = \mu^{-1} \sum z_k L_k$$

has the needed properties if μ is defined as above, since a straightforward calculation gives

$$[L_j, \overline{M}] = -ig_{z_j} T + ((L_j \mu^{-1})\mu)\, \overline{M},$$

and hence, writing φ' for φ_t, since we have seen that in effect, φ need only depend on t, we have

$$\begin{aligned}[L_j, \varphi T + \varphi'\overline{M}] &= (L_j\varphi)T + \varphi_t[L_j, \overline{M}] + L_j(\varphi')\overline{M} \\ &= ig_{z_j}\varphi' T + \varphi'\{-ig_{z_j}T + ((L_j\mu^{-1})\mu)\,\overline{M}\} + ig_{z_j}\varphi''\overline{M} \\ &= \varphi'((L_j\mu^{-1})\mu)\,\overline{M} + ig_{z_j}\varphi''\overline{M}.\end{aligned}$$

The next step should be (and is)

$$\begin{aligned}\left[L_j, \varphi T^2 + \varphi'\overline{M}T + \varphi''\frac{\overline{M}^2}{2!}\right] &= \left[L_j, \varphi T + \varphi'\overline{M}\right]T + \left[L_j, \varphi''\overline{M}^2/2\right] \\ &= \varphi'((L_j\mu^{-1})\mu)\,\overline{M}T + ig_{z_j}\varphi''\overline{M}T \\ &\quad + ig_{z_j}\varphi''\overline{M}^2/2 + \varphi''[L_j, \overline{M}]\overline{M}/2 \\ &\quad + \varphi''\overline{M}[L_j, \overline{M}]/2 \\ &= \varphi'((L_j\mu^{-1})\mu)\,\overline{M}T + ig_{z_j}\varphi''\overline{M}T \\ &\quad + ig_{z_j}\varphi''\overline{M}^2/2 + \varphi''\{-ig_{z_j}T. \\ &\quad + ((L_j\mu^{-1})\mu)\,\overline{M}\}\,\overline{M} + \varphi''\left[\overline{M}, -ig_{z_j}T \right. \\ &\quad \left. + ((L_j\mu^{-1})\mu)\,\overline{M}\right]/2 \\ &= \varphi'((L_j\mu^{-1})\mu)\,\overline{M}T + ig_{z_j}\varphi''\overline{M}^2/2 \\ &\quad + ((L_j\mu^{-1})\mu)\,\overline{M}^2 - i\varphi''(\overline{M}g_{z_j})T/2 \\ &\quad + \varphi''\overline{M}((L_j\mu^{-1})\mu)\overline{M}/2,\end{aligned}$$

but here the algebra seems to become too complex to handle. And the first term on the right lacks the essential g_{z_j} to go with T, which is all that we have been able to correct so far.

But we do have

$$[\overline{M}, L_k] = i\,g_{z_k}T + e_k\overline{M}, \qquad \text{where } e_k = (L_k\mu^{-1})\mu,$$

and

$$\begin{aligned}\overline{M}g_{z_j} &= \mu^{-1}\sum \overline{z_k}\,\frac{\partial}{\partial\overline{z_k}}g_{z_j} = \mu^{-1}\sum \overline{z_k}\,\frac{\partial}{\partial\overline{z_k}}\,\overline{z_j}h'(|z|^2) \\ &= \mu^{-1}\sum \overline{z_k}\left\{\delta_{jk}h'(|z|^2) + \overline{z_j}\,z_k\,h''(|z|^2)\right\} \\ &= \mu^{-1}\overline{z_j}\left\{h'(|z|^2) + |z|^2\,h''(|z|^2)\right\} = \mu^{-1}\,(\mu/2)\,\overline{z_j}\,h'(|z|^2) = g_{z_j}/2,\end{aligned}$$

so that

$$[\overline{M}, [\overline{M}, L_j]] = i(g_{z_j}/2)T + (\overline{M}e_j)\overline{M}$$

etc., and in general,

$$\mathrm{ad}^{\ell}_{\overline{M}} L_j = i(g_{z_j}/2^{\ell-1})T + (\overline{M}^{\ell-1}e_j)\overline{M}.$$

Our aim, then, is to build more complicated sums of powers of M and $\overline{M}$ to correct φT^p. But since each step is to correct the previous one, perhaps modulo some terms of precisely balanced lower order, we are obliged to seek coefficients $\tilde{A}^a_{a'}, \overline{A}^a_{a'}$ and set

$$\tilde{N}_a = \sum_{0 \le a' \le a} \tilde{A}^a_{a'} \frac{M^{a'}}{a'!}, \qquad \overline{N}_b = \sum_{0 \le b' \le b} A^b_{b'} \frac{\overline{M}^{b'}}{b'!}.$$

For suitably chosen $\tilde{A}^a_{a'}$ and $A^a_{a'}$, to be made precise below, we have the following result.

Proposition 9.1. *The operators $\tilde{N}_a$ and $\overline{N}_b$ satisfy the following commutation relations: for $a, b \ge 1$,*

$$[L_k, \overline{N}_b] = -ig_{z_k}\overline{N}_{b-1}T - \sum_{j+b''=1}^{b} S_j^{b''+j} 2^{-j} \frac{e_k^{(b''-1)}\overline{M}}{b''!} \overline{N}_{b-(b''+j)}, \tag{9.2}$$

where

$$e_i^{(d-1)} = \overline{M}^{d-1}((L_j\mu^{-1})\mu), \qquad \mu = \mu(|z|^2),$$

$$[\overline{L}_i, \overline{N}_b] = -\sum_{k+b''=1}^{b} S_k^{b''+k} 2^{-k} \overline{N}_{b-(b''+k)} \times \left(\frac{e_{i,1}^{(b''-1)}\overline{L}_i + e_{i,2}^{(b''-1)}\overline{M}}{b\text{"}!} \right), \tag{9.3}$$

where

$$e_{i,1}^{(b''-1)}\overline{L}_i + e_{i,2}^{(b''-1)}\overline{M}^{d-1} = \mathrm{ad}^{b''}_{\overline{M}}(\overline{L}_i),$$

$$[\overline{L}_i, \tilde{N}_a] = \tilde{N}_{a-1} i g_{z_i} T + \sum_{j+a''=1}^{a} S_j^{a''+j} (-2)^{-j} \frac{e_{i,3}^{(a''-1)}M}{a''!} \tilde{N}_{a-(a''+j)}, \tag{9.4}$$

where

$$e_{i,3}^{(a''-1)} = M^{a''-1}((\overline{L}_i\mu^{-1})\mu), \qquad \mu = \mu(|z|^2),$$

and

$$[L_i, \tilde{N}_a] = \sum_{k+a''=1}^{a} \left(\frac{e_{i,4}^{(a''-1)} L_i + e_{i,5}^{(a''-1)} M}{a''!} \right) \times (-2)^{-k} S_k^{a''+k} \tilde{N}_{a-(a''+k)}, \quad (9.5)$$

where

$$e_{i,4}^{(a''-1)} L_i + e_{i,5}^{(a''-1)} M = -\widetilde{\mathrm{ad}}_{\overline{M}}^{a''} (L_i),$$

and

$$\mathrm{ad}_A^c B = [[\ldots [B, A], A], \ldots, A], \qquad \widetilde{\mathrm{ad}}_A^c B = [A, [A, [\ldots, B]] \ldots].$$

Here the S_c^d are suitably bounded constants that permit us to raise and lower indices:

$$A_q^p = \sum_{j=0}^{p-q} S_k^{c+j} A_{q-c}^{p-(c+j)},$$

and the $A_{a'}^a$ and $\tilde{A}_{a'}^a$, are actually chosen by recursion, for in order to have, for example, (9.2), it turns out by a simple calculation that we need

$$\sum_{b' \leq b} A_{b'}^b \sum_{b''=1}^{b'} \frac{1}{b''!} \frac{1}{2^{b''-1}} g_{z_k} T \frac{\overline{M}^{b'-b''}}{(b'-b'')!} = g_{z_k} \overline{N}_{b-1} T,$$

and so we must have

$$\sum_{b' \leq b} A_{b'}^b \sum_{b''=1}^{b'} \frac{1}{b''!} \frac{1}{2^{b''-1}} \frac{\overline{M}^{b'-b''}}{(b'-b'')!} = \overline{N}_{b-1} = \sum_{b_1 \leq b-1} A_{b_1}^{b-1} (\overline{M}^{b_1} / b_1!),$$

which in turn requires, equating coefficients of like powers of $\overline{M}$, that for each b_1 (summing over all pairs (b', b'') with $b' - b'' = b_1$),

$$\sum_{b''=1}^{b-b_1} A_{b_1+b''}^b \frac{1}{b''!} \frac{1}{2^{b''-1}} = A_{b_1}^{b-1}.$$

Similarly, the condition on $\tilde{A}_{a'}^a$ turns out to be that

$$\tilde{A}_{a'}^a = (-1)^{(a-a_1)} A_{a_1}^a.$$

Things become a bit more transparent if we write $A_{b'}^b = 2^{b'-b} B_{b'}^b$. Then the requirement is that we have $B_q^q = 1$ for each q and, for each b and b_1,

$$\sum_{b''=1}^{b-b_1} B_{b_1+b''}^b \frac{1}{b''!} = B_{b_1}^{b-1}. \quad (9.6)$$

Lemma 9.1. *The unique solution to (9.6) with $B_q^q = 1$ for all q and $B_0^q = (-1)^q$ for all q is given by*

$$B_s^r = \left(\left(\frac{t}{e^t - 1}\right)^{r+1}\right)^{(r-s)}(0)/(r-s)!$$
$$= \text{the coefficient of } t^{r-s} \text{ in } \left(\frac{t}{e^t-1}\right)^{r+1}.$$

Proof. A straightforward expansion, using Leibnizs' rule, of

$$\left(\left(\frac{t}{e^t-1}\right)^{b-1}\left(\frac{e^t-1}{t}\right)\right)^{(b-1-b_1)}(0)/(b-1-b_1)!$$

together with the fact that $\left(\frac{e^t-1}{t}\right)^{(s)}(0)/s! = 1/(s+1)!$ yields (3.8). Uniqueness is clear from the recurrence formula. We merely refer to Hirzebruch's book [Hi], Lemma 1.7.1, for the initial conditions $B_0^q = (-1)^q$, since we make no use of them. □

Proposition 9.2. *The results of commuting $\tilde{N}_a$ and $\overline{N}_b$ with functions are as follows:*

$$[\tilde{N}_a, f(z,\bar{z},t)] = \sum_{j+k=1}^{a} \frac{M^k f}{k!}(-2)^{-j} S_j^{k+j}\tilde{N}_{a-(j+k)}$$

and

$$[\overline{N}_b, f(z,\bar{z},t)] = \sum_{j+k=1}^{b} \frac{\overline{M}^k f}{k!} 2^{-j} S_j^{k+j}\overline{N}_{b-(j+k)}.$$

We will see the proof in the next chapter.

Lemma 9.2. *There exists a constant C such that for $s \le r$ we have*

$$|B_s^r| \le C^r.$$

Proof. The Cauchy integral formula reads, in view of the definition of $\overline{N}$,

$$B_s^r = \frac{1}{2\pi i}\int \frac{\left(\frac{\zeta}{e^\zeta - 1}\right)^{r-1}}{\zeta^{r-s+1}}\, d\zeta,$$

since $\zeta/(e^\zeta - 1)$ is holomorphic, where the integration is over a small contour containing the origin. On such a contour the integrand is bounded by C^r (recall that $r \ge s > 0$ in the definition of $\overline{N}$). □

Lemma 9.3. *There exists a constant C such that for $s \le r$,*

$$|S_s^r| \le C^r,$$

Proof. This follows from a close examination of the definition of B_s^r above, which allows us to see that S_s^r is the sum of all $d_{i_1} d_{i_2} \cdots d_{i_r}$ such that $\sum i_j = s$, and for each k, $i_k \in \{0, 1, \ldots, k-1\}$, where $d_j = (r/(e^t - 1))^{(j)}(0)/j!$. Clearly, $|d_j| \leq A^j$ for some universal constant A, so each product is bounded by A^s in absolute value.

Thus S_s^r is bounded by A^s times the number of partitions of s into r nonnegative integers subject to the above conditions. But this number of partitions is the coefficient of t^s in $(1)(1 + t)(1 + t + t^2) \cdots (1 + t + \cdots + t^{r-1})$, which is less than the coefficient of t^s in $(1 + t + t^2 + \cdots)^{r-1} = (1/(1-t))^{r-1}$.

But this coefficient in absolute value is less than 4^r, as may be seen by differentiation or using the Cauchy integral formula again. Since $s \leq r$, there is a constant satisfying the lemma. □

In earlier chapters, a lot of technical work was avoided by picking a simple real change of coordinates that transformed the vector fields in such a way that the new vector fields, called X and Y, had the property that the Y were pure partial derivatives and would thus commute with the partial derivatives on X, - i.e., so that the "derivative part" of X could pass through Y and onto the function to the right,

$$XYw = \text{coeff}(Yw'),$$

and the coefficients taken to the right if desired,

$$XYw = \text{coeff}(Yw') = (Yw') \circ \text{coeff},$$

or more specifically,

$$\begin{aligned} XYw &= \left\{ \frac{\partial}{\partial x} - y\frac{\partial}{\partial t} \right\} \frac{\partial}{\partial y} w = \frac{\partial}{\partial y}\frac{\partial}{\partial x} w - y \frac{\partial}{\partial y}\frac{\partial}{\partial t} w \\ &= Yw_x - yYw_t = Yw_x - (Yw_t) \circ y. \end{aligned}$$

In this sense we were able to "commute" Y past X and onto w, and this served our purposes.

This drastic simplification is not available to us in the complex case, unfortunately, since no such coordinate change is available.

A substitute exists, however, by splitting the $\tilde{N}_a$ and $\overline{N}_b$, for in so doing when the indices aren't balanced as needed for $(T^p)_\varphi$ there will be either extra M's or L's on the extreme left of φ or extra $\overline{M}$'s or $\overline{L}$'s to the extreme right - in between will be the balanced $(T^{\tilde{p}})_{\tilde{\varphi}}$:

Definition 9.1. For any $m \geq 0$ and any function $\varphi(z, t)$ let

$$(T^m)_\varphi = \sum_{a+b \leq m} \tilde{N}_a \circ ((\partial/\partial t)^{a+b} \varphi) \circ \overline{N}_b \, T^{m-(a+b)},$$

and

$$G_{A;\varphi} = L^I T^p (T^m)_{(\partial/\partial t)^r \varphi} T^s \overline{L}^q.$$

Then, much as in Chap. 4, with

$$
\begin{aligned}
A &= (m, p, q, r, I, q), \\
|A| &= m + p + q + r + |I| + q, \\
\|A\| &= |A| + p + q,
\end{aligned}
$$

(counting free T derivatives twice because they arise from two Z's), we recall the following definition

Definition 9.2. An expression such as $G_{A;\varphi}$ in the above proposition is *admissible* if $|I + q| > 0$, i.e., if it contains an L or an $\overline{L}$) and *simple* if it is inadmissible and $m = 0$ (i.e., if it contains no L or $\overline{L}$ and no $(T^m)_{\tilde{\varphi}}$).

Proposition 9.3. *Let* $G_{A;\varphi} = L^I T^p (T^m)_{(\partial/\partial t)^r \varphi} T^s \overline{L}^q$. *Then*

$$[G_{A;\varphi}, L_j \text{ or } \overline{L}_j] = -G_{A_{[0]};\varphi_{z_j \text{ or } \bar{z}_j}} + \tilde{\mathcal{G}}_{A_{[0]};\varphi}$$

where $G_{A;\varphi} = \sum_{|\rho| \geq 0} f_{(\rho)} G_{A_{(\rho)};\varphi}$ *with* $|D^\sigma f_{(\rho)}| \leq C^{|\rho+\sigma|} |\sigma|! \; \forall \alpha$,

$$|A_{(\rho)}| \leq |A| - |\rho|, \quad \|A_{(\rho)}\| \leq \|A\| + 1,$$

and

$$|||\tilde{\mathcal{G}}_{A_{[0]};\varphi}|||_N \leq |||L_j G_{A;\varphi}|||_N.$$

If $m \neq 0$, *and* $|I + q| = 0$ *(in the original A), then* $\tilde{\mathcal{G}}_{A_{[0]};\varphi}$ *is admissible.*

For brackets with functions we have, as in Chap. 4,

$$
\begin{aligned}
[(T^p)_\varphi, a] = \sum C_{p,\alpha_1,\beta_1,j} (y)^{\beta_2'' + \alpha_{12}} & \frac{(Y^{\alpha_1} X^{\beta_1} T^j a)}{\alpha_1! \beta_1! j!} \\
& \circ L^{\beta_2'} \left(T^{p-j-|\alpha_{12}+\beta_{12}|} \right)_{\varphi^{(\beta_2''+\beta_{12}+\alpha_{12})}} \overline{L}^{\alpha_2},
\end{aligned}
$$

and $C_{p,\alpha_1,\beta_1,j} \lesssim C^j p^{j-t}$,

$$= \sum_{\substack{t+j+|\gamma| \leq p \\ t = t_1 + t_2' + t_2''}} F_{(0)}^{(\alpha_1+\beta_1+j)} (y)^{t_2''+\gamma} \frac{L^{t_{2'}}}{p^{t_{2'}}} \frac{p^j}{p^{t_2''}} (T^{p-j-|\gamma|})_{\varphi^{(|\gamma|+t_2'')}} \frac{\overline{L}^{t_1}}{p^{t_1}} ZZu,$$

where each Z is either an L or an $\overline{L}$, and with the same norm results. From this point onward, the analysis continues as in Chap. 4, permitting us to conclude that the solution is real analytic when this is true of the right hand side.

Chapter 10
Details of the Previous Chapter

In this chapter we provide the details of the previous chapter, some of which may seem rather forbidding. The aim was to prove local real analytic hypoellipticity (at the origin) in both the $\bar{\partial}$-Neumann problem and for $\Box_b$, for pseudoconvex domains that are locally of the form $\Im w > h(|z|^2)$ with $h(z)$ real analytic (and not constant), $h(0) = 0$. These results were announced in [DT9] and presented in extended form in [DT1]. Again, we use purely L^2 methods generalizing those of Tartakoff [T4], [T5], and use results of Derridj and Grigis–Sjöstrand [D3], [GS] that show that one has "maximal" estimates for such domains. (It is worth noting that we make no use of subellipticity per se; a compactness estimate would suffice.) We also need to construct a special holomorphic vector field M that in a sense takes the place of the modifications of ψT cited above and used in [T4].

We must apologize for quoting here a few pages of the previous chapter. This chapter is in fact self-contained, but only by doing so, whereas the previous chapter was intended to give more of an outline, with much left to be filled in, specifically proofs.

10.1 Statement of the Theorems

Theorem 10.1. *Let V be a neighborhood of a point p_0 in a real analytic, pseudoconvex C-R manifold Γ that is locally given in $\mathbb{C}^n$ by $\Im w = h(|z^2|)$ with h real analytic and not identically constant. Let u solve $\Box_b u = \alpha$ in V with α real analytic in V. Then u is real analytic in V.*

Theorem 10.2. *In a neighborhood V of $p_0 \in \partial\Omega$, let Ω be pseudoconvex and given by $\Im w > h(|z|^2)$ with h real analytic and not identically constant. Then the $\bar{\partial}$-Neumann problem is analytic hypoelliptic up to the boundary near* 0; *that is, if u is a solution to the $\bar{\partial}$-Neumann problem and if $\alpha \in \mathcal{A}(\overline{\Omega} \cap V)$ then $u \in \mathcal{A}(\overline{\Omega} \cap V)$.*

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_10,

10.2 The Vector Field M and the Localization

The first proposition prepares the important vector field M, which will allow us to correct errors that arise when derivatives fall on the localizing functions.

Proposition 10.1. *There exists a real, real analytic function $\mu(s)$, nonzero at the origin, such that $M = \mu^{-1}\sum_{j<n} z_j L_j$ satisfies*

$$[L_j, \overline{M}] = -ig_{z_j}T + (L_j\mu^{-1})\mu\overline{M}. \tag{10.1}$$

Proof. The function $h(s)$ that occurs in the defining function $\Im w - h(|z|^2)$ is real analytic and hence satisfies a relation of the form

$$h'(s) + sh''(s) = (\mu(s)/2)h'(s), \quad \mu(0) \neq 0, \tag{10.2}$$

as may be seen by power series: if $h'(s) = \sum_m^\infty a_j s^j$ with $a_m \neq 0$, then $h'(s) + sh''(s) = \sum_m^\infty (1+j)a_j s^j = (\mu(s)/2)h'(s)$. With this choice of the function μ, since $L_j = \partial/\partial z_j + ig_{z_j}\partial/\partial t$ and $g(z,\overline{z}) = h(|z|^2), z = (z_1, \ldots, z_n)$, we have

$$\begin{aligned}
[L_j, \overline{M}] &= \left[L_j, \mu^{-1}\sum_{k<n}\overline{z}_k\overline{L}_k\right] = (L_j\mu^{-1})\mu\overline{M} + \mu^{-1}(|z|^2)\sum_{k<n}\overline{z}_k[L_j, \overline{L}_k] \\
&= (L_j\mu^{-1})\mu\overline{M} + \mu^{-1}\sum_{k<n}\overline{z}_k(-2i)\{(z_k\overline{z}_j h''(|z|^2) + \delta_{jk}h'(|z|^2)\}T \\
&= (L_j\mu^{-1})\mu\overline{M} - 2i\mu^{-1}\overline{z}_j\{|z|^2h''(|z|^2) + h'(|z|^2)\}T \\
&= (L_j\mu^{-1})\mu\overline{M} - 2i\mu^{-1}\overline{z}_j\{(\mu/2)h'(|z|^2)\}T \\
&= (L_j\mu^{-1})\mu\overline{M} - ig_{z_j}T.
\end{aligned}$$

□

Note that in particular, for any function ζ, in view of (10.1),

$$\begin{aligned}
[L_j, \zeta T + \zeta_t\overline{M}] &= L_j(\zeta)T - ig_{z_j}\zeta_t T + L_j(\zeta_t)\overline{M} + \zeta_t(L_j\mu^{-1})\mu\overline{M} \\
&= \zeta_{z_j}T + L_j(\zeta_t)\overline{M} + \zeta_t(L_j\mu^{-1})\mu\overline{M}.
\end{aligned}$$

We shall choose our cut-off functions so that ζ_{z_j} is zero near the degeneracies of the Levi form, so that where ζ_{z_j} is nonzero, the analyticity of the solution is known by [T4], [T5], and [Tr4], so that we shall be able to treat terms with ζ_{z_j} easily. This is the "first correction" of ζT as in [T4]. But the higher-order generalization is not straightforward: the "simple" sum

$$\sum_{a+b\le p} M^a(T^{a+b}\zeta)\overline{M}^b T^{p-(a+b)}/a!b!,$$

which was so useful in [T4], will not work here, because of the presence of non-trivial higher-order commutators. Still, there is a replacement for this expression, which we proceed to formulate.

Definition 10.1. Let

$$\tilde{N}_a = \sum_{0 \le a' \le a} \tilde{A}^a_{a'} (M^{a'}/a'!) \quad \text{and} \quad \overline{N}_b = \sum_{0 \le b' \le b} A^b_{b'} (\overline{M}^{b'}/b'!),$$

where

$$\tilde{A}^a_{a'} = (-2)^{a'-a} B^a_{a'} \quad \text{and} \quad A^b_{b'} = 2^{b'-b} B^b_{b'}$$

with

$$\begin{aligned} B^r_s &= \left(\left(\frac{t}{e^t - 1} \right)^{r+1} \right)^{(r-s)} (0)/(r-s)! \\ &= \text{the coefficient of } t^{r-s} \text{ in } \left(\frac{t}{e^t - 1} \right)^{r+1} \\ &= \sum_{\sum i_j = r-s} d_{i_1} d_{i_2} \cdots d_{i_{r+1}}, \end{aligned}$$

where

$$d_k = \left(\frac{t}{e^t - 1} \right)^{(k)} (0)/k!.$$

We also define

$$S^r_s = \sum_{\substack{\sum i_j = s \\ i_k \in \{0,1,\dots,k-1\}}} d_{i_1} d_{i_2} \cdots d_{i_{r+1}}.$$

Proposition 10.2. *The operators $\tilde{N}_a$ and $\overline{N}_b$ satisfy the following commutation relations: for $a, b \ge 1$,*

$$[L_k, \overline{N}_b] = -i g_{z_k} \overline{N}_{b-1} T - \sum_{j+b''=1}^{b} S^{b''+j}_j 2^{-j} \frac{e_k^{(b''-1)} \overline{M}}{b''!} \overline{N}_{b-(b''+j)}, \tag{10.3}$$

where

$$\begin{aligned} e_i^{(d-1)} &= \overline{M}^{d-1} ((L_j \mu^{-1}) \mu), \quad \mu = \mu(|z|^2), \\ \overline{L}_i, \overline{N}_b &= - \sum_{k+b''=1}^{b} S^{b''+k}_k 2^{-k} \overline{N}_{b-(b''+k)} \times \left(\frac{e_{i,1}^{(b''-1)} \overline{L}_i + e_{i,2}^{(b''-1)}}{b''!} \right), \end{aligned} \tag{10.4}$$

where

$$e_{i,1}^{(b''-1)}\overline{L}_i + e_{i,2}^{(b''-1)}\overline{M}^{d-1} = \mathrm{ad}_{\overline{M}}^{b''}(\overline{L}_i),$$

$$\overline{L}_i, \tilde{N}_b = \tilde{N}_{a-1} i g_{z_i} T + \sum_{j+a''=1}^{a} S_j^{a''+j}(-2)^{-j}\, \frac{e_{i,3}^{(a''-1)} M}{a''!}\, \tilde{N}_{a-(a''+j)}, \tag{10.5}$$

where

$$e_{i,3}^{(a''-1)} = M^{a''-1}((\overline{L}_i \mu^{-1})\mu), \qquad \mu = \mu(|z|^2),$$

and

$$[L_i, \tilde{N}_a] = \sum_{k+a''=1}^{a} \left(\frac{e_{i,4}^{(a''-1)} L_i + e_{i,5}^{(a''-1)} M}{a''!} \right) \times (-2)^{-k} S_k^{a''+k} \tilde{N}_{a-(a''+k)}, \tag{10.6}$$

where

$$e_{i,4}^{(a''-1)} L_i + e_{i,5}^{(a''-1)} M = -\,\widetilde{\mathrm{ad}}_{M}^{a''}(L_i)$$

and

$$\mathrm{ad}_A^c B = [[\ldots[B, A], A], \ldots, A], \qquad \widetilde{\mathrm{ad}}_A^c B = [A, [A, [\ldots, B]] \ldots].$$

Proof. First we show that $[L_k, \overline{N}_b] \equiv -i g_{z_k} \overline{N}_{b-1} T$ modulo terms with all $\overline{M}$'s. We have

$$\frac{[L_k, \overline{M}^c]}{c!} = -\sum_{c' \le c} \left(\frac{\mathrm{ad}_{\overline{M}}^{c'} L_k}{c'!} \right) \left(\frac{\overline{M}^{c-c'}}{(c-c')!} \right).$$

Now

$$[\overline{M}, L_k] = i g_{z_k} T + e_k \overline{M}, \quad \text{ where } e_k = (L_k \mu^{-1})\mu,$$

and as in the proof of Proposition 10.1,

$$[\overline{M}, i g_{z_k}] = i g_{z_k}/2.$$

Thus

$$[\overline{M}, [\overline{M}, L_k]] = \frac{1}{2} i g_{z_k} T + e_k^{(1)} \overline{M} e_k,$$

and in general

$$\mathrm{ad}_{\overline{M}}^{d}(L_k) = \frac{1}{2^{d-1}} i g_{z_k} T + e_k^{(d-1)} \overline{M},$$

where $e_k^{(d-1)} = \overline{M}^{d-1} e_k = \overline{M}^{d-1}(L_k \mu^{-1})\mu$. Thus

$$[L_k, \overline{N}_b] = \sum_{b' \le b} A^b_{b'} [L_k, \overline{M}^{b'}/b'!]$$
$$= -\sum_{b' \le b} A^b_{b'} \sum_{b''=1}^{b'} \frac{1}{b''!} \left(\frac{i}{2^{b''-1}} g_{z_k} T + e_k^{(b''-1)} \overline{M} \right) \frac{\overline{M}^{b'-b''}}{(b'-b'')!}. \quad (10.7)$$

Now in order to have

$$\sum_{b' \le b} A^b_{b'} \sum_{b''=1}^{b'} \frac{1}{b''!} \frac{-i}{2^{b''-1}} g_{z_k} T \frac{\overline{M}^{b'-b''}}{(b'-b'')!} = -i g_{z_k} \overline{N}_{b-1} T,$$

we must have

$$\sum_{b' \le b} A^b_{b'} \sum_{b''=1}^{b'} \frac{1}{b''!} \frac{i}{2^{b''-1}} \frac{\overline{M}^{b'-b''}}{(b'-b'')!} = \overline{N}_{b-1} = \sum_{b_1 \le b-1} A^{b-1}_{b_1} (\overline{M}^{b_1}/b_1!),$$

which in turn requires, equating coefficients of like powers of $\overline{M}$, that for each b_1 (summing over all pairs (b', b'') with $b' - b'' = b_1$),

$$\sum_{b''=1}^{b-b_1} A^b_{b_1+b''} \frac{1}{b''!} \frac{i}{2^{b''-1}} = A^{b-1}_{b_1}.$$

Things become a bit more transparent if we write $A^b_{b'} = 2^{b'-b} B^b_{b'}$. Then the requirement is that we have $B^q_q = 1$ for each q and, for each b and b_1,

$$\sum_{b''=1}^{b-b_1} B^b_{b_1+b''} \frac{1}{b''!} = B^{b-1}_{b_1}. \quad (10.8)$$

Lemma 10.1. *The unique solution to (10.8) with $B^q_q = 1$ for all q and $B^q_0 = (-1)^q$ for all q is given by*

$$B^r_s = \left(\left(\frac{t}{e^t - 1} \right)^{r+1} \right)^{(r-s)} (0)/(r-s)!$$
$$= \textit{the coefficient of } t^{r-s} \textit{ in } \left(\frac{t}{e^t - 1} \right)^{r+1}.$$

□

Proof. A straightforward expansion, using Leibniz's rule, of

$$\left(\left(\frac{t}{e^t - 1} \right)^{b-1} \left(\frac{e^t - 1}{t} \right) \right)^{(b-1-b_1)} (0)/(b-1-b_1)!$$

together with the fact that $\left(\frac{e^t-1}{t}\right)^{(s)}(0)/s! = 1/(s+1)!$ yields (3.8). Uniqueness is clear from the recurrence formula. We merely refer to Hirzebruch's book [Hi], Lemma 1.7.1, for the initial conditions $B_0^q = (-1)^q$, since we make no use of them.

For the second half of the first part of Proposition 10.2, we must consider

$$\sum_{b'\le b} A_{b'}^b \sum_{b''=1}^{b'} \frac{1}{b''!}\left(e_k^{(b''-1)}\overline{M}\right)\frac{\overline{M}^{b'-b''}}{(b'-b'')!}$$

and must use the proposition below to lower the lower index b' on A to make it agree with the factorials and powers of $\overline{M}$. Assuming this for the moment, we complete the proof of (10.3):

$$\begin{aligned}
&\sum_{b'\le b} A_{b'}^b \sum_{b''=1}^{b'} \frac{1}{b''!}\left(e_k^{(b''-1)}\overline{M}\right)\frac{\overline{M}^{b'-b''}}{(b'-b'')!}\\
&= \sum_{\substack{1\le b'\le b\\ j\le b-b'}} S_j^{b''+j} A_{b'-b''}^{b-(b''+j)} \sum_{b''=1}^{b'} \frac{1}{b''!}\left(e_k^{(b''-1)}\overline{M}\right)\frac{\overline{M}^{b'-b''}}{(b'-b'')!}\\
&= \sum_{j+b''=1}^{b} S_j^{b''+j}\left(e_k^{(b''-1)}\overline{M}\right)\overline{N}_{b-(b''+j)}.
\end{aligned}$$

The proof of (10.4) is simpler and follows the same lines. The proofs of the last two parts of the proposition are virtually the same. □

Proposition 10.3. *For any p, q, and c, we may shift lower indices:*

$$B_q^p = \sum_{j=0}^{p-q} S_k^{c+j} B_{q-c}^{p-(c+j)},$$

and so

$$A_q^p = \sum_{j=0}^{p-q} S_k^{c+j} B_{q-c}^{p-(c+j)}$$

and

$$\tilde{A}_q^p = \sum_{j=0}^{p-q} S_k^{c+j} \tilde{A}_{q-c}^{p-(c+j)}.$$

Proof. We recall the alternative definition $B_q^p = \sum d_{i_1} d_{i_2} \cdots d_{i_{p+1}}$, where the sum is over all $\{i_j\}$ with $\sum i_j = p - q$. Then it is easy to see that we may write

$$\begin{aligned} B_q^p &= d_0 B_{q-1}^{p-1} + d_1 B_q^{p-1} + d_2 B_{q+1}^{p-1} + \cdots + d_{p-q} B_{p-1}^{p-1} \\ &= d_0 \left(d_0 B_{q-2}^{p-2} + d_1 B_{q-1}^{p-2} + d_2 B_q^{p-2} + \cdots + d_{p-q} B_{p-2}^{p-2} \right) \\ &\quad + d_1 \left(d_0 B_{q-1}^{p-2} + d_1 B_q^{p-2} + d_2 B_{q+1}^{p-2} + \cdots + d_{p-q-1} B_{q-2}^{p-2} \right) \\ &\quad + \cdots + d_{p-q} d_0 B_{q-2}^{p-2} = \cdots . \end{aligned}$$

We continue to expand each B until it reaches $q' = q - c$. Since this implies that in each case the next-to-last step must have been of the form $B_{q-c+1}^{p'+1}$, the last d_j must have been zero. And so just before reaching this next-to-last step, such a term must have been either $B_{q-c+2}^{p'+2}$ or $B_{q-c+1}^{p'+2}$, since the lower index can decrease by at most one, and if the previous term had had lower index $q-c$, we would have stopped then. This means that the next-to-last d_j must have been either zero or one. Continuing in this fashion, and reversing the ordering of the d_j, we see that the restriction on the size of each d_j in the proposition must hold. Finally, it is clear from the successive reductions above of the *upper* index in the proof that at the jth step, the sum of the lower index plus the indices of all the d_j that go with a given term equals $q - j$. Thus when we stop reducing a given term because its lower index has reached $q-c$, the sum of the indices on the d_j that accompany it must be c. □

Remark 10.1. We may express the coefficient S_c^b as

$$S_c^b = \underbrace{G(t)\frac{\partial}{\partial t}G(t)\frac{\partial}{\partial t}\cdots G(t)\frac{\partial}{\partial t}}_{c} G(t)^{(b-c)}(0)/c!, \tag{10.9}$$

where

$$G(t) = \left(\frac{t}{e^t - 1}\right).$$

Proposition 10.4. *The results of commuting $\tilde{N}_a$ and $\overline{N}_b$ with functions are as follows:*

$$[\tilde{N}_a, f(z, \bar{z}, t)] = \sum_{j+k=1}^{a} \frac{M^k f}{k!}(-2)^{-j} S_j^{k+j} \tilde{N}_{a-(j+k)}$$

and

$$[\overline{N}_b, f(z, \bar{z}, t)] = \sum_{j+k=1}^{b} \frac{\overline{M}^k f}{k!} 2^{-j} S_j^{k+j} \overline{N}_{b-(j+k)}.$$

Proof. A direct application of Proposition 10.3 and Definition 10.1. □

Lemma 10.2. *There exists a constant C such that for $s \leq r$ we have*

$$|B_s^r| \leq C^r.$$

Proof. The Cauchy integral formula reads, in view of Definition 10.1,

$$B_s^r = \frac{1}{2\pi i}\int \frac{\left(\frac{\zeta}{e^\zeta - 1}\right)^{r-1}}{\zeta^{r-s+1}}\, d\zeta,$$

since $\zeta/(e^\zeta - 1)$ is holomorphic, where the integration is over a small contour containing the origin. On such a contour the integrand is bounded by C^r (recall that $r \geq s > 0$ in Definition 10.1). □

Lemma 10.3. *There exists a constant C such that for $s \leq r$,*

$$|S_s^r| \leq C^r.$$

Proof. By definition, S_s^r is the sum of all $d_{i_1} d_{i_2} \cdots d_{i_r}$ such that $\sum i_j = s$, and for each k, $i_k \in \{0, 1, \ldots, k-1\}$, where $d_j = (r/(e^t - 1))^{(j)}(0)/j!$ Clearly, $|d_j| \leq A^j$ for some universal constant A, so each product is bounded by A^s in absolute value. Thus S_s^r is bounded by A^s times the number of partitions of s into r nonnegative integers subject to the above conditions. But this number of partitions is the coefficient of t^s in $(1)(1+t)(1+t+t^2)\cdots(1+t+\cdots+t^{r-1})$, which is less than the coefficient of t^s in $(1+t+t^2+\cdots)^{r-1} = (1/(1-t))^{r-1}$. But this coefficient in absolute value is less than 4^r, as may be seen by differentiation or using the Cauchy integral formula again. Since $s \leq r$, there is a constant satisfying the lemma.

Definition 10.2. For any m, a, b, d and any function $\Psi(z, t)$, let

$$(T^m)_\Psi = \sum_{a+b \leq m} \tilde{N}_a \circ ((\partial/\partial t)^{a+b}\Psi) \circ \overline{N}_b\, T^{p-(a+b)}$$

and

$$(T^m)_{\Psi; X^d} = \sum_{a+b \leq m} \tilde{N}_a \circ ((\partial/\partial t)^{a+b}\Psi) \circ X^d \circ \overline{N}_b\, T^{p-(a+b)}.$$

Before we can prove lemmas concerning commutators of $(T^m)_{\Psi; X^d}$ with coefficients or vector fields, we need to prove some elementary combinatorial results, which we do next.

Lemma 10.4. *For any $a' \leq a \leq m$ and $k \leq m - a$,*

$$\binom{m-a}{k} = \sum (-1)^{j+s} \binom{m-s}{k-j-s}\binom{a'}{s}\binom{a-a'}{j},$$

where the sum is over $s \leq a', k-j$ and $j \leq a-a', k$.

Proof.

$$\binom{m-a}{k} = \sum_{j \le k, a-a'} (-1)^j \binom{a-a}{j}\binom{m-a+(a-a')}{k-j}$$
$$= \sum (-1)^j \binom{a-a'}{j} (-1)^s \binom{m-s}{k-j-s}\binom{a'}{s},$$

with the last sum over $s \le a', k-j$, and $j \le a-a', k$. □

Lemma 10.5. *For any $a'' \le a,\ b'' \le b$, and $k \le m-(a+b)$, we may write*

$$\binom{m-(a+b)}{k} = \sum_{S_1} R^{m;a;b;k}_{a'';a_2;b'';b_2} \binom{a''}{a_2}\binom{b''}{b_2},$$

with $S_1 = \{a_2 \le a'', b_2 \le b''\}$, where

$$R^{m;a;b;k}_{a'';a_2;b'';b_2} = \sum_{S_2} (-1)^{j_1+j_2+a_2+b_2} \binom{a-a''}{j_1}\binom{b-b''}{j_2}\binom{m-a_2-b_2}{k-j_1-j_2-a_2-b_2}$$

with $S_2 = \{j_1 \le a-a'', j_2 \le b-b'', j_1+j_2+a_2+b_2 \le k\}$ and

$$\left| R^{m;a;b;k}_{a'';a_2;b'';b_2} \right| \le 2^{a-a''+b-b''} \frac{m^{k-a_2-b_2}}{(k-a_2-b_2)!}.$$

Proof. Two applications of the previous lemma. □

The following lemmas are straightforward combinatorial results that use the definition of M above. We collect them here for later reference.

Lemma 10.6. *$M^{a'}$ is a sum of $C^{a'}$ terms of the form $f^{(a'')}_{(a')} X^{I_{a'-a''}}$ with $a'' \le a'$ and*

$$|D^\sigma f^{(a'')}_{(a')}| \le C^{|\sigma|+a'}(|\sigma|+a')! \qquad \forall \sigma.$$

Lemma 10.7. *$\tilde{N}_a$ is a sum of C^a terms of the form $\tilde{f}_{a';a''} X^{I_{a'-a''}}$ with $a'' \le a' \le a$ and*

$$|D^\sigma \tilde{f}_{a';a''}| \le C^{|\sigma|+a'}(|\sigma|+a'')!/a'! \qquad \forall \sigma.$$

Lemma 10.8. *$\tilde{N}_a \circ \theta \circ \overline{N}_b$ is a sum of C^{a+b} terms, each of the form $\tilde{f}_{a';a'';a''';b';b''} X^{I_{a'-a''-a'''}} \circ \theta \circ X^{q_{b'-b''}}$ with*

$$|D^\sigma \tilde{f}_{a';a'';a''';b';b''}| \le C^{|\sigma|+a'+b'}(|\sigma|+a''+a''')!/a'!b'! \qquad \forall \sigma.$$

Lemma 10.9. *For any a and b,*

$$ad_T^a ad_X^b(D_{tg}) = \sum f_j^{(a+b)} X_j + \sum f_{2n-1}^{(a+b)} T$$

where for all σ,

$$|D^\sigma f_j^{(a+b)}| \leq C^{|\sigma|+a+b}(|\sigma| + a + b)!$$

10.3 Behavior of the Localized Operators

To facilitate keeping track of errors and shorten the estimates, we follow some of the notation of [T5] and introduce some formal expressions and norms. We recall that

$$\begin{aligned} L_j &= \partial/\partial z_j + i g_{z_j} \partial/\partial t, \qquad j < n, \\ X_{2j-1} &= \Re L_j \text{ and } X_{2j} = \Im L_j, \qquad j < n, \\ X^I &= X_{i_1} X_{i_2} \cdots X_{I_{|I|}}, \qquad I = (i_1, \ldots, i_{|I|}). \end{aligned}$$

Now for any $A = (I, p, m, r, d, q, r)$ and any θ we set (cf. Definition 10.2)

$$G_{A,\theta} = C_A X^I T^p (T^m)_{T^r\theta; X^d} T^s X^q, \tag{10.10}$$

and define the following norms:

$$|A| = |I| + p + m + d + q + s = |G_{A;\theta}| \tag{10.11}$$

and

$$||A|| = |A| + p + s = ||G_{A;\theta}||, \tag{10.12}$$

the orders of G with T given single and double weight, respectively.

We also assign a norm to (variable-coefficient) sums of the $G_{A;\theta}$ as follows: N being given, if

$$\mathcal{G} = \sum F_A G_{A;\theta} \quad \text{with } |A| \leq N, \tag{10.13}$$

we set

$$|||\mathcal{G}|||_N = \sup_{A;\,\text{supp}\,\theta} |F_A| C_A K_0^m K_1^{|I|+d+q} K_2^{p+q} N^{|A|+r+m}/m!, \tag{10.14}$$

where K_0, K_1, K_2 will be chosen later, large relative to other constants that enter and subject to

$$K_0 \gg K_1 \gg K_2 \gg 1. \tag{10.15}$$

We shall always assume (10.15) to be satisfied without further mention.

We shall use the notation $F_{(k)}^{(j)}$ for a function or sum of functions such that in the support of all localizing functions, and for all σ,

$$|D^\sigma F_{(k)}^{(j)}| \le C^{|\sigma|+j+k+1}(|\sigma|+k)!, \tag{10.16}$$

with the constant C universal for the present problem. For example, $F_{(k)}^{(j)}$ could stand for any $(j+k)$th derivative of any coefficient divided by $j!$ Likewise, we might denote the function $e_k^{(b''-1)}$ appearing in (10.7) by $F_{(0)}^{(b''-1)} = F^{(b''-1)}$.

Definition 10.3. An operator $G_{A;\theta}$ will be called *admissible* if it contains an X (and at most one $\overline{L}_n$ in the case of Theorem 8.2). $G_{A;\theta}$ will be called *simple* if it is not admissible and $m=0$. A variable-coefficient sum of such operators will be said to have one of these properties if it is true of each member of the sum.

We may now state and prove the propositions that show the effect of commuting $(T^m)_\Psi$, or more generally $G_{A;\theta}$, with a function or a (tangential) vector field.

Proposition 10.5. *Let $G_{A;\Psi} = X^I T^p (T^m)_{(\partial/\partial t)^r\Psi;X^d} T^s X^q$, and let $f(z,\bar{z},t)$ be a real analytic function in the support of Ψ (which will only be smooth and generally of compact support).Then*

$$G_{A;\Psi,}f = \sum_{|\rho|\ge 0} f_{(\rho)} G_{A(-\rho);\Psi,} = \tilde{\mathcal{G}}_{A;\Psi}$$

with

$$f_{(0)} = f, A_{(0)} = A, |D^\sigma f_{(\rho)}| \le C^{|\sigma+\rho|+1}|\sigma|! \quad \forall\sigma$$

and

$$|A_{(-\rho)}| \le |A| - |\rho|, \quad \|A_{(-\rho)}\| \le \|A\| - |\rho|,$$

and

$$|||\tilde{\mathcal{G}}_{A;\Psi}|||_N \le C_f|||G_{A;\Psi}|||_N.$$

Proof. The T derivatives of f present the greatest difficulty. We have

$$T^{m-(a+b)}f = \sum_{k=0}^{m-(a+b)} \binom{m-(a+b)}{k} f^{(k)}T^{m-k-(a+b)},$$

where $f^{(k)}$ will denote any k-fold derivative of f, each derivative taken from a fixed, finite set (the X's, M, $\overline{M}$, or T). Next,

$$(\overline{M}^{b'}/b'!)f^{(k)} = \sum_{b''\le b'} \frac{f^{(k+b'-b'')}}{(b'-b'')!}(\overline{M}^{b''}/b''!),$$

so that

$$(T^m)_\Psi f = \sum_S \frac{f^{(k+b'-b''+a'-a'')}}{(b'-b'')!(a'-a'')!}\binom{m-(a+b)}{k}\tilde{A}^a_{a'}(M^{a''}/a''!)$$
$$\times((\partial/\partial t)^{a+b}\Psi)\sum_{b'\le b} A^b_{b'}(\overline{M}^{b''}/b''!)T^{m-k-(a+b)},$$

$S = \{a'' \le a' \le a, b'' \le b' \le b, k \ge 1\}$. Next we write $\binom{m-(a+b)}{k}$ in terms of $\binom{a''}{a'''}\binom{b''}{b'''}$:

$$\binom{m-(a+b)}{k} = \sum_{S'} R^{m;a;b;k}_{a'';a''-a''';b'';b''-b'''}\binom{a''}{a'''}\binom{b''}{b'''},$$

$S' = \{a'' \le a; b'' \le b; a''' \le a'', b''' \le b''\}$ by (10.5). Then we need to express $\tilde{A}^a_{a'}$ in terms of $A^{\hat{a}}_{a'''}$ for some $\hat{a}$ and the same for $A^b_{b'}$. This is done by Proposition 10.3. The result is that

$$\binom{m-(a+b)}{k}\tilde{A}^a_{a'}A^b_{b'}/a''!b''!$$
$$= \sum_{S''} R^{m;a;b;k}_{a'';a''-a''';b'';b''-b'''}\frac{2^{-k_2}(-2)^{-k_3}}{(a''-a''')!(b''-b''')!}S^{a'-a'''+k_2}_{k_2}$$
$$\times S^{b'-b'''+k_3}_{k_3}\tilde{A}^{a-(a'-a'''+k_2)}_{a'''}A^{b-(b'-b'''+k_3)}_{b'''}/a'''!b'''!,$$

$S'' = \{a''' \le a''; b''' \le b'', a'''+b''' \le k, k_2 \le a-a', k_3 \le b-b'\}$. Thus,

$$(T^m)_\Psi f = \sum_{S'''} \frac{f^{(k+a'-a''+b'-b'')}}{(a''-a''')!(b''-b''')!}KM^{a''-a'''}$$
$$\times(T^{m-k-a'-a'''-k_2-b'-b'''-k_3})_{(\partial/\partial t)^{a-a'''+k_2+b'-b'''+k_3}\Psi;\overline{M}^{b'-b'''}},$$

where $S''' = \{a''' \le a'' \le a' \le a, b''' \le b'' \le b' \le b, k+a+b \le m, k_2 \le a-a', k_3 \le b-b'\}$, and

$$K = R^{m;a;b;k}_{a'';a''-a''';b'';b''-b'''}\frac{2^{-k_2}(-2)^{-k_3}}{(a''-a''')!(b''-b''')!}S^{a'-a'''+k_2}_{k_2}S^{b'-b'''+k_3}_{k_3}.$$

Things are clearer if we use new indices $a_0 = a-a'$, $a_1 = a'-a''$, $a_2 = a''-a'''$, $a_3 = a'''$, $b_0 = b-b'$, $b_1 = b'-b''$, $b_2 = b''-b'''$, $b_3 = b'''$. Then we have

$$(T^m)_\Psi f = \sum_{S''''} \frac{f^{(k+a_1+b_1)}}{a_1!b_1!}$$
$$\times K'M^{a_2}(T^{m-k-a_1-a_2-k_2-b_1-b_2-k_3})_{(\partial/\partial t)^{a_1+a_2+k_2+b_1+b_2+k_3}\Psi;\overline{M}^{b_2}},$$

where $S'''' = \{k + a_0 + a_1 + a_2 + b_0 + b_1 + b_2 \leq m, k_2 \leq a_0, k_3 \leq_0\}$,

$$K' = R^{m;a;b;k}_{a'';a_2;b'';b_2} \frac{2^{-k_2}(-2)^{-k_3}}{a_2!b_2!} S^{a_1+a_2+k_2}_{k_2} S^{b_1+b_2+k_3}_{k_3},$$

and M^{a_2} is of course a sum of C^{a_2} terms, each of the form $\tilde{f}^{(a')}_{a_2} X^{I_{a_2-a'}}$ for some $a' \leq a_2$ and

$$|D^\sigma \tilde{f}^{(a')}_{a_2}| \leq C^{a_2+|\sigma|+1}(a' + |\sigma|)! \quad \forall\sigma.$$

Thus

$$(T^m)_\Psi f = \sum_{\tilde{S}} \frac{f^{(k+c_1)} \tilde{f}^{(a')}_{a_2}}{c_1!} K'' M^{a_2}$$
$$\times (T^{m-k-a_2-b_2-c_1-c_2})_{(\partial/\partial t)^{c_1+c_2+a_2+b_2}\Psi;\overline{M}^{b_2}},$$

where $\tilde{S} = \{k + c_1 + c_2 + a_2 + b_2 \leq m, a' \leq a_2\}$ and

$$K'' = \frac{C^{k+c_1+c_2+a_2+b_2}}{a'!(a_2 - a')!b_2!}.$$

If the constants K_j are well chosen and satisfy (10.15), and we let $(T^m)_\Psi f$ stand for the whole sum above, then uniformly in $m \leq N$ we have

$$|||(T^m)_\Psi f|||_N \leq C|||(T^m)_\Psi|||_N.$$

When the entire expression $G_{A;\Psi} = X^I T^p (T^m)_{(\partial/\partial t)^r\Psi;X^d} T^s X^q$ is considered instead of just $(T^m)_\psi$, the conclusion is unchanged and the only difference in the proof consists in introducing expressions such as

$$T^p f = \sum \binom{|p|}{|p'|} f^{(p')},$$

summed over all $p' \leq p$, etc.

Proposition 10.6. *Let* $G_{A;\Psi} = X^I T^p (T^m)_{(\partial/\partial t)^r\Psi;X^d} T^s X^q$. *Then*

$$[G_{A;\Psi}, L_j \text{ } or \text{ } \overline{L}_j] = -G_{A_{(0)};\Psi_{z_j} \text{ } or \text{ } \bar{z}_j} + \tilde{\mathcal{G}}_{A_{(0)};\Psi}$$

where $G_{A;\Psi} = \sum_{|\rho|\geq 0} f_{(\rho)} G_{A_{(\rho)};\Psi}$, *with y* $|D^\sigma f_{(\rho)}| \leq C^{|\rho+\sigma|}|\sigma|!\ \forall\alpha$ *and*

$$|A_{(\rho)}| \leq |A| - |\rho|, \quad \|A_{(\rho)}\| \leq \|A\| + 1,$$

and

$$|||\tilde{\mathcal{G}}_{A_{(0)};\Psi}|||_N \le |||L_j G_{A;\Psi}|||_N .$$

If $m \neq 0$ *and* $|I+q|+d=0$ *(in A), then* $\tilde{\mathcal{G}}_{A_{(0)};\Psi}$ *is admissible.*

Proof. First we consider the case $G_{A;\Psi} = (T^m)_\Psi$:

$$\begin{aligned}
&\sum_{a+b\le m} \left[\tilde{N}_a((\partial/\partial t)^{(a+b)}\Psi)\overline{N}_b T^{m-(a+b)}, L_j\right] \\
&\quad = \sum_{a+b\le m} \left[\tilde{N}_a, L_j\right]((\partial/\partial t)^{(a+b)}\Psi)\overline{N}_b T^{m-(a+b)} \\
&\qquad + \sum_{a+b\le m} \tilde{N}_a\left[((\partial/\partial t)^{(a+b)}\Psi)\overline{N}_b, L_j\right] T^{m-(a+b)} \\
&\qquad + \sum_{a+b\le m} \tilde{N}_a((\partial/\partial t)^{(a+b)}\Psi)\overline{N}_b \sum_{k_1\ge 1}\binom{m-(a+b)}{k_1} ad_T^{k_1}(L_j)T^{m-k_1-(a+b)} \\
&\quad = \sum_{\substack{a-k_1+b\le m-k_1\\ k_1\ge 1}} F_{(k_1)}(L \text{ or } M)\tilde{N}_{a-k_1}((\partial/\partial t)^{a-k_1+b}(\partial/\partial t)^{k_1}\Psi)\overline{N}_b T^{m-(a+b)} \\
&\qquad - \sum_{a+b\le m} \tilde{N}_a((\partial/\partial t)^{a+b}\Psi_{z_j})\overline{N}_b T^{m-(a+b)} \\
&\qquad + \sum_{a+b=m} \tilde{N}_a((\partial/\partial t)^{m+1}\Psi)\overline{N}_b + E_{L_j} \\
&\qquad + \sum_{\substack{a-k_1+b-k_2\\ \le m-d_1-k_2\\ k_2\ge 1}} F_{(k_1+k_2)}\tilde{N}_{a-k_1}((\partial/\partial t)^{a-k_1+b-k_2}(\partial/\partial t)^{k_1+k_2}\Psi)\overline{N}_{b-k_2} X T^{m-(a+b)}
\end{aligned}$$

from (10.6), Definition 10.2, and Proposition 10.5, where $E_{L_j} = \sum_{a_b\le m} E_{L_j;a,b}$ with

$$\begin{aligned}
E_{L_j;a,b} &= \sum_{\substack{a'\le a\\ b'\le b}} \tilde{A}^a_{a'}\frac{M^{a'}}{a'!}((\partial/\partial t)^{a+b}\Psi)A^b_{b'}\frac{\overline{M}^{b'}}{b'!} \\
&\qquad \times \sum_{k_1\ge 1}\binom{m-(a+b)}{k_1} F^{(k_1)}(L \text{ or } \overline{L})T^{m-k_1-(a+b)} \\
&\quad = \sum_S F^{(k_1)}_{a'-a''+b'-b''}\tilde{A}^a_{a'}\frac{M^{a''}}{a''!}\Psi^{(a+b)} \\
&\qquad \times A^b_{b'}\frac{\overline{M}^{b''}}{b''!}\binom{m-(a+b)}{k_1}(L \text{ or } \overline{L})T^{m-k_1-(a+b)}
\end{aligned}$$

$$\times \big(S = \{a'' \le a' \le a; b'' \le b' \le b; k_1 \ge 1\}\big)$$

$$= \sum_{S'} F^{(k_1)}_{a'-a''+b'-b''} R^{m;a;b;k_1}_{a'';a''-a''';b'';b''-b'''} S^{a'-a'''+k_2}_{k_2} S^{b'-b'''+k_3}_{k_3} 2^{-k_2}(-2)^{-k_3}$$

$$\times \frac{M^{a''-a'''}}{(a''-a''')!} \tilde{A}^{a-(a'-a'''+k_2)}_{a'''} \frac{M^{a'''}}{a'''!} \Psi^{(a+b)} A^{b-(b'-b'''+k_3)}_{b'''} \frac{\overline{M}^{b'''}}{b'''!}$$

$$\times \frac{\overline{M}^{b''-b'''}}{(b''-b''')!} \binom{m-(a+b)}{k_1} (L \text{ or } \overline{L}) T^{m-k_1-(a+b)}$$

$$\big(S' = \{a''' \le a'' \le a' \le a; b''' \le b'' \le b'vb; k_2 \le a-a';$$

$$k_3 \le b-b'; k_1 \le 1\}\big)$$

$$= \sum_{S''} F^{(k_1)}_{(a_1+a_2)} L M^{a_2} (T^{m-k_1-a_1-a_2-k_2-b_1-b_2-k_3})_{\Psi^{(a_1+a_2+k_2+b_1+b_2+k_3)}}$$

$$\times \overline{M}^{b_2} (L \text{ or } \overline{L}),$$

where

$$S'' = \{a_1 + a_2 + k_2 \le a, b_1 + b_2 + k_3 \le b, k_1 \ge 1\}$$

and

$$K = R^{m;a;b;k_1}_{a'';a_2;b'';b_2} S^{a_1+a_2+k_2}_{k_2} S^{b_1+b_2+k_3}_{k_3} 2^{-k_2}(-2)^{-k_3}/a_2!b_2!.$$

□

It is convenient to replace $\overline{M}^{b_2}$ and the L's or $\overline{L}$'s by X's, using Lemma 10.6 and the observation that commuting the resulting coefficients to the left does not introduce new $\overline{M}^{b_2}$ on the right, since the coefficients are already to the left of the free T's in $(T^{m'})_{\Psi'}$ (cf. the proof of Proposition 10.5).

Thus $[(T^m)_\Psi, L_j \text{ or } \overline{L}_j] = -(T^m)_{\Psi_{z_j} \text{ or } \bar{z}_j} + \tilde{\mathcal{G}}_{A;\Psi}$, where

$$\tilde{\mathcal{G}}_{A;\Psi} = \sum_{|\gamma| \ge 0} F_\gamma G_{A(-\gamma);\Psi} \quad \text{with} \quad |D^\sigma F_\delta| \le C^{|\delta+\sigma|+1}|\sigma|!$$

and

$$|A_{(-\gamma)}| \le |A| - |\gamma|, \qquad \|A_{(-\gamma)}\| \le \|A\| + 1,$$

$$|||\tilde{\mathcal{G}}_{A;\Psi}|||_N \le |||L_j (T^m)_\Psi|||_N.$$

Note that every term in $\tilde{\mathcal{G}}_{(0)}$ is admissible (if $m \ne 0$). When more general $G_{A;\Psi} = X^I T^p (T^m)_{\Psi^{(r)};X^d} T^s X^q$ are considered, bracketing with $L_j u$ or $\overline{L}_j$ may also yield as many as d terms of the form $X^I T^p (T^m)_{\Psi^{(r)};\text{coeff}X^{d-1}} T^s X^q$. The coefficients may be commuted to the left using Proposition 10.5 again, while the new T increases q by 1. In this case the norm $|\cdot|$ still drops, while the norm $\|\cdot\|$ may stay the same. Note that if $d = 0$, this does not happen.

Proposition 10.7. *Let $H_{A;\Psi} = X^I T^p (T^m)_{\Psi^{(r)};X^d} T^s X^q$ be admissible. Then*

$$H_{A;\Psi} = XG_{A_{(-1)};\Psi} + G_{A_{(-1)};\Psi_x} + \tilde{\mathcal{G}}^{sim}_{A_{(-1)};\Psi},$$

where

$$\tilde{\mathcal{G}}^{sim}_{A_{(-1)};\Psi} = \sum_{|\rho|\geq 1} f_{(\rho)} G^{sim}_{A_{(-1)};\Psi}$$

is simple,

$$|D^\sigma F_{(\rho)}| \leq C^{\sigma+|\rho|}|\sigma|!, \quad |A_{(-\rho)}| \leq |A| - |\rho|, \qquad \|A_{(-\rho)}\| \leq \|A\| + 1,$$

and

$$|||\tilde{\mathcal{G}}^{sim}_{A_{(-1)};\Psi}|||_N \leq |||XG_{A_{(-1)};\Psi}|||_N \quad \textit{and} \quad |||G_{A_{(-1)};\Psi_x}|||_N \leq |||H_{A;\Psi}|||_N.$$

Proof. The proof of Proposition 10.6 applies except that it is never necessary to bracket one X with another; it suffices to bring any X to the left. As long as $m \neq 0$, after application of the proof of Proposition 10.6, all terms except those in $G_{A_{(-1)};\Psi_x}$ retain an X. When m drops to zero, further iterations of the proof of Proposition 10.6 may contribute simple terms. □

10.4 Proof of Theorem 10.1

Proposition 10.8. *Let $H_{A;\Psi}$ be admissible and let u solve $\Box_b u = \alpha$ in V. Then*

$$\|H_{A;\Psi}u\|^2_{L^2} \leq C\left(\|G_{A_{(-1)};\Psi}\alpha\|^2_{L^2} + \sup \|G_{A'_{(-1)};\Psi}u\|^2_{L^2}\right),$$

where the supremum is over all $G_{A'_{(-1)};\Psi}$ whose norms satisfy

$$|||G_{A'_{(-1)};\Psi}|||_N \leq |||H_{A;\Psi}|||_N, \quad |A_{(-1)}| \leq |A| - 1, \quad \|A_{(-1)}\| \leq \|A\|.$$

Proof. We write $H_{(A;\Psi)} = XG_{A_{(-1)};\Psi)} + G'_{A_{(-1)};\Psi)}$ from (10.7). Then a priori estimate (9.1), applied to $G_{A_{(-1)};\Psi)}$, yields a right-hand side

$$Q(G_{A_{(-1)};\Psi})u, G_{A_{(-1)};\Psi}\big) u = Q(u, (G_{A_{(-1)};\Psi})^* G_{A_{(-1)};\Psi}u) + \sum_1^4 E_j$$

$$= (G_{A_{(-1)};\Psi}\alpha, G_{A_{(-1)};\Psi}u)_{L^2} + \sum_1^4 E_j,$$

where the E_j are the error terms

$$E_1 = ([\overline{\partial}_b, G_{A_{(-1)};\Psi}]u, \overline{\partial}_b G_{A_{(-1)};\Psi}u)_{L^2},$$

$$E_2 = \overline{\partial}_b u, [\overline{\partial}_b, (G^*_{A_{(-1)};\Psi}]G_{A_{(-1)};\Psi}u)_{L^2},$$

$$E_3 = ([\overline{\partial}^*_b, G_{A_{(-1)};\Psi}]u, \overline{\partial}^*_b G_{A_{(-1)};\Psi}u)_{L^2},$$

$$E_2 = (\overline{\partial}^*_b u, [\overline{\partial}^*_b, (G_{A_{(-1)};\Psi})^*]G_{A_{(-1)};\Psi}u)_{L^2}.$$

Now it suits us to write $\overline{\partial}_b$ as the sum of two terms of the form coeff$X$. If we do this, then the first and third terms above and the second and fourth are the same. Thus the errors that must be majorized are of the two types:

$$F_1 = \|[\,\mathrm{coeff}X, G_{A_{(-1)};\Psi}]u\|^2_{L^2}$$

and

$$F_2 = |([\mathrm{coeff}X, G_{A_{(-1)};\Psi}]\ \mathrm{coeff}X u, G_{A_{(-1)};\Psi}u)_{L^2}|.$$

These are handled by Propositions 10.5 and 10.14 above. Using these propositions, we have at once

$$|F_1| \leq \sup \|G'_{A_{(-1)};\Psi}u\|^2_{L^2},$$

with the supremum over all $G'_{A_{(-1)};\Psi}$ as in the statement of Proposition 10.8. For F_2 we need to take a bit more time, since Proposition 10.7 must be used. But

$$\begin{aligned}
&([\mathrm{coeff}X, G_{A_{(-1)};\Psi}]\ \mathrm{coeff}X u, G_{A_{(-1)};\Psi}u)_{L^2}\\
&\quad = (G'_{A_{(-1)};\Psi}\ \mathrm{coeff}X u, G_{A_{(-1)};\Psi}u)_{L^2}\\
&\quad = (\mathrm{coeff}X G_{A'_{(-1)};\Psi}u, G_{A_{(-1)};\Psi}u)_{L^2} + (G_{A''_{(-1)};\Psi}u, G_{A_{(-1)};\Psi}u)_{L^2}\\
&\quad = (G_{A''_{(-1)};\Psi}u, G_{A_{(-1)};\Psi}u)_{L^2} + (G_{A''_{(-1)};\Psi}u, XG_{A_{(-1)};\Psi}u)_{L^2},
\end{aligned}$$

and a weighted Schwarz inequality yields the result at once and proves the proposition. □

We may iterate this procedure until one of the resulting terms is no longer admissible. If $m = 0$ for an inadmissible term, we do not process it further at this point. On the other hand, when $m \neq 0$ in an inadmissible term, such a term is necessarily of the form

$$G'_{A_{(-k)};\Psi} = T^p(T^m)_{(\partial/\partial t)^r\Psi}T^s$$

(or $G'_{A_{(-k)};\Psi} = T^p(T^m)_{(\partial/\partial t)^r\Psi_x}T^s$, a type that we shall deal with later) with $|||G'_{A_{(-k)};\Psi}|||_N \leq |||H_{A;\Psi}|||_N, |A_{(-k)}| \leq |A| - k, \|A_{(-k)}\| \leq \|A\|$, and it is not immediately clear how to proceed, since to use (9.1) again with maximal advantage we would have to adjoin an X to the left of $T^p(T^m)_{(\partial/\partial t)^r\Psi_x}T^s$, a process that would increase all three norms. Now in fact we shall do just that, but we must show that:

- in arriving at $T^p(T^m)_{(\partial/\partial t)^r\Psi_x}T^s$, we gain a *factor* of N in $|||\cdot|||_N$, and
- after adjoining an X to $T^p(T^m)_{(\partial/\partial t)^r\Psi_x}T^s$, the next iteration of (9.1) drops $|\cdot|$ *and* $\|\cdot\|$ again, so that after these two applications of (9.1) the $|||\cdot|||_N$ norm is unchanged, $|\cdot|$ has dropped by one, and $\|\cdot\|$ is once again no greater than at the start.

For the first of these, there are two ways in which all free X's are lost. The first is when, in Proposition 10.14, two X's are bracketed to generate a T. In general, the form of this bracket is $[X^I, X]$ with $|I| \le N$, i.e., at most $|I|$ terms with a T and at most $|I|-1$ remaining X's *but at least one if* $|I| \ge 2$. Hence if such a bracketing uses up all remaining X's, $|I|$ had to be 1, and the result of the operation is to *decrease* $|||\cdot|||_N$ by a *factor of* N. And the second way that *all* free X's or $\overline{L}_n$ might be lost is as derivatives of a coefficient, and in this case we merely need to shift the count of derivatives by one: after all, the term in $[X^I, f] = \sum \binom{|I|}{|I'|} f^{(|I'|)} X^{I-I'}$ with $I = I'$ has coefficient $f^{(I)}$, and $|f^{(|I|)}| \le C^{|I|}|I|! \le C'^{|I|-1}(|I|-1)! \le C'^{|I|-1}N^{|I|-1}$.

Thus, modulo a constant for each derivative that falls on a function (which is always permitted by taking the K_i to have the proper ratios in (10.15)), the loss of the last X in this fashion does not need to result in a corresponding factor of N in the norm. For the second bulleted point above, we merely note that in the application of (9.1), i.e., from Propositions 10.5 and 10.14 when $|I|+q+d = 0$, not only does $|\cdot|$ drop, as it always does, but also $\|\cdot\|$ will drop. That is, two X's never bracket to generate a T, since there is no second X. And Proposition 10.7 will not generate T either.

Proposition 10.9. *Let $H_{A;\Psi}$ be admissible and let u solve $\Box_b u = \alpha$ in V. Then*

$$\|H_{A;\Psi}u\|^2_{L^2} \le C^N \left(\|G_{A_{(-1)};\Psi}\alpha\|^2_{L^2} + \sup \|G^{sim}_{A';\Psi}u\|^2_{L^2} + \sup \|G_{A'';\Psi_x}u\|^2_{L^2}\right),$$

where the first supremum is over all simple $G^{sim}_{A';\Psi}$ whose norms satisfy

$$|||G^{sim}_{A';\Psi}|||_N \le |||H^{sim}_{A;\Psi}|||_N, \quad |A'| \le |A|-1, \quad \textit{and } \|A'\| \le \|A\|,$$

and where the second supremum is over all $G_{A'';\Psi_x}$ whose norms satisfy the same bounds:

$$|||G_{A'';\Psi_x}|||_N \le |||H^{sim}_{A;\Psi}|||_N, \quad |A''| \le |A|-1, \quad \textit{and } \|A''\| \le \|A\|.$$

Note that since the singularities of the Levi form are contained in $\{x = 0\}$ $(= \{z = 0\})$, we may take Ψ to be a product $\Psi = \Psi_1(x)\Psi_2(t)$ with $\Psi_1(x) \equiv 1$ in a neighborhood of 0; thus since u is known to be real analytic in the complement of the set where the Levi form degenerates (cf. [T4], [Tr4]), we have

$$|D^\sigma u| \le C_u^{|\sigma|+1}(|\sigma|!) \quad \text{in supp} \quad \Psi_x.$$

For $G_{A'';\Psi_x}$ as above, then, we have the estimate

$$\|G_{A'';\Psi_x} u\|_{L^2}^2 \le \left(\sup \|G_{A''';\Psi_x} u\|_{L^2}^2\right) |||H_{A;\Psi}|||_N,$$

where the supremum is over all $G_{A'';\Psi_x}$ with $|A'''| \le |A| - 1, \|A'''\| \le \|A\|$ but $|||G_{A''';\Psi_x}|||_N \le 1$. Since for real analytic functions w satisfying (10.4) we have

$$\|G_{B;\Psi} w\|_{L^2}^2 \le C_w |||G_{B;\Psi}|||_N^2 \qquad \text{under (10.15)},$$

the conclusion of Proposition 10.9 reads

$$\begin{aligned}\frac{\|H_{A,\Psi} u\|_{L^2}}{|||H_{A,\Psi}|||_N} &\le C_u C^N \left(1 + \sup_S \frac{\|T^{p'} \circ \Psi^{(r')} \circ T^{q'} u\|_{L^2}}{N^{p'+q+r'}}\right) \\ &\le C_u C^N \left(1 + \sup_{|r'| \le N+1} \frac{|\Psi^{(r')}|}{N^{(r')}} \sup_{S'} \frac{\|T^{m'} u\|_{L^2(supp\,\Psi)}}{N^{m'}}\right), \\ S = \{2(p'+q') &\le \|A\|\}, S' = \{m' \le \|A\|/2\|.\end{aligned}$$

That is, if $|I| + p \le N$, then

$$\frac{\|X^I T^p u\|_{L^2_{(\Psi=1)}}}{N^{|I|+p}} \le C^N \left(1 + \sup_{\substack{|r| \le N+1 \\ 2m \le N}} \frac{|\Psi^{(r)}|}{N^{(r')}} \frac{\|T^m u\|_{L^2(supp\,\Psi)}}{N^m}\right). \tag{10.17}$$

Lastly, we control the localizing functions as in [T4] and [T5]:

Proposition 10.10. *There exists a constant K such that if Ω_1, Ω_2 are open sets in $\mathbb{C}^n$ with distance d from Ω_1 to the complement of Ω_2, then for any N there exists $\Psi = \Psi_N$ in $C_0^\infty(\Omega_2)$ equal to one in a neighborhood of Ω_1 with*

$$|\Psi^{(\beta)}| \le K K^{|\beta|} d^{-|\beta|} N^{|\beta|} \qquad \textit{for } |\beta| \le 2N. \tag{10.18}$$

The first use of these functions seems to be due to Ehrenpreis. Using such Ψ, (10.19) becomes

$$\max_{p+|I| \le N} \frac{\|X^I T^p u\|_{L^2(\Omega_1)}}{N^{|I|+p}} \le C^N \left(1 + d^{-2N} \sup_{q+s \le N/2} \frac{\|X^q T^s u\|_{L^2(\Omega_2)}}{(N/2)^{q+q}}\right). \tag{10.19}$$

We iterate *this* estimate by nesting $\log_2 N$ such pairwise relatively compact open sets $\Omega_1 \Subset \Omega_2 \cdots \Subset \Omega_{\log_2 N}$ with separations $d_j = d_0/2^j$. The iterates of (10.19) then yield

$$\max_{p+|I|\leq N} \frac{\|X^I T^p u\|_{L^2(\Omega_1)}}{N^{|I|+p}} \leq C'^N \Pi \left(2^j\right)^{N/2^{j-2}} C_u$$

with constants independent of N. This implies analyticity in Ω_1. Since Ω_1 was an arbitrary compact subset of V, this proves Theorem 10.1.

10.5 Proof of Theorem 10.2

We recall and rewrite the $\overline{\partial}$-Neumann boundary condition: if we split an arbitrary form $w = \sum w_{IJ} \lambda^I \overline{\lambda}^q$,

$$w = w' + w'',$$

where w' is the sum of all terms $\sum w_{IJ} \lambda^I \overline{\lambda}^q$ with $n \notin J$, while w'' consists of those with $n \in J$, and write

$$\widetilde{\overline{L}}_n w = (\overline{L} w' + Aw)$$

for some analytic matrix A, then we may rewrite the boundary conditions as

$$u'' = 0 \text{ on } \Gamma \quad \text{ and } \quad \widetilde{\overline{L}}_n u' = 0 \text{ on } \Gamma. \tag{10.20}$$

We modify the notation for our scalar operators $G_{A;\Psi}$ slightly from that given in (10.10) for the boundary value problem, allowing at most one $\overline{L}_n$:

$$(A = (I, p, m, r, d, q, r))$$

$$G_{A;\theta} = C_A \overline{L}_n^{i'} X^I T^p (T^m)_{\theta^{(r)};X^d} T^s X^q \overline{L}_n^{i''}$$

with $i' + i'' \leq 1$ and constants C_A introduced to adjust the norms:

$$|A| = |I| + p + s + m + d + q + r + i' + i'' = |G_{A;\theta}|,$$

$$\|A\| = |A| + p + s = \|G_{A;\theta}\|,$$

and

$$|||G|||_N = \sup_{supp\theta} C_A K_0^{\nu}.$$

If $\mathcal{G} = \sum F_{A'} G_{A';\theta}$ denotes a variable-coefficient sum, then its norms are given by

$$|\mathcal{G}| = \sup_{A'} |G_{A';\theta}|, \quad \|\mathcal{G}\| = \sup_{A'} \|G_{A';\theta}\|,$$

and

$$|||\mathcal{G}|||_N = \sup_{\substack{A' \\ \text{supp } \theta}} |F_{A'}| C_A K_0^m K_1^{|I|+d+q+I'+I''} K_2^{p+q} N^{|A'|+r+m}/m!$$

$(T^m)_{\theta^{(r)};X^d}$ has been introduced in Definition 10.2 above. When $m \neq 0$, all of the derivatives on θ are assumed to be sequences of T, L_n, and $\overline{L}_n$.

Definition 10.4. $\mathcal{G} = \sum F_{A'} G_{A';\theta}$ will be called

- *admissible* if each term contains an X or an $\overline{L}_n$;
- *simple* if $m = 0 = |I| + q + d + i' + i''$ in each term;
- *well normed relative to* $G'_{A;\theta}$ if $|||G|||_N \leq |||G'_{A;\theta}|||_N$, $|\mathcal{G}| \leq |A|$ and $\|\mathcal{G}\| \leq \|A\|$. For the purposes of these last two inequalities, we may admit L_n for $\overline{L}_n$ in a $G_{A';\theta}$.

Definition 10.5. $\hat{\mathcal{G}}$ will denote a sum $\sum F_{A'} G_{A';\theta}$ in which in each term, θ has been differentiated by either $\partial/\partial x_j$ or $\overline{L}_n$. The notation is intended to suggest that we may ignore such terms, since in the first instance, x-derivatives of our localizing functions are zero near the singularities of the Levi form, and in the second, $\overline{L}_n\theta$ vanishes (to high order) on the boundary, so that further consideration of such terms amounts to a consideration of the analyticity of solutions of the Dirichlet problem, which is well understood.

Definition 10.6. Let $H = H_{A;\theta}$ be given. By $\mathcal{H}_{(-1)} = \mathcal{H}_{(-1);\theta}$ we denote any expression of the form $\sum F_{A'} H_{A';\theta}$ that is well normed relative to H but $|A'| \leq |A| - 1$ for each A'. Similarly, $H_{(-2)} = \mathcal{H}_{(-2);\theta}$ will stand for any expression of the form $\sum F_{A''} H_{A'';\theta}$ that is well normed relative to H but $|A'| \leq |A| - 2$ for each A'' etc. That is, the subscript denotes the decrease in free derivatives (measured by $|\cdot|$) with no increase in the other norms.

Definition 10.7. Let $H = H_{A;\theta}$ be given. By writing $G = G_{(0)} = G_{A_{(-1)},\theta}$ we shall mean any operator $G_{B;\theta}$ with

$$|B| = |A| - 1, \quad \|B\| = \|A\| - 1, \quad \text{and} \quad |||G|||_N \leq \frac{|||H|||_N}{N}.$$

Thus, $N\mathcal{G}_{(-k)}$ is of the form $\mathcal{H}_{(-k-1)}$ for any $k \geq 0$.

Proposition 10.11. *The following relations hold:*

- *Let $H = H_{A;\theta}$ be admissible. Then*

$$H = \{X \text{ or } \overline{L}_n\} G_{(0)} + \hat{\mathcal{H}}_{(-1)} + \mathcal{H}_{(-1)}$$

(with $\mathcal{H}_{(-1)}$ admissible if $m_A \neq 0$)

$$= \{X \text{ or } \overline{L}_n\} G_{(0)} + \hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)}$$

by iteration. We may assume that $G_{(0)}$ is tangential, i.e., that $i' + i'' = 0$ in $G_{(0)}$.

- *Let* $H = H_{A;\theta}$ *be arbitrary except that* $i' + i'' = 0$. *Then*

$$HX = X'H_{(0)} + N\{\hat{\mathcal{H}}_{(0)} + \mathcal{H}_{(0)}\} = X'H_{(0)} + N\{\hat{\mathcal{H}}_{(0)} + \mathcal{H}^{sim}_{(0)}\}.$$

- *For any* $H' = H_{B;\theta}$ *with* $|B| \neq 0$, *we have (for short)*

$$H' = DG' + \hat{\mathcal{H}}'_{(-1)} = (\textstyle\sum) D_k G'^{,k} + \mathcal{H}'_{(-|H'|)} = D\mathcal{G}' + \mathcal{H}'_{(-|H'|)}.$$

We may assume that G' *and the* $G'^{,k}$ *are tangential. Similarly,*

$$H'X = DH'^{,adm} + N\mathcal{H}'_{(-|H'|)}.$$

Proof. As long as $m \neq 0$, Propositions 10.5 and 10.14 show that any errors incurred in bringing an X to the left of $(T^m)_{\theta(r)}$ still contain an X unless they fall within $\mathcal{H}'_{(-1)}$ (or $\mathcal{H}^{sim}_{(-1)}$). In the first part, if H contained an $\overline{L}_n$, we may assume that this was brought out, and hence $G_{(0)}$ is tangential. In proving the third part, we (attempt to) pull out a derivative; when the derivative is lost to the localizing function, we attempt to pull out another, etc., each time with fewer free derivatives remaining. □

In the sequel, $a(z, \overline{z}, t)$ will stand for any of a finite set of real analytic functions of $(z, \overline{z}, t)$, and G and H (or $\mathcal{G}$ and $\mathcal{H}$) will always be related as in Definition 10.7. Use of Propositions 10.5 and 10.14 will be assumed in the proofs below without explicit mention (e.g., $\mathcal{G}_{(0)}a(z, \overline{z}, t) = \mathcal{G}_{(0)}$).

Proposition 10.12. *If* $G = G_{A_{(-1)};\theta}$, *then*

$$[a(z, \overline{z}, t)X, G] = N\mathcal{G}_{(0)} = \mathcal{H}_{(-1)}.$$

Proof. The terms that arise in this commutator are the same as those that arise in Propositions 10.5 and 10.14 above. □

Proposition 10.13. *If* $G = G_{A_{(-1)};\theta}$, *then*

$$[a(z, \overline{z}, t)\overline{L}_n, G] = (\overline{L}_n)\mathcal{G}_{(-1)} = \hat{\mathcal{H}}_{(-1)} = (\overline{L}_n)N^{-1}\mathcal{H}_{(-2)} = \hat{\mathcal{H}}_{(-1)}$$

(where parenthesized vector fields may not be present).

Proof. Evident from the definitions and Proposition 10.5. □

Proposition 10.14. *If* $G = G_{A_{(-1)};\theta}$, *then*

$$[a(z, \overline{z}, t)L_n, G] = (L_n)\mathcal{G}_{(-1)} = N\widetilde{\mathcal{G}}_{(0)} = (L_n)N^{-1}\mathcal{H}_{(-2)} = \grave{\mathcal{H}}_{(-1)},$$

where $\grave{\mathcal{H}}_{(-1)}$ *indicates that the localizing function may have been differentiated by* L_n.

Proposition 10.15. *If* $G = G_{A_{(-1)};\theta}$ *(recall that each occurrence of* $a(z,\bar{z},t)$ *denotes a generic coefficient, and similarly with* X *for a generic* X*) then*

$$[a(z,\bar{z},t)X, G]a(z,\bar{z},t)X = NX\mathcal{G}_{(0)} + N\hat{\mathcal{H}}_{(-1)} + N\mathcal{H}^{sim}_{(-1)}$$
$$= X\mathcal{H}_{(-1)} + N\hat{\mathcal{H}}_{(-1)} + N\mathcal{H}^{sim}_{(-1)}.$$

Proof. From Propositions 10.12 and 10.11,

$$[a(z,\bar{z},t)X, G]a(z,\bar{z},t)X = N\mathcal{G}_{(0)}a(z,\bar{z},t)X$$
$$= Na(z,\bar{z},t)X\mathcal{G}_{(0)} + N\hat{\mathcal{H}}_{(-1)} + N\mathcal{H}^{sim}_{(-1)}.$$

□

Proposition 10.16. *If* $G = G_{A_{(-1)};\theta}$*, then*

$$[a(z,\bar{z},t)X, G]a(z,\bar{z},t)L_n = L_n\mathcal{H}_{(-1)} + N\hat{\mathcal{H}}_{(-1)}.$$

Proof.

$$[a(z,\bar{z},t)X, G]a(z,\bar{z},t)L_n = N\mathcal{G}_{(0)}L_n$$
$$= NL_n\mathcal{G}_{(0)} + N^2\grave{\mathcal{G}}_{(0)} = L_n\mathcal{H}_{(-1)} + N\grave{\mathcal{H}}_{(-1)}.$$

□

Proposition 10.17. *If* $G = G_{A_{(-1)};\theta}$*, then*

$$[a(z,\bar{z},t)\overline{L}_n, G]a(z,\bar{z},t)X = \overline{L}_n\mathcal{H}^{adm}_{(-1)} + (X \text{ or } N)\hat{\mathcal{H}}_{(-1)}.$$

Proof.

$$[a(z,\bar{z},t)\overline{L}_n, G]a(z,\bar{z},t)X = \overline{L}_n\mathcal{G}_{(-1)}X + N\hat{\mathcal{G}}_{(0)}X$$
$$= \overline{L}_n\mathcal{G}_{(-1)}X + XN\hat{\mathcal{G}}_{(0)} + N^2\hat{\mathcal{G}}_{(0)}$$
$$= \overline{L}_n\mathcal{H}^{adm}_{(-1)} + X\hat{\mathcal{H}}_{(-1)} + N\hat{\mathcal{H}}_{(-1)}.$$

□

Proposition 10.18. *If* $G = G_{A_{(-1)};\theta}$*, then*

$$[a(z,\bar{z},t)\overline{L}_n, G]a(z,\bar{z},t)L_n = L_n\mathcal{H}_{(-1)} + \overline{L}_n\grave{\mathcal{H}}_{(-2)} + N\hat{\mathcal{H}}_{(-1)}.$$

Proof. First we commute $\overline{L}_n$ to the left of the coefficient:

$$[a(z,\bar{z},t)\overline{L}_n, G]a(z,\bar{z},t)L_n = (\overline{L}_n\mathcal{G}_{(-1)} + N\hat{\mathcal{G}}_{(0)})a(z,\bar{z},t)L_n$$
$$= (L_n)((\overline{L}_n)\mathcal{G}_{(-1)} + N\hat{\mathcal{G}}_{(0)})$$
$$+\overline{L}_n(L_n\hat{\mathcal{G}}_{(-2)} + N\grave{\mathcal{G}}_{(-1)}) + N(L_n\hat{\mathcal{G}}_{(-1)} + N\hat{\mathcal{G}}_{(0)})$$
$$= L_n\grave{\mathcal{H}}_{(-1)} + \overline{L}_n\grave{\mathcal{H}}_{(-2)} + N\grave{\mathcal{H}}_{(-1)}.$$

□

Proposition 10.19. *If* $G = G_{A_{(-1)};\theta}$, *then*

$$[a(z,\bar{z},t)L_n, G](a(z,\bar{z},t)Xu'' + a(z,\bar{z},t)\tilde{\overline{L}}_n u)$$
$$= L_n\mathcal{H}_{(-1)}u'' + (L_n\mathcal{G}_{(-1)} + N\tilde{\mathcal{G}}_{(0)})\tilde{\overline{L}}_n u + X\grave{\mathcal{H}}_{(-1)}u'' + N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u''.$$

Proof.

$$[a(z,\bar{z},t)L_n, G](a(z,\bar{z},t)Xu'' + a(z,\bar{z},t)\tilde{\overline{L}}_n u)$$
$$= (L_n\mathcal{G}_{(-1)} + N\tilde{\mathcal{G}}_{(-1)}\tilde{\overline{L}}_n u) + N\grave{\mathcal{G}}_{(0)}(aXu'' + a\tilde{\overline{L}}_n u)$$
$$= L_n\mathcal{H}_{(-1)} + (L_n\mathcal{G}_{(-1)} + N\grave{\mathcal{G}}_{(0)})\ \tilde{\overline{L}}_n u\Big) + XN\grave{\mathcal{G}}_{(0)}u'' + N\grave{\mathcal{H}}_{(-1)}u''$$
$$= L_n\mathcal{H}_{(-1)}u'' + (L_n\mathcal{G}_{(-1)} + N\grave{\mathcal{G}}_{(0)})\tilde{\overline{L}}_n + X\grave{\mathcal{H}}_{(-1)}u'' + N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u''.$$

Note that $[\grave{\mathcal{H}}_{(-1)}, \partial/\partial x] = N\hat{\mathcal{H}}_{(-1)}$. □

Proposition 10.20. *If* $G = G_{A_{(-1)};\theta}$, *then*

$$[a(z,\bar{z},t)XG]a(z,\bar{z},t)(\overline{L}_n u' \ or\ \tilde{\overline{L}}_n u)$$
$$= N\mathcal{G}_{(0)}\tilde{\overline{L}}_n u + N\mathcal{G}_{(0)}u = \mathcal{H}_{(-1)}\tilde{\overline{L}}_n u + \mathcal{H}_{(-1)}u.$$

Proof. Propositions 10.12 and 10.13. □

10.6 End of the Proof of Theorem 10.2

Let H^{adm} be admissible. Then from Proposition 10.11 and the a priori estimate (9.1) we know that (writing G for $G_{(0)}$, which, as in Proposition 10.11, is tangential)

$$\|H^{adm}u\|^2_{L^2} \le \|(X \text{ or } \overline{L}_n)Gu\|^2_{L^2} + \|\hat{\mathcal{H}}_{(-1)}u\|^2_{L^2} + \|\mathcal{H}^{sim}_{(-1)}u\|^2_{L^2}, \tag{10.21}$$

and in obvious notation, using (10.20) for all $\mathcal{K}$,

$$\|XGu\|^2_{L^2} + \|\overline{L}_n Gu\|^2_{L^2} + \|Gu''\|^2_{H^1} + \mathcal{K}\|Gu\|^2_{L^2}$$
$$\le CQ(Gu, Gu) + C_{\mathcal{K}}\|Gu\|^2_{-1} \tag{10.22}$$

and

$$Q(Gu, Gu) \le C\left(\|G\alpha\|^2_{L^2} + \|Gu\|^2_{L^2} + \sum |E_j|\right), \tag{10.23}$$

where

- $E_1 = ([\overline{\partial}, G]u, \overline{\partial}Gu)_{L^2}$,
- $E_2 = (\overline{\partial}u, [\overline{\partial}, G^*]Gu)_{L^2}$,
- $E_3 = ([\overline{\partial}^*, G]u, \overline{\partial}^* Gu)_{L^2}$,
- $E_4 = (\overline{\partial}^* u, [\overline{\partial}^*, G^*]Gu)_{L^2}$.

Now from the action of $\overline{\partial}$ and its adjoint, not all components of u are subjected to all vector fields. In particular, we may write

$$\overline{\partial}u = a(z, \bar{z}, t)(X)u + (\overline{\partial}u)'', \tag{10.24}$$

where

$$(\overline{\partial}u)'' = a(z, \bar{z}, t)Xu'' + a(z, \bar{z}, t)\overline{L}_n u' + Au \equiv a(z, \bar{z}, t)Xu'' + \widetilde{\widetilde{\overline{L}}}_n u$$

defines the operator $\widetilde{\widetilde{\overline{L}}}_n$ and

$$\overline{\partial}^* u = a(z, \bar{z}, t)(X)u + a(z, \bar{z}, t)(L_n)u''. \tag{10.25}$$

Thus for any $\rho > 0$ there exists a constant C_ρ such that

$$|E_1| + |E_3| \leq \rho Q(Gu, Gu) + C_\rho(E_1' + E_3'), \tag{10.26}$$

where from Propositions 10.12 and 10.13,

$$E_1' = \|[a(z, \bar{z}, t)(X \text{ or } \overline{L}_n), G]u\|_{L^2}^2 \leq C\left(\|(\overline{L}_n)\mathcal{G}_{(-1)}u\|_{L^2}^2 + \|\mathcal{H}_{(-1)}u\|_{L^2}^2\right) \tag{10.27}$$

and

$$\begin{aligned} E_3' = \|[\text{coeff}\, L_n, Gu]u''\|_{L^2}^2 &\leq C\left(\|L_n\mathcal{G}_{(-1)}u''\|_0^2 + \|\hat{\mathcal{H}}_{(-1)}u''\|_0^2\right) \\ &\leq C\left(\|L_n\mathcal{G}_{(-1)}u''\|_0^2 + \|\hat{\mathcal{H}}_{(-|H|)}u''\|_0^2\right) \end{aligned}$$

by Proposition 10.11. The remaining two terms are expanded as follows:

$$\begin{aligned} E_2 &= (\overline{\partial}u, [\overline{\partial}, G^*]Gu)_{L^2} \\ &= (\text{coeff}\,((X) + \widetilde{\widetilde{\overline{L}}}_n)u, [\text{coeff}\,((X) + \widetilde{\widetilde{\overline{L}}}_n), G^*]Gu)_{L^2} \\ &= ([\text{coeff}\,(X), G](\text{coeff}\,((X) + \widetilde{\widetilde{\overline{L}}}_n)u), Gu)_{L^2} \\ &\quad + ([\text{coeff}\,(L_n), G](\text{coeff}\,((X) + \widetilde{\widetilde{\overline{L}}}_n)u), Gu)_{L^2} \\ &= E_{21} + E_{22} \end{aligned}$$

and

$$\begin{aligned}E_4 &= \overline{\partial}^* u, [\overline{\partial}^*, G^*]Gu_{L^2}\\ &= (\text{coeff}\,(Xu + L_n u''), [\text{coeff}\,X, G^*]Gu + [\text{coeff}\,L_n, G^*]Gu'')_{L^2}\\ &= ([\text{coeff}\,X, G_{A;\theta}]\text{coeff}\,(Xu + L_n u''), Gu)_{L^2}\\ &\quad +([\text{coeff}\,\overline{L}_n, G]\text{coeff}\,(Xu + L_n u''), Gu'')_{L^2}\\ &= E_{41} + E_{42},\end{aligned}$$

where the integrations by parts are justified, since u'' and $\widetilde{\widetilde{\overline{L}}}_n u = 0$ on Γ and G is tangential. Now E_{21} is treated by Proposition 10.15, E_{22} by Proposition 10.19, E_{41} by Propositions 10.17 and 10.18. The results are

$$E_{21} = (\mathcal{H}_{(-1)}u, XGu)_{L^2} + (N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u, Gu)_{L^2} + (\mathcal{H}_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u, Gu)_{L^2},$$

which lends itself to the estimate

$$\begin{aligned}|E_{21}| &\le \|\mathcal{H}_{(-1)}u\|_{L^2}\|XGu\|_{L^2}\\ &\quad +\|\mathcal{H}_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u\|_{L^2}\|Gu\| + \|N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u\|_{L^2}\|Gu\|^{L^2}\end{aligned} \tag{10.28}$$

(note that the expression $\mathcal{H}_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u$ is of a new type),

$$\begin{aligned}E_{22} &= (\mathcal{H}_{(-1)}u'', \overline{L}_n Gu)_{L^2} + ((XN\grave{\mathcal{G}}_{(0)} + N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)}))u'', Gu)_{L^2}\\ &\quad + (N\grave{\mathcal{G}}_{(0)}\widetilde{\widetilde{\overline{L}}}_n u, Gu)_{L^2} + L_n\mathcal{G}_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u, Gu)_{L^2},\end{aligned}$$

which lends itself to the estimate

$$\begin{aligned}|E_{22}| &\le \|\grave{\mathcal{H}}_{(-1)}u''\|_{L^2}\|(X \text{ or } \overline{L}_n \text{ or } N)Gu\|_{L^2}\\ &\quad +\|(\grave{\mathcal{H}}_{(-1)} \text{ or } L_n\mathcal{G}_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u\|_{L^2}\|Gu\|_{L^2}.\end{aligned} \tag{10.29}$$

Next,

$$\begin{aligned}E_{41} &= (\mathcal{H}_{(-1)}u, XGu)_{L^2} + (N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u, Gu)_{L^2}\\ &\quad +([\text{coeff}\,X, G]\text{coeff}\,L_n u'', Gu)_{L^2},\end{aligned}$$

and using Proposition 10.11 to expand the third term below,

$$\begin{aligned}[\text{coeff}\,X, G]\text{coeff}\,L_n &= \text{coeff}\,L_n[\text{coeff}\,X, G]\\ &\quad +[\text{coeff}\,(X \text{ or } L_n, G] + [\text{coeff}\,X, \text{coeff}\,L_n, G]]\\ &= (\overline{L}_n)\mathcal{H}_{(-1)} + \hat{\mathcal{H}}_{(-1)} + X\grave{\mathcal{H}}_{(-1)} + N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)}),\end{aligned} \tag{10.30}$$

so that

$$|E_{41}| \leq \left(\|\mathcal{H}_{(-1)}u\|_0 + \|\grave{\mathcal{H}}_{(-1)}u''\|_0\right)\|(X \text{ or } \overline{L}_n)Gu\|_0 \tag{10.31}$$

and

$$E_{42} = (\overline{L}_n\mathcal{H}^{adm}_{(-1)}u + (X \text{ or } L_n \text{ or } N)\hat{\mathcal{H}}_{(-1)}u + \overline{L}_n\hat{\mathcal{H}}_{(-2)}u'', Gu'')_0,$$

so that

$$|E_{42}| \leq \left(\|\mathcal{H}^{adm}_{(-1)}u\|_0 + \|N\hat{\mathcal{H}}_{(-1)}u\|_0 + \|\grave{\mathcal{H}}_{(-2)}u''\|_0\right)\|(D)Gu''\|_0.$$

Thus, using a weighted Schwarz inequality on the sum of the absolute values of $E_1, E_3, E_{21}, E_{22}, E_{41}$, and E_{42}, we have, for any $\rho > 0$,

$$\begin{aligned}(1-\rho)&\left(\|XGu\|_0^2 + \|\overline{L}_nGu\|_0^2 + \|Gu''\|_{H^1}^2 + Q(Gu, Gu)\right) + (\mathcal{K} - C_\rho)\|Gu\|_0^2 \\ &\leq C\|G\alpha\|_0^2 + \rho\left(\|\grave{\mathcal{H}}_{(-1)}\widetilde{\widetilde{L}}_nu\|_0^2 + \|N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u\|_0^2\right) \\ &\quad + C_\rho\left(\|\mathcal{H}_{(-1)}u\|_0^2 + \|\grave{\mathcal{G}}_{(-1)}u''\|_0^2 + \|\mathcal{H}_{(|H|}u\|_0^2\right) + C_{\mathcal{K}}\|Gu\|_{-1}^2.\end{aligned} \tag{10.32}$$

Note that we have not written NGu as $\mathcal{H}_{(-1)}u$ on the right. The term $\|\mathcal{H}_{(-1)}u\|_0^2$ arises naturally and unavoidably in (10.27) and E_{21} but always as a result of a bracket, never as NGu. Also, note that

$$C_{\mathcal{K}}\|Gu\|_{-1}^2 \leq \|Gu\|_0^2 + C(\mathcal{K})^{|H|}\|Gu\|_{-|H|}^2 \tag{10.33}$$

by the logarithmic convexity of the Sobolev norms.

The second term on the right of (10.32) is new; to handle it we use coercive estimates on $\widetilde{\widetilde{L}}_nu$ but first must extract a tangential derivative from $\mathcal{H}_{(-1)}$: using Proposition 10.11 on each of the terms in $\mathcal{H}_{(-1)}\widetilde{\widetilde{L}}_n$,

$$H_{(-1)}\widetilde{\widetilde{L}}_n = DG_{(-1)}\widetilde{\widetilde{L}}_n + H_{(-|H|)}\widetilde{\widetilde{L}}_n. \tag{10.34}$$

Noting that (10.20) implies that $\widetilde{\widetilde{L}}_nu = 0$ on $\partial\Omega$, we apply the coercive estimate

$$\|DG_{(-1)}\widetilde{\widetilde{L}}_nu\|_0^2 \leq C\left(Q(G_{(-1)}\widetilde{\widetilde{L}}_nu, G_{(-1)}\widetilde{\widetilde{L}}_nu)_0 + \|G_{(-1)}\widetilde{\widetilde{L}}_nu\|_0^2\right) \tag{10.35}$$

with

$$|Q(G_{(-1)}\widetilde{\widetilde{L}}_nu, G_{(-1)}\widetilde{\widetilde{L}}_nu)_0| \leq C\left(\|G_{(-1)}\widetilde{\widetilde{L}}_n\alpha\|_0^2 + \|G_{(-1)}\widetilde{\widetilde{L}}_nu\|_0^2 + \sum|F_j|x\right), \tag{10.36}$$

where

$$
\begin{aligned}
F_1 &= ([\bar{\partial}, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}]u, \bar{\partial} G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0, \\
F_2 &= (\bar{\partial} u, [\bar{\partial}, (G_{(-1)}\widetilde{\widetilde{\overline{L}_n}})^*]G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0, \\
F_3 &= ([\bar{\partial}^*, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}]u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0, \\
F_4 &= (\bar{\partial}^* u, [\bar{\partial}^*, (G_{(-1)}\widetilde{\widetilde{\overline{L}_n}})^*]G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0.
\end{aligned}
$$

Now for F_1 and F_3, again, we use a weighted Schwarz inequality and have the estimate, for any $\sigma > 0$,

$$
\begin{aligned}
|F_1| &= |([\bar{\partial}, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}]u, \bar{\partial} G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0| \\
&\le \sigma Q(G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u) + C_\sigma \|[\text{coeff}\,(X \text{ or } \overline{L}_n)G_{(-1)}]\widetilde{\widetilde{\overline{L}_n}}u\|_0^2 \\
&\quad + C_\sigma \|G_{(-1)}[\text{coeff}\,(X \text{ or } \overline{L}_n), \overline{L}_n + A]u\|_0^2 \\
&\le \sigma Q(G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u) + C_\sigma \left(\|\mathcal{H}_{(-2)}\widetilde{\widetilde{\overline{L}_n}}u\|_0^2 + \|\mathcal{H}_{(-1)}^{adm}\|_0^2 \right)
\end{aligned}
$$

and

$$
\begin{aligned}
|F_3| &= |([\bar{\partial}^*, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}]u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0|, \text{ and} \\
&\le C|([\text{coeff}\,(X \text{ or } L_n), G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}]u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0| \\
&\le C|(D\grave{\mathcal{G}}_{(-2)}\widetilde{\widetilde{\overline{L}_n}} + \mathcal{H}_{(-1)}^{adm} + \mathcal{H}_{(-|H|)})u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u_0| \\
&\quad + C|(G_{(-1)}\text{coeff}\, L_n u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0|.
\end{aligned}
$$

But $G_{(-1)}\text{coeff}\, L_n = L_n G_{(-1)} + \grave{\mathcal{H}}_{(-1)} = L_n G_{(-1)} + D\grave{\mathcal{G}}_{(-2)} + \hat{\mathcal{H}}_{(-1)} + \grave{\mathcal{H}}_{(-|H|)}$, and if we expand $\bar{\partial}^*$ as coeff(X) + coeff L_n in this last term, we may integrate by parts (twice) to obtain

$$
\begin{aligned}
&(G_{(-1)}\text{coeff}\, L_n u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0 \\
&\quad = (((\text{coeff}\,(\overline{L}_n + X)\mathcal{G}_{(-1)} + \grave{\mathcal{H}}_{(-|H|)})u, \text{coeff}\,((X) + \overline{L}_n)G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0 \\
&\qquad + \text{coeff}\,\hat{\mathcal{H}}_{(-1)}u, \bar{\partial}^* G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0,
\end{aligned}
$$

so that we obtain, with the coercive estimate (10.35),

$$
\begin{aligned}
|F_3| &\le \sigma Q(G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u, G_{(-1)}\widetilde{\widetilde{\overline{L}_n}}u)_0 \\
&\quad + C_\sigma \left(\|\grave{\mathcal{H}}_{(-2)}\widetilde{\overline{L}_n}u\|_0^2 + \|(\mathcal{H}_{(-1)}^{adm} \text{ or } \hat{\mathcal{H}}_{(-1)})u\|_0^2 + \|\grave{\mathcal{H}}_{(-|H|)}u\|_0^2 \right).
\end{aligned}
$$

For the terms F_2 and F_4 we have

$$
\begin{aligned}
F_2 &= (\overline{\partial} u, [\overline{\partial}, (G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n)^* G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= (\text{coeff}\,((X) + \overline{L}_n)u, [\text{coeff}\,(X), (G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n)^*]G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +(\text{coeff}\,((X) + \overline{L}_n)u, [\text{coeff}\,(\overline{L}_n), (G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)^*]G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= ([\text{coeff}\,(X \text{ or } L_n), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u]\text{coeff}\,\widetilde{\widetilde{\overline{L}}}_n u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +([\text{coeff}\,(X), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n]\text{coeff}\,(X)u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +([\text{coeff}\,(L_n), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n]\text{coeff}\,(X)u'', G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= (((D)\grave{\mathcal{H}}_{(-2)} + N\mathcal{H}_{(-2)})L_n u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +((\mathcal{H}_{(-2)}(\widetilde{\widetilde{\overline{L}}}_n) + H_{(-1)})(X)u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +(((\grave{\mathcal{H}}_{(-2)} + L_n G_{(-2)})\widetilde{\widetilde{\overline{L}}}_n + L_n G_{(-1)})\text{coeff}\,(X)u'', G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= (((D)\grave{\mathcal{H}}_{(-2)} + N\mathcal{H}_{(-2)})L_n u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +((\mathcal{H}_{(-2)}X(\widetilde{\widetilde{\overline{L}}}_n) + H_{(-1)}X)u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +(((N \text{ or } X \text{ or } L_n)\grave{\mathcal{H}}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n + (X \text{ or } N \text{ or } L_n\grave{\mathcal{H}}_{(-2)})u'', G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= (((D)\grave{\mathcal{H}}_{(-2)} + N\mathcal{H}_{(-2)})L_n u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +(((X + N)\mathcal{H}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n + D\mathcal{H}^{adm}_{(-1)} + \mathcal{H}_{(-|H|)})u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +(((N \text{ or } X \text{ or } L_n)\grave{\mathcal{H}}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n + (X \text{ or } N \text{ or } L_n\grave{\mathcal{H}}_{(-2)})u'', G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0,
\end{aligned}
$$

and so

$$
\begin{aligned}
|F_2| \le {}& \|(D \text{ or } N)G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u\|_0 \\
&\times \left(\|\grave{\mathcal{H}}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n u\|_0 + \|G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u\|_0 + \|\mathcal{G}^{adm}_{(-1)} u\|_0 + \grave{\mathcal{G}}_{(-2)} u''\|_0 \right),
\end{aligned}
$$

while

$$
\begin{aligned}
F_4 &= (\overline{\partial}^* u, [\overline{\partial}^*, (G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n)^*]G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= (\text{coeff}\,(X)u + \text{coeff}\,L_n u'', [\text{coeff}\,(X), (G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n)^*]G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +(\text{coeff}\,(X)u + \text{coeff}\,L_n u'', [\text{coeff}\,L_n, (G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n)^*]G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&= ([\text{coeff}\,(X), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n](\text{coeff}\,(X)u + \text{coeff}\,L_n u''), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0 \\
&\quad +([\text{coeff}\,(\overline{L}_n), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n](\text{coeff}\,(X)u + \text{coeff}\,L_n u''), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0
\end{aligned}
$$

$$= ((\mathcal{H}^{adm}_{(-1)} \text{ or } \mathcal{H}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n)(\text{coeff}\,(X)u + \text{coeff}\, L_n u''), G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0$$

$$= ((X \text{ or } N)(\mathcal{H}^{adm}_{(-1)} \text{ or } \mathcal{H}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n)u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0$$

$$+((L_n(\mathcal{H}^{adm}_{(-1)} \text{ or } \mathcal{H}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n) + (\grave{\mathcal{H}}^{adm}_{(-1)} \text{ or } \grave{\mathcal{H}}_{(-2)})\widetilde{\widetilde{\overline{L}}}_n)u'', G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)_0.$$

Thus we also have

$$|F_4| \le C\,\|(D \text{ or } N)G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)\|_0$$
$$\times \left(\|\grave{\mathcal{H}}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n u\|_0 + \|G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u\|_0 + \|\grave{\mathcal{H}}^{adm}_{(-1)}u\|_0 + \|\grave{\mathcal{H}}_{(-2)}u''\|_0\right).$$

Collecting the estimates of the F_j, we have

$$\sum |F_j| \le \sigma Q(G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u, G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u)$$
$$+C_\sigma \left(\|\grave{\mathcal{H}}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n u\|_0^2 + \|\mathcal{H}^{adm}_{(-1)}u\|_0^2 + \|\grave{\mathcal{H}}_{(-2)}u''^2_0 + \|\grave{\mathcal{H}}_{(-|H|)}u\|_0^2\right),$$

so that (10.35) and (10.36) may be rewritten

$$\|\grave{\mathcal{H}}_{(-1)}\widetilde{\widetilde{\overline{L}}}_n u\|_0^2 \le C\left(\|G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n \alpha\|_0^2 + \|\grave{\mathcal{H}}_{(-2)}\|_0^2\right.$$
$$\left. +\|\mathcal{H}^{adm}_{(-1)}u\|_0^2 + \|\mathcal{H}_{(-2)}u''\|_0^2 + \|\grave{\mathcal{H}}_{(|H|)}\|_0^2\right). \tag{10.37}$$

Together with (10.32), This yields

$$(1-\rho)(\|XGu\|_0^2 + \|\overline{L}_n u\|_0^2 + \|Gu''\|_{H^1}^2 + Q(Gu, Gu))$$
$$\le C(\|G\alpha\|_0^2 + \|G_{(-1)}\widetilde{\widetilde{\overline{L}}}_n \alpha\|_0^2) + C_\rho\|\grave{\mathcal{H}}_{(-2)}\widetilde{\widetilde{\overline{L}}}_n u\|_0^2$$
$$+C_{\mathcal{K}}\|Gu\|_{-1}^2 + \|N(\hat{\mathcal{H}}_{(-1)} + \mathcal{H}^{sim}_{(-1)})u_0^2)$$
$$+C_\rho\left(\|\mathcal{H}_{(-1)}u\|_0^2 + \|\grave{\mathcal{G}}_{(-1)}u''\|_0^2 + \|\mathcal{H}_{(|H|)}u\|_0^2\right). \tag{10.38}$$

As long as the terms in $\mathcal{H}_{(-1)}$ on the right are admissible (or act on u''), we may iterate (10.38). If $m = 0$ for an inadmissible term, we do not process it further at this point. On the other hand, when $m \neq 0$ in an inadmissible term, such a term is necessarily of the form

$$G'_{A';\Psi} = C_{A'}T^p(T^m)_{(\partial/\partial t)^r\Psi}T^s \tag{10.39}$$

(or $G'_{A_{(-k)};\Psi} = C_{A'}T^p(T^m)_{(\partial/\partial t)^r\{\Psi_x \text{ or } \overline{L}_n\Psi\}}T^s$; cf. below) with

$$|||G'_{A_{(-k)};\Psi}|||_N \le |||H_{A,\Psi}|||_N, |A_{(-k)}| \le |A| - k, \|A_{(-k)}\| \le \|A\|,$$

since we may always write $L_n = \partial/\partial t$ modulo $\overline{L}_m$. It is not immediately clear how to proceed, since to use (9.1) again with maximal advantage we would have to adjoin an X or an $\overline{L}_n$ to the left of $T^p(T^m)_{(\partial/\partial t)^r\Psi}T^s$, a process that would increase all three norms. Now in fact we shall do just that, but we must show that, modulo acceptable errors (namely when $\overline{L}_n$ lands on the localizing function):

(1) in arriving at $T^p(T^m)_{(\partial/\partial t)^r\Psi}T^s$ in fact we have already gained a factor of N in $|||\cdot|||_N$, and
(2) after adjoining an X or $\overline{L}_n$ to $T^p(T^m)_{(\partial/\partial t)^r\Psi}T^s$, the next iteration of (9.1) drops $|\cdot|$ and $\|\cdot\|$ *both* by at least one, so that that after these two applications of (9.1) the $|||\cdot|||_N$ norm is unchanged, $|\cdot|$ has dropped by one, and $\|\cdot\|$ is once again no greater than at the start.

For (1), there are five ways in which all free X's or $\overline{L}_n$ may be lost. The first and second occur when $\overline{L}_n$ or L_n lands on the localizing function, and we shall discuss such terms below; the third occurs when in Proposition 10.14, two X's are bracketed to generate a T. In general, the form of this bracket is $[X^I, X]$ with $|I| \leq N$, i.e., at most $|I|$ terms with a T and at most $|I|-1$ remaining X's *but at least one if* $|I| \geq 2$. Hence if such a bracketing uses up all remaining X's, $|I|$ had to be 1, and the result of this operation is to *decrease* $|||\cdot|||_N$ *by a factor of* N. And the fourth way that *all* free X's or $\overline{L}_n$ might be lost is as derivatives of a coefficient, and in this case we merely need to shift the count of derivatives by one: after all, the term in

$$[X^I, f] = \sum \binom{|I|}{|I'|} f^{(|I'|)} X^{I-I'} \quad \text{with } I = I'$$

has coefficient $f^{(|I|)}$, and

$$|f^{(|I|)}| \leq C^{(|I|)}|I|! \leq C'^{|I|-1}(|I|-1)! \leq C'^{|I|-1}N^{|I|-1}.$$

Thus, modulo a constant for each derivative that falls on a function (which is always permitted by taking the K_i to have the proper ratios in (10.15)), the loss of the last X in this fashion does not need to result in a corresponding factor of N in the norm. When $\overline{L}_n$ lands on the localizing function we are in a favorable situation, since then coercive estimates apply. When L_n lands on the localizing function the situation is not favorable, and we examine this case in detail below. Finally, on the right-hand side of many of the estimates for the E_{ij}, we often have just $\|G\|_{L^2}$, which may be in the form (10.40). But in every case we have given a form of the estimate that *omits* a factor of N in front of such a term, hence the gain of N in this case as well. To balance this, we must be willing to treat the term $\|N\hat{\mathcal{H}}_{(-1)}\mathcal{H}^{sim}_{(-1)}u\|_0$ or $\|N\mathcal{H}^{sim}_{(-1)}u\|_0$. This, too, we discuss below.

For (2), we merely note that in the application of p. 120, i.e., from Propositions 10.5 and 10.14, when $|I|+q+d+i'+i''=0$, not only does $|\cdot|$ drop, as it always does, but also $\|\cdot\|$ will drop. That is, two X's do not bracket to generate a T, since there is no second X! Even in F_2, where we appeared to use a second bracket to get an X to the right, we didn't need to use a bracket, merely to extract an X on the left from $G_{A'_{(-1)};\Psi}X$, which never generates a T.

The difficult second case under (1) above occurs when L_n (i.e., a T modulo an $\overline{L}_n$,) lands on the localizing function Ψ, leaving a term of the form

$$G'_{a';\Psi} = C_{A'}T^p(T^m)_{(\partial/\partial t)^r\Psi}T^s. \tag{10.40}$$

A first remark is that if such a term operates on u'', it should be considered admissible after all, since u'' is zero on the boundary and thus satisfies coercive estimates (i.e., we may extract *any* derivative, for example a T, as in Proposition 10.11). In the expansions above, E'_4 contains such a term, but either against $(X)u''$ again or with $\widetilde{\overline{L}}_n u$. If E_{41} or E_{42} generates such a term by commuting the L_n to the left prior to integrating by parts, once again it is combined with u''. Thus we are led to consider only the second type of terms, namely $F_1, \ldots, F_4$. Of these, F_1 never contributes L_n on a localizing function, while F_3 also never does; the terms of the form $\mathcal{H}_{(-1)}u$ that F_3 contributes are all in fact admissible.

There remain two types of terms to handle: those that vanish on the boundary and those that we have called "simple". Terms that vanish on the boundary, whether of the form $\|(N)\hat{\mathcal{H}}_{(-1)}u''\|_0$ or $\|\mathcal{H}_{(-1)}u''\|_0$, satisfy coercive estimates that are much stronger than the maximal estimate. The additional factor of N here merely stays with such a term until a new localizing function must be introduced, as in the previous section, and does not disturb the final estimates. A small additional problem is posed by u'', in that while u'' is also a solution to the $\overline{\partial}$-Neumann problem, its right-hand side is not α''; i.e., we shall have to examine $Q(u'', G^*Gu'')$ and relate it to α. But $Q(u, G^*Gu'') - Q(u'', G^*Gu'') = Q(u', G^*Gu'')$, and since the principal part of Q is diagonal, $Q(u', G^*Gu'')$ is a sum of terms of the form $(\text{coeff}\, u, DG^*Gu'')_0$, which will be absorbed easily.

Finally, there are the simple terms, those without X's and with $m = 0$. But these are handled just as above: a new localizing function is introduced that will need to handle only half the original number of derivatives, though in a somewhat larger open set. The process terminates after $\log_2 N$ such iterations.

Chapter 11
Nonsymplectic Strata and *Germ* Analytic Hypoellipticity

In all of the above cases, the characteristic variety for the operator has been symplectic, in fact, a symplectic manifold. This is in agreement with the spirit of Treves' conjecture that in order to have analytic hypoellipticity, the characteristic variety and all the subsidiary Poisson strata should be symplectic.

However, if one alters the definition of analytic hypoellipticity and replaces it by analytic hypoellipticity in the sense of germs, there are situations in which not all the strata are symplectic yet one has analyticity in this sense.

In a recent paper [Han], Hanges considered the operator

$$P_H = \partial_t^2 + t^2 \Delta_x + \partial_{\theta(x)}^2 = \sum_1^4 X_j^2 \tag{11.1}$$

in $\mathbb{R}^3$, where $\partial_{\theta(x)} = x_1 \partial/\partial x_2 - x_2 \partial/\partial x_1$, and made the interesting distinction between analytic hypoellipticity (AHE) in the germ sense and AHE in the *strict* sense. While it is well known that the operator in (11.1) is not microlocally analytic hypoelliptic, Hanges gave a proof in [Han] by means of explicit constructions that the operator P_H is not analytic hypoelliptic *in the strict sense* in any open set U containing the origin, i.e., does not have the property that for any open subset V of U, if Pu is analytic in V then so is the solution u, yet has the property that if Pu is analytic in some neighborhood of the origin then so is u in a (possibly smaller) neighborhood of the origin. He ties this result to the conjecture of Treves concerning the Poisson strata of the operator P, namely that if one writes $P = \sum_1^4 X_j^2$, and considers the successive strata where (1) all X_j vanish, (2) all X_j and their first brackets vanish, (3) all X_j and their first and second brackets vanish, etc., then the operator should be analytic hypoelliptic in the strict sense if and only if all these strata are symplectic. In the case of the particular operator being considered here, not even the characteristic variety is symplectic, being given by $t = \tau = x_1 \xi_2 - x_2 \xi_1 = 0$.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_11,

Here we give a very elementary, and flexible, proof of the affirmative part of his result and argue that the negative part is entirely reasonable as well, though in our proof we avoid completely any mention of so-called Treves curves, which foliate the characteristic variety of P.

We also consider more general cases in which there is a stratum of the characteristic manifold that is not symplectic: the symplectic form vanishes on that stratum. Our proofs, as usual, are entirely "elementary" and use little beyond L^2 estimates and careful localization in certain variables.

11.1 Proof for Hanges' Operator (11.1)

As remarked above, we may take localizing functions to be independent of t, since were a derivative in t to land on such a localizer, one would be in the region where the operator was clearly elliptic and the analyticity of the solution u was well known. We denote such an Ehrenpreis-type localizing function by $\varphi(x) = \varphi_N(x)$ subject to the usual growth of its derivatives: $|D^\alpha \varphi| \le C^{|\alpha|+1} N^{|\alpha|}$ for $|\alpha| \le N$, where the constant C is (universally) inversely proportional to the width of the band separating the regions where $\varphi \equiv 0$ and $\varphi \equiv 1$.

Next, since P is C^∞ hypoelliptic, we may assume that u is smooth and proceed to obtain estimates for $D_t^p u$ and $D_{x_j}^p u$ near 0.

The a priori estimate for P, while subelliptic, is more importantly maximal: for $v \in C_0^\infty$,

$$\|D_t v\|_{L^2}^2 + \sum_1^2 \|t D_{x_j} v\|_{L^2}^2 + \|D_{\theta(x)} v\|_{L^2}^2 \; (+ \, \|v\|_{1/2}^2) \le C |\langle P v, v\rangle| + C \|v\|_{L^2}^2. \tag{11.2}$$

Setting $v = \varphi D_t^p u$, to begin with, we obtain

$$\begin{aligned}
&\|D_t \varphi D_t^p u\|_{L^2}^2 + \sum_1^2 \|t D_{x_j} \varphi D_t^p u\|_{L^2}^2 + \|D_{\theta(x)} \varphi D_t^p u\|_{L^2}^2 \; (+ \, \|\varphi D_t^p u\|_{1/2}^2) \\
&\quad \le C |\langle P \varphi D_t^p u, \varphi D_t^p u\rangle| + C \|\varphi D_t^p u\|_{L^2}^2 \\
&\quad \le C |\langle \varphi D_t^p P u, \varphi D_t^p u\rangle_{L^2}| + C \sum_1^4 |([X_j^2, \varphi D_t^p] u, \varphi D_t^p u)| + C \|\varphi D_t^p u\|_{L^2}^2.
\end{aligned}$$

Now crucial among the brackets are those in the next two displays (recall that we may take φ independent of t, and clearly to localize in x we may take it to be purely *radial* in (x_1, x_2), i.e., we choose φ to be constant on the integral curves of X_4), so that $X_4 \varphi = 0$,

$$[X_1, \varphi D_t^p] = [X_4, \varphi D_t^p] = 0,$$

and

$$[X_j, \varphi D_t^p] = t\varphi' D_t^p - \underline{p}\varphi D_x D_t^{p-1}, \qquad j = 2, 3.$$

In the first case, we may ignore the factor t and recognize the passage from one power of D_t to a derivative on φ as an acceptable swing, which, upon iteration, will lead to $C^{p+1}N^p \sim C^{p+1}p!$ when $p \sim N$. The second term takes two powers of D_t (e.g., X_1 from the estimate and one power of D_t) and produces a factor of p and a "bad" vector field D_x. Iterating this will yield $p!!D_x^{p/2}u \sim p!^{1/2}D_x^{p/2}u$ on the support of φ.

On the other hand, setting $v = \varphi D_{x_j}^q u$, with perhaps $q = p/2$, or, better, $v = \varphi\Delta_x^{q/2}u$, where we write $\Delta_x = \sum_j D_{x_j}^2$, we obtain

$$\begin{aligned}
&\|D_t\varphi\Delta_x^{q/2}u\|_{L^2}^2 + \sum_1^2 \|tD_{x_j}\varphi\Delta_x^{q/2}u\|_{L^2}^2 + \|D_{\theta(x)}\varphi\Delta_x^{q/2}u\|_{L^2}^2 (+ \|\varphi\Delta_x^{q/2}u\|_{1/2}^2) \\
&\quad \le C|\langle P\varphi\Delta_x^{q/2}u, \varphi\Delta_x^{q/2}u\rangle| + C\|\varphi\Delta_x^{q/2}u\|_{L^2}^2 \\
&\quad \le C|\langle \varphi\Delta_x^{q/2}Pu, \varphi\Delta_x^{q/2}u\rangle_{L^2}| + C\sum_1^4 |\langle [X_j^2, \varphi\Delta_x^{q/2}]u, \varphi\Delta_x^{q/2}u\rangle| \\
&\qquad + C\|\varphi\Delta_x^{q/2}u\|_{L^2}^2,
\end{aligned}$$

and now the crucial brackets are

$$[X_1^2, \varphi\Delta_x^{q/2}] = 0, [X_4^2, \varphi\Delta_x^{q/2}] = 0$$

and

$$[X_j^2, \varphi\Delta_x^{q/2}] = 2X_j t\varphi'\Delta_x^{q/2} - t^2\varphi^{(2)}\Delta_x^{q/2}, \qquad j = 2, 3$$

(where we have used rather heavily the fact that $X_4\varphi = 0$, since φ depends only on x, and radially so, and that in fact $[D_\theta, \Delta_x] = 0$).

This last line leads to two kinds of terms, namely, for $j = 2, 3$,

$$\langle 2X_j t\varphi'\Delta_x^{q/2}u, \varphi\Delta_x^{q/2}u\rangle$$

and

$$\langle t^2\varphi^{(2)}\Delta_x^{q/2}u, \varphi\Delta_x^{q/2}u\rangle.$$

Morally, these terms show the correct gain to lead to analytic growth of derivatives, namely one must think of $t\Delta^{1/2}$ as an X_j with $j = 2$ or $j = 3$, and so in the first term above one merely integrates by parts, noting that $X_j^* = -X_j$, and obtains, after a weighted Schwarz inequality, a small multiple of the left-hand side of the a priori inequality and the square of a term with one derivative on φ and $t\Delta^{1/2}$ and q reduced by one, though one more commutator is required to make the order correct, and this will introduce another derivative on φ and q again decreased by one unit, etc.

The second term is of a different character, though the same observation reduces us essentially to

$$\langle X\varphi^{(2)}\Delta_x^{(q-1)/2}u, X\varphi\Delta_x^{(q-1)/2}u\rangle,$$

in which instead of each copy of φ receiving one derivative, we have two derivatives on one copy and none on the other. Fortunately, the Ehrenpreis-type cut-off functions may be differentiated not merely N times with the usual growth, but $2N$ or $3N$ with no change—so in the above inner product we include a factor CN with the copy of φ that remains undifferentiated and a factor of $(CN)^{-1}$ with the other. The estimates work out just as before.

This completes the proof of Hanges' result.

11.2 A More Complicated Example

We consider

$$P(t,x,D_t,D_x) = D_t^2 + [x_1D_2 - x_2D_1 + t^k(x_1D_1 + x_2D_2)]^2 = X_1^2 + X_2^2, \quad (11.3)$$

for some $k \geq 2$ in the region $0 < r_1 \leq r \leq r_2,\ r = |x|,\ x = (x_1,\ x_2) \in \mathbb{R}^2$.

The so-called Poisson–Treves stratification for P begins with the characteristic manifold $\Sigma_1 = \Sigma_1' \cup \Sigma_1''$ and is given by

$$\begin{aligned}
\Sigma_{1'} &= \{\tau = 0, x_1\xi_2 - x_2\xi_1 + t^k(x_1\xi_1 + x_2\xi_2)) = 0, t \neq 0\};\\
\Sigma_{1''} &= \{\tau = 0, t = 0, x_1\xi_2 - x_2\xi_1 = 0, \xi = (\xi_1,\xi_2) \neq 0\}\\
\Sigma_j &= \{\tau = 0, x_1\xi_2 - x_2\xi_1 = 0, \}, \quad \text{for } 1 < j \leq k,\\
\Sigma_{k+1} &= \{0\}, \qquad (11.4)
\end{aligned}$$

where the last equation means that Σ_{k+1} is just the zero section of $T^*\mathbb{R}^3$.

The stratification consists of (1) the characteristic manifold, which is the common vanishing set of the symbols of the vector fields, (2) the common vanishing set of the first "layer" above with the zeros of the symbols of the first brackets of the vector fields, (3) the common vanishing set of the second "layer" above with the zeros of the symbols of the second brackets of the vector fields, etc.

(For the sake of simplicity we ignore here that Σ_1 has two connected components: $\Sigma_1 = \Sigma_{1,-} \cup \Sigma_{1,+}$, according to $t \lessgtr 0$.)

We explicitly remark that Σ_1' is a symplectic submanifold of codimension 2, while Σ_1'' is not symplectic. And $\Sigma_1' \cup \Sigma_1'' = \text{CharP}$.

Let us define

$$X_1 = D_t \qquad (11.5)$$

$$\begin{aligned}
X_2 &= x_1D_2 - x_2D_1 + t^k(x_1D_1 + x_2D_2)\\
&= D_\theta + t^kD_r = D_\theta + R, \qquad (11.6)
\end{aligned}$$

so that

$$P = X_1^2 + X_2^2 \tag{11.7}$$

with

$$[X_1, X_2] = kt^{k-1}R \quad \text{and } [R, X_j] = 0, \quad j = 1, 2. \tag{11.8}$$

We have the a priori estimate

$$\|v\|^2_{\frac{1}{k+1}} + \|X_1 v\|_0^2 + \|X_2 v\|_0^2 \le C\left\{|\langle Pv, v\rangle| + \|v\|_0^2\right\}, \qquad v \in C_0^\infty, \tag{11.9}$$

where $\|\cdot\|_0$ denotes the L^2-norm in $\mathbb{R}_t \times \mathbb{R}_x^2$.

Theorem 11.1. *P is analytic hypoelliptic in any open set of the form $\Omega = \{(t, x) \in \mathbb{R}^3 \mid r_1 < |x| < r_2, t \in (-\delta, \delta)\}$, $\delta > 0$.*

11.3 The General Scheme

Since the operator is subelliptic, the solution will be in C^∞. Additionally, since for $t \neq 0$, the characteristic manifold of P is symplectic, we know that the solution is analytic for $t \neq 0$.

With a localizing function $\varphi(r)$ to be made precise below (but of Ehrenpreis type), and exploiting the maximality of the a priori estimate satisfied by P, we will study $\|\varphi X^{p+1}u\|^2$, each occurance of X being X_1 or X_2. Using the a priori estimate effectively will require moving one X to the left of φ, but this will not present a problem in the ensuing recurrence.

We will immediately be led to estimate the bracket

$$|\langle [P, \varphi X^p]u, \varphi X^p u\rangle|.$$

Upon iteration, using (11.8), we arrive, after at most p iterations of the a priori estimate, at terms of the form

$$\underline{C^p}\|(X)\varphi^{(p+1)}u\|_0^2 \quad \text{or} \quad \underline{C^p}Cp!!\|(X)\varphi R^{p/2}u\|_0^2, \tag{11.10}$$

and of course all the intermediate terms with some derivatives on φ, some powers of p, and some powers of R, all with the generic bounds

$$CC^p p^a \|(X)\varphi^{(b)} R^a u\| \quad \text{with } p \sim b + 2a.$$

Here $p!! = p(p-2)(p-4)\ldots$, the value of C may change from line to line but always independently of u and the order of differentiation, and underlining a coefficient indicates the number of terms of the form that follows that occur. Finally, writing (X) means that an $X = X_1$ or X_2 may or may not be present.

When all X's have been consumed in this way, we may no longer iterate effectively, and we must turn our attention to pure powers of R, suitably localized.

This will require a new localizing function and a construction we denote by $(R^q)_\psi$, reminiscent of [T4] and [T5], or more precisely [DT1], which requires a special vector field M that commutes especially well with both X_1 and X_2, namely reproducing X_1 or generating R.

11.4 The Vector Field M and the Localization

We are fortunate to have a "good" vector field M at our disposal that reproduces R by bracketing that X_2: with

$$M = \frac{t}{k} D_t, \tag{11.11}$$

we have

$$[M, X_2] = R. \tag{11.12}$$

As localizing functions we shall use a nested family of Ehrenpreis-type functions as used by as we have used before in [T4], [T5]. Given $N \in \mathbb{N}$, the band $r \in (r_1, r_2)$ will contain $\log_2 N$ nested subbands, $\Omega_k = \{r : r_{1_k} \le r \le r_{2_k}\}$, $r_{1_0} = r_1, r_{2_0} = r_2$, $k \le \log_2 N$, with

$$d_k = r_{1_k} - r_{1_{k-1}} = r_{2_k} - r_{2_{k-1}} = (r_2 - r_1)\frac{1}{4k^2} \tag{11.13}$$

(so that $\sum d_k \le r_2 - r_1$) and functions $\varphi_k \equiv 1$ on Ω_k and supported in Ω_{k+1}, such that with a constant $C = C_{r_2 - r_1}$,

$$|\varphi_k^{(\ell)}(r)| \le (C/d_k)^{\ell+1} N_k^\ell \quad \text{for } \ell \le N_k = N/2^{k-1}. \tag{11.14}$$

The functions φ_k, but not the constant C, depend on the choice of N. In fact, we shall double the number of these functions, for technical reasons, $\varphi_1, \tilde{\varphi}_1, \varphi_2, \tilde{\varphi}_2, \ldots$ with φ_j and $\tilde{\varphi}_j$ satisfying the same growth estimates.

Given the definition of M above, and for $p \in \mathbb{R}$ large and $j \in \{0, 1, \ldots, p\}$, we define the expressions

$$N_j = \sum_{j'=0}^{j} a_{j'}^j \frac{M^{j'}}{j'!}, \tag{11.15}$$

where the $a_{j'}^j$ denote rational numbers satisfying properties, which shall be made precise below, that optimize commutation relations.

Finally, we define our localizing operator, which is equal to R^p where $\varphi \equiv 1$. We let

$$R^p_\varphi = \sum_{j=0}^{p} \varphi^{(j)} N_j R^{p-j} = \sum_{j=0}^{p} (R^j \varphi) N_j R^{p-j}. \tag{11.16}$$

11.5 The Commutation Relations for R^p_φ

For two vector fields Z and $\tilde{Z}$ we shall frequently use the formula

$$[Z^j, \tilde{Z}] = \sum_{k=1}^{j} \binom{j}{k} \mathrm{ad}_Z^k(\tilde{Z}) Z^{j-k}, \tag{11.17}$$

where

$$\mathrm{ad}_Z(\tilde{Z}) = [Z, \tilde{Z}], \quad \mathrm{ad}_Z^2(\tilde{Z}) = [Z, [Z, \tilde{Z}]],$$

and so forth.

11.6 The Bracket $[X_2, R^p_\varphi]$

We first compute the commutator of X_2 with N_j. We have

$$[X_2, N_j] = \sum_{j'=1}^{j} a^j_{j'} \left[X_2, \frac{M^{j'}}{j'!} \right] = -R^k \sum_{j'=1}^{j} a^j_{j'} \sum_{\ell=1}^{j'} \frac{1}{\ell!} \frac{M^{j'-\ell}}{(j'-\ell)!} R,$$

since

$$[X_2, M] = -t^k R.$$

We seek to find coefficients $a^j_{j'}$ such that

$$[X_2, N_j] = -t^k N_{j-1} R,$$

which will ensure that the bracket $[X_2, R^p_\varphi]$ is free of the (poorly controlled) vector field R (see below). Using (11.10), the necessary condition is that the $a^j_{j'}$ must satisfy

$$\sum_{j'=1}^{j} \sum_{\ell=1}^{j'} a^j_{j'} \frac{1}{\ell!} \frac{M^{j'-\ell}}{(j'-\ell)!} = \sum_{j_1=1}^{j-1} a^{j-1}_{j_1} \frac{M^{j_1}}{j_1!}, \tag{11.18}$$

or

$$\sum_{s=1}^{j-\ell} a^{j}_{\ell+s} \frac{1}{s!} = a^{j-1}_{\ell}, \tag{11.19}$$

for $\ell = 0, 1, \dots, j-1$.

We shall come back to condition (11.19) later; for the time being we may conclude the following:

Lemma 11.1. *With the coefficients $a^{j}_{j'}$ chosen as above,*

$$[X_2, N_j] = -t^k N_{j-1} R,$$

for every $j \in \mathbb{N}$.

Proposition 11.1.

$$[X_2, R^p_\varphi] = t^k \varphi^{(p+1)} N_p.$$

Proof. Using the above lemma, we have, for these $a^{j}_{j'}$,

$$\begin{aligned}
[X_2, R^p_\varphi] &= t^k \sum_{j=0}^{p} \varphi^{(j+1)} N_j R^{p-j} - t^k \sum_{j=1}^{p} \varphi^{(j)} N_{j-1} R^{p+1-j} \\
&= t^k \varphi^{(p+1)} N_p + t^k \sum_{j=1}^{p} \varphi^{(j)} N_{j-1} R^{p-(j+1)} - t^k \sum_{j=1}^{p} \varphi^{(j)} N_{j-1} R^{p+1-j} \\
&= t^k \varphi^{(p+1)} N_p.
\end{aligned}$$

□

11.7 The Bracket $[X_1, R^p_\varphi]$

First observe that

$$\operatorname{ad}^{\ell}_{M}(X_1) = \left(-\frac{1}{k}\right)^{\ell} X_1, \tag{11.20}$$

for every $\ell \in N$.

Therefore,

$$[X_1, N_j] = \sum_{j'=1}^{j} a^{j}_{j'} \left[X_1, \frac{M^{j'}}{j'!} \right] = -X_1 \sum_{j'=1}^{j} \sum_{\ell=1}^{j'} a^{j}_{j'} \left(-\frac{1}{k}\right)^{\ell} \frac{1}{\ell!} \frac{M^{j'-\ell}}{(j'-\ell)!},$$

so that we have

$$[X_1, R^p_\varphi] = \sum_{j=1}^{p} \varphi^{(j)} (-X_1) \sum_{j'=1}^{j} \sum_{\ell=1}^{j'} a^{j}_{j'} \left(-\frac{1}{k}\right)^{\ell} \frac{1}{\ell!} \frac{M^{j'-\ell}}{(j'-\ell)!} R^{p-j}. \tag{11.21}$$

Our next goal is to prove the following lemma:

Lemma 11.2. *For every $j \in \mathbb{N}$ and $\ell \in \{1, \dots, j\}$ there exist real constants δ_s, $s = 0, \dots, j-2$, such that*

$$\sum_{h=1}^{j-\ell} a^j_{\ell+h} \left(-\frac{1}{k}\right)^h \frac{1}{h!} = \sum_{h=1}^{j-\ell} \delta_{j-\ell-h} a^{\ell+h-1}_\ell. \tag{11.22}$$

The above lemma has an easy consequence:

Lemma 11.3. *For every $j \in \mathbb{N}$ there exist real constants γ_s, with $s = 0, \dots, j-1$, such that*

$$\sum_{j'=1}^{j} \sum_{\ell=1}^{j'} a^j_{j'} \left(-\frac{1}{k}\right)^\ell \frac{1}{\ell!} \frac{M^{j'-\ell}}{(j'-\ell)!} = \sum_{s=0}^{j-1} \gamma_{j-s} N_s. \tag{11.23}$$

Proof of Lemma 11.3. The identity (11.23) can be restated as

$$\sum_{s=0}^{j-1} \sum_{j'=s+1}^{j} a^j_{j'} \left(-\frac{1}{k}\right)^{j'-s} \frac{1}{(j'-s)!} \frac{M^s}{s!} = \sum_{s=0}^{j-1} \sum_{h=0}^{s} \gamma_{j-s} a^s_h \frac{M^h}{h!},$$

or

$$\sum_{s=0}^{j-1} \sum_{j'=s+1}^{j} a^j_{j'} \left(-\frac{1}{k}\right)^{j'-s} \frac{1}{(j'-s)!} \frac{M^s}{s!} = \sum_{\ell=0}^{j-1} \sum_{s=\ell}^{j-1} \gamma_{j-s} a^s_\ell \frac{M^\ell}{\ell!},$$

from which we get

$$\sum_{j'=\ell+1}^{j} a^j_{j'} \left(-\frac{1}{k}\right)^{j'-\ell} \frac{1}{(j'-\ell)!} = \sum_{s=\ell}^{j-1} \gamma_{j-s} a^s_\ell.$$

Now the latter identity can be rewritten as

$$\sum_{h=1}^{j-\ell} a^j_{\ell+h} \left(-\frac{1}{k}\right)^h \frac{1}{h!} = \sum_{h=1}^{j-\ell} \delta_{j-\ell-h} a^{\ell+h-1}_\ell,$$

for any $\ell = 1, \dots, j-1$, and this is in the statement of Lemma 11.2. □

Proof of Lemma 11.2. In order to prove Lemma 11.2 we must analyze the recurrence relation (11.19):

$$\sum_{h=1}^{j-\ell} a^j_{\ell+h} \frac{1}{h!} = a^{j-1}_\ell,$$

for $\ell = 0, 1, \dots, j-1$.

Another way of rewriting the above relation is the following:

$$\begin{bmatrix} 1 & \frac{1}{2!} & \cdots & \frac{1}{(j-1)!} & \frac{1}{j!} \\ 0 & 1 & \cdots & \frac{1}{(j-2)!} & \frac{1}{(j-1)!} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \frac{1}{2!} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1^j \\ \vdots \\ a_j^j \end{bmatrix} = \begin{bmatrix} a_0^{j-1} \\ \vdots \\ a_{j-1}^{j-1} \end{bmatrix}. \tag{11.24}$$

Note that on the left-hand side there are no terms of the form a_0^j, which means that we are free to choose those coefficients. We shall choose $a_0^0 = 1$ for the sake of simplicity, leaving the others undetermined.

We point out that the matrix in the above formula is clearly invertible and that it can be written as

$$I_j + \frac{1}{2!} J_j + \frac{1}{3!} J_j^2 + \cdots + \frac{1}{j!} J_j^{j-1} = \int_0^1 e^{tJ_j}\, dt,$$

where J_j denotes the standard $j \times j$ Jordan matrix

$$q_j = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Using, for example, formula (11.24), we may easily see that by inverting the matrix, we obtain

$$\begin{bmatrix} a_1^j \\ \vdots \\ a_j^j \end{bmatrix} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{j-2} & c_{j-1} \\ 0 & c_0 & \cdots & c_{j-3} & c_{j-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_0 & c_1 \\ 0 & 0 & \cdots & 0 & c_0 \end{bmatrix} \begin{bmatrix} a_0^{j-1} \\ \vdots \\ a_{j-1}^{j-1} \end{bmatrix}, \tag{11.25}$$

where $c_0 = 1$, $c_1 = -\frac{1}{2!}$, and the other c_m can be computed by a triangular relation. In particular, using the structure of the matrix, we obtain that

$$\begin{bmatrix} a^j_{\ell+1} \\ \vdots \\ a^j_j \end{bmatrix} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{j-\ell-2} & c_{j-\ell-1} \\ 0 & c_0 & \cdots & c_{j-\ell-3} & c_{j-\ell-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & c_0 & c_1 \\ 0 & 0 & \cdots & 0 & c_0 \end{bmatrix} \begin{bmatrix} a^{j-1}_\ell \\ \vdots \\ a^{j-1}_{j-1} \end{bmatrix}, \tag{11.26}$$

for $\ell = 0, 1, \dots, j-1$.

Another way of writing the above identity is

$$a^j_{\ell+h} = \sum_{s=h}^{j-\ell} c_{s-h} a^{j-1}_{\ell-1+s}, \tag{11.27}$$

for $h = 1, 2, \dots, j-\ell$.

Iterating, we get

$$a^{j-t}_{\ell-t+h} = \sum_{s=h}^{j-\ell} c_{s-h} a^{j-t-1}_{\ell-t-1+s},$$

for $t = 0, 1, \dots, j-\ell-1$.

Let us now fix an $h \in \{1, \dots, j-\ell\}$. Then we have

$$\begin{aligned} a^j_{\ell+h} &= \sum_{s_1=h}^{j-\ell} c_{s_1-h} a^{j-1}_{\ell-1+s} \\ &= \sum_{s_1=h}^{j-\ell} \sum_{s_2=s_1}^{j-\ell} c_{s_1-h} c_{s_2-s_1} a^{j-2}_{\ell-2+s_2} \\ &= \sum_{s_1=h}^{j-\ell} \sum_{s_2=s_1}^{j-\ell} \cdots \sum_{s_h=s_{h-1}}^{j-\ell} c_{s_1-h} c_{s_2-s_1} \cdots c_{s_h-s_{h-1}} a^{j-h}_{\ell-h+s_h}. \end{aligned}$$

The latter sum can be written as

$$\begin{aligned} a^j_{\ell+h} &= c_0^h a^{j-h}_\ell + \sum_{s_1=h}^{j-\ell} \sum_{s_2=s_1}^{j-\ell} \cdots \sum_{\substack{s_h=s_{h-1} \\ s_h>h}}^{j-\ell} c_{s_1-h} c_{s_2-s_1} \cdots c_{s_h-s_{h-1}} a^{j-h}_{\ell-h+s_h} \\ &= c_0^h a^{j-h}_\ell + \sum_{s_1=h}^{j-\ell} \sum_{s_2=s_1}^{j-\ell} \cdots \sum_{\substack{s_h=s_{h-1} \\ s_h>h}}^{j-\ell} \sum_{s_{h+1}=s_h}^{j-\ell} c_{s_1-h} \cdots c_{s_{h+1}-s_h} a^{j-h-1}_{\ell-h-1+s_{h+1}}. \end{aligned}$$

The latter sum allows us to compute the coefficient of a_ℓ^{j-h-1}, by picking all terms for which one of the s_j is equal to $h+1$, for $j = 1, 2, \dots, h+1$.

Iterating this procedure, i.e., using the recurrence relation until we obtain a coefficient a_*^* where the lower index is equal to ℓ, we may express the coefficient $a_{\ell+h}^j$ as a linear combination of a_ℓ^*; the above formulas show that we may actually write

$$a_{\ell+h}^j = \sum_{\sigma=0}^{j-\ell-h} \alpha_{j-\ell-\sigma} a_\ell^{\ell+\sigma}, \tag{11.28}$$

for $h = 1, 2, \dots, j-\ell$.

We point out explicitly that up to this point we have used only the recurrence relation (11.19). Let us now denote by A_h the collection of real numbers $A_h = (-k^{-1})^h h!^{-1}$. Then it is evident that

$$\sum_{h=1}^{j-\ell} a_{\ell+h}^j A_h = \sum_{\sigma=0}^{j-\ell-1} \delta_{j-\ell-\sigma} a_\ell^{\ell+\sigma},$$

where $\delta_{j-\ell-\sigma} = \sum_{h=1}^{j-\ell-\sigma} A_h$, and this is the statement of the lemma. □

Lemma 11.4 (See [DT1] and [Hi]). *Let us consider the recurrence relation (11.19):*

$$\sum_{s=1}^{j-\ell} a_{\ell+s}^j \frac{1}{s!} = a_\ell^{j-1}.$$

Setting

$$a_\ell^j = \frac{1}{(j-\ell)!} \left(\left[\frac{t}{e^t - 1} \right]^{j+1} \right)^{(j-\ell)} (0), \tag{11.29}$$

we obtain a solution of the above recurrence satisfying the boundary conditions $a_j^j = 1$ and $a_0^j = (-1)^j$, $j \geq 0$. Moreover, this is the only power series with rational coefficients satisfying (11.19) and the above boundary conditions.

Proof. By a simple computation we have

$$\begin{aligned} a_\ell^{j-1} &= \frac{1}{(j-1-\ell)!} \left(\left(\frac{t}{e^t-1} \right)^j \right)^{(j-1-\ell)} (0) \\ &= \frac{1}{(j-1-\ell)!} \left(\left(\frac{t}{e^t-1} \right)^{j+1} \frac{e^t-1}{t} \right)^{(j-1-\ell)} (0) \\ &= \sum_{h=0}^{j-1-\ell} \frac{1}{(j-1-\ell-h)!} \left(\left(\frac{t}{e^t-1} \right)^{j+1} \right)^{(j-1-\ell-h)} (0) \frac{1}{(h+1)!} \end{aligned}$$

$$= \sum_{p=1}^{j-\ell} \frac{1}{(j-1-p)!} \left(\left(\frac{t}{e^t-1}\right)^{j+1}\right)^{(j-\ell-p)} (0) \frac{1}{p!}$$

$$= \sum_{p=1}^{j-\ell} a^j_{\ell+p} \frac{1}{p!},$$

where we used the fact that

$$\frac{1}{h!} \left(\frac{e^t-1}{t}\right)^{(h)} (0) = \frac{1}{(h+1)!}.$$

Moreover, we have $a^j_j = 1$ for every $j \geq 0$. As for the other boundary condition, first we remark that

$$a^j_0 = \frac{1}{j!} \left(\left[\frac{t}{e^t-1}\right]^{j+1}\right)^{(j)} (0),$$

i.e., a^j_0 is the coefficient of t^j in the power series of $Q(t)$,

$$Q(t) = \left(\frac{t}{e^t-1}\right)^{j+1}.$$

Thus

$$a^j_0 = \frac{1}{2i\pi} \int_\gamma \left(\frac{1}{e^z-1}\right)^{j+1} dz,$$

where γ is a smooth curve encircling the origin in $\mathbb{C}$.

Changing variables $w = e^z - 1$, so that the origin is mapped to the origin and γ is mapped to another smooth curve encircling the origin that we still denote by γ, we have

$$a^j_0 = \frac{1}{2i\pi} \int_\gamma w^{-(j+1)} (w+1)^{-1} dz = (-1)^j.$$

The uniqueness is proved in [Hi]. This ends the proof of the lemma. □

As a consequence of the preceding lemmas we may now state the following result:

Proposition 11.2. *The commutator of X_1 with the localizing operator R^p_φ has the form*

$$[X_1, R^p_\varphi] = -X_1 \sum_{\ell=0}^{p-1} \delta_\ell R^{p-\ell-1}_{\varphi^{(\ell+1)}}, \tag{11.30}$$

where $\varphi^{(j)}$ denotes the jth derivative $(r\partial_r)^j \varphi$ and $\delta_\ell = \sum_1^\ell \frac{1}{k^h h!} \leq 1$.

From Lemma 11.4 we have the following corollary:

Corollary 11.1. *For every* $j \geq 0$ *and* $\ell \in \{0, \dots, j\}$ *we have*

$$|a_\ell^j| \leq c^j,$$

for a suitable universal positive constant c.

Proof. From (11.29) we have that

$$a_\ell^j = \frac{1}{2i\pi} \int_\gamma \left(\frac{t}{e^t - 1} \right)^{j+1} t^{-(j-\ell+1)} \, dt,$$

where γ is a circle of fixed radius around the origin. Since the function under the integral sign may be estimated by a positive constant (depending on the radius of γ) raised to the power j, the corollary follows. □

11.8 Proof of Theorem 11.1

In this section we prove that P is analytic hypoelliptic in any open set of the form $\Omega = \{(t, x) \in \mathbb{R}^3 \mid r_1 < |x| < r_2, t \in (-\delta, \delta)\},\ \delta > 0$.

The maximal estimate may be restated to allow X to appear to the right or left of the localizing function (where $\psi' = X\psi$):

$$\begin{aligned}
&\|\psi X^p u\|^2_{\frac{1}{k+1}} + \|X\psi X^p u\|_0^2 + \|\psi X^{p+1} u\|_0^2 \\
&\quad \lesssim |\langle P\psi X^p u, \psi X^p u\rangle| + \|\psi' X^p u\|_0^2 \\
&\quad \lesssim |\langle \psi X^p P u, \psi X^p u\rangle| + |\langle [X^2, \psi X^p] u, \psi X^p u\rangle| + \|\psi' X^p u\|_0^2. \qquad (11.31)
\end{aligned}$$

Now

$$\begin{aligned}
|\langle [X^2, \psi X^p] u, \psi X^p u\rangle| &\leq |\langle [X, \psi X^p] u, X\psi X^p u\rangle| + |\langle [X, \psi X^p] X u, \psi X^p u\rangle| \\
&\lesssim |\langle \psi' X^p u, X\psi X^p u\rangle| + \underline{p} |\langle t^{k-1} \psi R X^{p-1} u, X\psi X^p u\rangle| \\
&\quad + |\langle \psi X^p X u, \psi' X^p u\rangle| + \underline{p} |\langle t^{k-1} \psi X^p R u, \psi X^p u\rangle| \\
&\leq \varepsilon \left\{ \|X\psi X^p u\|_0^2 + \|\psi X^{p+1} u\|_0^2 \right\} \\
&\quad + C_\varepsilon \left\{ \|\psi' X^p u\|_0^2 + (\underline{p} \|\psi X^{p-1} R u\|_0)^2 \right\},
\end{aligned}$$

where we have freely exchanged ψ and ψ' on the two sides of the inner product when no derivatives intervened. Note that $[X, R] = 0$.

In all,

$$\|\psi X^p u\|^2_{\frac{1}{k+1}} + \|X\psi X^p u\|_0^2 + \|\psi X^{p+1}u\|_0^2$$
$$\lesssim \|\psi X^p P u\|_0^2 + \|\psi' X^p u\|_0^2 + (\underline{p}\|\psi R X^{p-1} u\|_0)^2.$$

Here the L^2 norms on the left-hand side (the "maximal" part of the estimate as opposed to the subelliptic part) are the principal ones. And (11.32) has the effect of reducing the number of X's by one while adding a derivative to the localizing function or reducing the number of X's by two, introducing a factor of at most a multiple of p and the operator R, which commutes with P and X.

Iterating this process as long as at least one X remains to profit from the maximal portion of the estimate yields

$$\|\psi X^p u\|^2_{\frac{1}{k+1}} + \|X\psi X^p u\|_0^2 + \|\psi X^{p+1}u\|_0^2$$
$$\leq C^p \left\{ \sup_{j+2d\leq p} \{p^d \|\psi^{(j)} R^d X^{p-j-2d} P u\|_0^2\} + \sup_{j+2d=p} (p^d \|\psi^{(j)} R^d u\|_0)^2 \right\}. \tag{11.32}$$

The last term occurs as a sum, though with a larger constant C we may replace the sum with a supremum, and the expression $j + 2d = p$ reflects that either the X's that were present became derivatives on ψ or a pair of them yielded a factor of P and an R.

The first term on the right can be estimated directly (even taken to be zero, using the Cauchy–Kovalevskaya theorem). For the second, we will take the localizing function out of the norm and introduce one of the $(R^d)_{\tilde{\psi}} \equiv R^d$ on the support of ψ. Thus for such $\tilde{\psi}$, and taking $Pu = 0$ for simplicity,

$$\|\psi X^p u\|_{\frac{1}{k+1}} + \|X\psi X^p u\|_0 + \|\psi X^{p+1}u\|_0 \leq \sup_{j+2d=p} p^d \sup|\psi^{(j)}| \|(R^d)_{\tilde{\psi}} u\|_0. \tag{11.33}$$

For convenience we recall the bracket relations and a few important definitions (for generic φ):

$$\left[X_1, (R^b)_\varphi\right] = -X_1 \sum_{\ell=0}^{b-1} \delta_\ell \left(R^{b-\ell-1}\right)_{\varphi^{(\ell+1)}}, \quad |\delta_\ell| \leq 1,$$

$$\left[X_2, (R^b)_\varphi\right] = t^k \varphi^{(b+1)} N_b,$$

$$N_b = \sum_{b'=0}^{b} a^b_{b'} \frac{M^{b'}}{b'!}, \qquad M = \frac{t}{k} D_t, \qquad |a^b_{b'}| \leq c^b.$$

As above, we use the a priori estimate, but now on $v = (R^d)_\varphi u$:

$$\|(R^d)_\varphi u\|^2_{\frac{1}{k+1}} + \|X(R^d)_\varphi u\|_0^2 + \|(R^d)_\varphi X u\|_0^2$$
$$\lesssim |\langle P(R^d)_\varphi u, (R^d)_\varphi u\rangle| + \|[X, (R^d)_\varphi]u\|_0^2$$
$$\lesssim |\langle (R^d)_\varphi P u, (R^d)_\varphi u\rangle| + |\langle [X^2, (R^d)_\varphi]u, (R^d)_\varphi u\rangle| + \|[X, (R^d)_\varphi]u\|_0^2. \tag{11.34}$$

Again, taking $Pu = 0$, and expanding $[X^2, (R^d)_\varphi] = X[X, (R^d)_\varphi] + [X, (R^d)_\varphi]X$, we obtain, as before, with a weighted Schwarz inequality and integrating by parts since $X = -X^*$,

$$\|(R^d)_\varphi u\|^2_{\frac{1}{k+1}} + \|X(R^d)_\varphi u\|_0^2 + \|(R^d)_\varphi X u\|_0^2$$
$$\lesssim |\langle [X, (R^d)_\varphi] X u, (R^d)_\varphi u\rangle| + \|[X, (R^d)_\varphi]u\|_0^2. \tag{11.35}$$

Now on the right, when $X = X_1$, the result, as we saw above, still has an X_1, which we integrate by parts in the case of the inner product:

$$|\langle [X_1, (R^d)_\varphi] X_1 u, (R^d)_\varphi u\rangle| + \|[X_1, (R^d)_\varphi]u\|_0^2$$
$$\leq \varepsilon \|X_1 (R^d)_\varphi u\|_0^2 + C_\varepsilon \sum_{d_1=1}^{d} \|(R^{d-d_1})_{\varphi^{(d_1)}} X_1 u\|_0^2. \tag{11.36}$$

On the other hand, when $X = X_2$, we have nearly pure powers of tD_t, which it will be necessary to convert into pure powers of $X_1 = D_t$ (from which we started, but, we note, of at most half the order).

Proposition 11.3.

$$(tD_t)^j = \sum_{\ell=1}^{j} B_\ell^j \; t^\ell D_t^\ell,$$

where

$$B_\ell^j = \sum_{m=0}^{\ell-1} \frac{(-1)^m (\ell - m)^{j-1}}{m!(\ell - m - 1)!} = \sum_{m=0}^{\ell-1} \frac{(-1)^m (\ell - m)^j}{m!(\ell - m)!},$$

so that for all v, pointwise,

$$\frac{|(tD_t)^j v|}{j!} \leq C^j \sum_{\ell=1}^{j} \frac{|t^\ell D_t^\ell v|}{\ell!}$$

and hence in $|t| < 1$,

$$|N_b v| = \left| \sum_{b'=0}^{b} a_{b'}^b \frac{M^{b'}}{b'!} v \right| = \left| \sum_{b'=0}^{b} \left(\frac{1}{k}\right)^{b'} a_{b'}^b \frac{(tD_t)^{b'}}{b'!} v \right| \leq C^b \sup_{b' \leq b} \left| \frac{X_1^{b'} v}{b'!} \right|.$$

This particular expression for the coefficients B_ℓ^j is proved by induction and can be understood by a kind of overcounting/undercounting argument. One pleasing interpretation for the quantities $B_\ell^j\, j\,!$ is the number of ways to color j objects with ℓ colors, ensuring that all the colors are used. Thus, we have an alternative expression for B_ℓ^j:

$$B_\ell^j = \binom{j}{\ell} \ell^{j-\ell},$$

which is also easily proved by induction.

Thus for X_2,

$$\begin{aligned}|\langle [X_2, (R^d)_\varphi] X_2 u, (R^d)_\varphi u\rangle| &+ \|[X_2, (R^d)_\varphi] u\|_0^2 \\ &\leq |\langle t^k \varphi^{(d+1)} N_d X_2 u, (R^d)_\varphi u\rangle| + \|t^k \varphi^{(d+1)} N_d u\|_0^2 \\ &\leq C^d \sup_{d' \leq d} \left\| \varphi^{(d+1)} \frac{X_1^{d'} (X_2) u}{d'!} \right\|_0^2 + \varepsilon \|(R^d)_\varphi u\|_0^2, \end{aligned} \tag{11.37}$$

or in all,

$$\begin{aligned}\left\|(R^d)_\varphi u\right\|_{\frac{1}{k+1}}^2 &+ \|X (R^d)_\varphi u\|_0^2 + \|(R^d)_\varphi X u\|_0^2 \\ &\leq C \sum_{d_1=1}^{d} \|(R^{d-d_1})_{\varphi^{(d_1)}} X_1 u\|_0^2 + C^d \sup_{d' \leq d} \left\| \varphi^{(d+1)} \frac{X_1^{d'} (X_2) u}{d'!} \right\|_0^2 \end{aligned} \tag{11.38}$$

(i.e., with or without X_2 in the last term). Iterating on the first term on the right, eventually only the last term survives:

$$\begin{aligned}\left\|(R^d)_\varphi u\right\|_{\frac{1}{k+1}}^2 &+ \|X (R^d)_\varphi u\|_0^2 + \|(R^d)_\varphi X u\|_0^2 \\ &\leq C^d \sup |\varphi^{(d+1)}| \sup_{d' \leq d} \left\| \tilde{\varphi} \frac{X^{d'+1} u}{(d' + 1!)} \right\|_0^2 \end{aligned} \tag{11.39}$$

for any $\varphi, \tilde{\varphi}$ with $\tilde{\varphi} \equiv 1$ on the support of φ.

Recalling the previous bound

$$\begin{aligned}\left\|\psi X^p u\right\|_{\frac{1}{k+1}}^2 &+ \|X \psi X^p u\|_0^2 + \|\psi X^{p+1} u\|_0^2 \\ &\leq \sup_{j+2d=p} p^d \sup |\psi^{(j)}| \|(R^d)_{\tilde{\psi}} u\|_0^2, \end{aligned} \tag{11.40}$$

valid for any $\tilde{\psi} \equiv 1$ on the support of ψ, we have the choice of starting with X's, reducing the order by half, introducing $(R^d)_\varphi$ and iterating that until we are back to X's, or starting with $(R^d)_\varphi$, reducing to X's until they bracket to yield pure R's at half the order.

In either order, after one full cycle, we need a new localizing function each time $(R^d)_\varphi$ is put together. Thus in starting with N derivatives to estimate, after $\log_2 N$ full cycles, the number of free derivatives on u will be only a bounded number.

For definiteness, we follow the cycle starting with powers of X's, and introduce for a moment the new norms

$$|||\psi, X^p, u||| = \left\|\psi X^p u\right\|_{\frac{1}{k+1}} + \|X\psi X^p u\|_0 + \|\psi X^{p+1} u\|_0$$

and

$$|||(R^d)_\varphi, u||| = \left\|(R^d)_\varphi u\right\|_{\frac{1}{k+1}} + \|X(R^d)_\varphi u\|_0 + \|(R^d)_\varphi X u\|_0,$$

so that the above may be written

$$|||(R^d)_\varphi, u||| \le C^d \sup_{d' \le d} \frac{1}{(d'+1)!} |||\varphi^{(d+1)}, X^{d'+1}, u||| \tag{11.41}$$

for any φ and

$$|||\psi, X^p, u||| \le \sup_{j+2d=p} p^d \sup|\psi^{(j)}|\, |||(R^d)_{\tilde{\psi}}, u|||. \tag{11.42}$$

Thus we start with $\psi = \varphi_1$ (the first in the sequence of precisely nested localizing functions (cf. (11.14)) for a fixed $N = N_1 \in \mathbb{N}$:

$$\begin{aligned}
\frac{\|X^{N_1} u\|_{L^2(\Omega_1)}}{N_1!} &\le \frac{|||\varphi_1, X^{N_1}, u|||}{N_1!} \le \sup_{N_2 \le \widetilde{N_1} \le N_1} \frac{|||\varphi_1, X^{\widetilde{N_1}}, u|||}{\widetilde{N_1}!} \\
&\le \sup_{\substack{\ell_1 + 2\delta_1 = \widetilde{N_1} \\ N_2 \le \widetilde{N_1} \le N_1}} \frac{\widetilde{N_1}^{\delta_1} \sup|\varphi_1^{(\ell_1)}|\, |||(R^{\delta_1})_{\tilde{\varphi}}, u|||}{\widetilde{N_1}!} \\
&\le \sup_{\substack{\ell_1 + 2\delta_1 = \widetilde{N_1} \\ N_2 \le \widetilde{N_1} \le N_1}} \frac{\left(\frac{C}{d_1}\right)^{\ell_1} \widetilde{N_1}^{\ell_1+\delta_1} |||(R^{\delta_1})_{\tilde{\varphi}}, u|||}{\widetilde{N_1}!},
\end{aligned}$$

with any $\tilde{\varphi} \equiv 1$ near the support of φ_1.

Now there is some freedom in the choice of $\tilde{\varphi}$, since all that we have required is that it be equal to 1 on the support of φ_1, and we pick the largest index k consistent with the δ_1, k giving the supremum just above, i.e., $N_k \ge \delta_1 \ge N_{k+1}$ and $k \ge 2$, since $\ell_1 + 2\delta_1 = N_1$. Thus with $\tilde{\varphi} = \varphi_k$ and together with the other estimate,

$$|||(R^\delta)_{\varphi_k}, u||| \le C^\delta \sup|\varphi_k^{(\delta+1)}| \sup_{\delta' \le \delta} \frac{|||\widetilde{\varphi_k}, X^{\delta'+1}, u|||}{(\delta'+1)!}, \tag{11.43}$$

we arrive at

$$\frac{\|X^{N_1}u\|^2_{L^2(\Omega_1)}}{{N_1}^{N_1}} \leq \sup_{N_2\leq\widetilde{N_1}\leq N_1} \frac{|||\widetilde{\varphi_1}, X^{\widetilde{N_1}}, u|||}{\widetilde{N_1}^{\widetilde{N_1}}}$$

$$\leq \sup_{\substack{\ell+2\delta=\widetilde{N_1}\\ N_2\leq\widetilde{N_1}\leq N_1+1}} \frac{\left(\frac{C}{d_1}\right)^{\ell} \widetilde{N_1}^{\ell+\delta} C^{\delta} \sup|\varphi_k^{(\delta+1)}|}{\widetilde{N_1}^{\widetilde{N_1}}}$$

$$\times \sup_{N_{k+1}\leq\widetilde{N_k}\leq N_k+1, k\geq 2} \frac{|||\widetilde{\varphi_k}, X^{\widetilde{N_k}}, u|||}{\widetilde{N_k}^{\widetilde{N_k}}}$$

$$\leq \sup_{\substack{\ell+2\delta=\widetilde{N_1}\\ N_2\leq\widetilde{N_1}\leq N_1}} \frac{\left(\frac{C}{d_1}\right)^{\ell} \widetilde{N_1}^{\ell+\delta} C^{\delta} \left(\frac{C}{d_k}\right)^{\delta+1} N_k^{(\delta+1)}}{\widetilde{N_1}^{\widetilde{N_1}}}$$

$$\times \sup_{N_{k+1}\leq\widetilde{N_k}\leq N_k+1, k\geq 2} \frac{|||\widetilde{\varphi_k}, X^{\widetilde{N_k}}, u|||}{\widetilde{N_k}^{\widetilde{N_k}}}$$

$$\leq C^{\widetilde{N_1}} \sup_{\ell+2\delta=\widetilde{N_1}} \frac{d_1^{-\ell} d_k^{-(\delta+1)}}{2^{\ell+\delta}2^{k(\delta+1)}} \sup_{N_{k+1}\leq\widetilde{N_k}\leq N_k+1, k\geq 2} \frac{|||\widetilde{\varphi_k}, X^{\widetilde{N_k}}, u|||}{\widetilde{N_k}^{\widetilde{N_k}}}.$$

Now the expressions in the first supremum increase as k decreases, bounded by $d_1^{-(N_1+1)}/2^{N_1+1}$. Iteration will introduce another coefficient bounded by $C^{N_2}d_2^{-(N_2+1)}/2^{2(N_2+1)}$, then next by $C^{N_3}d_3^{-(N_3+1)}/2^{3(N_3+1)}$. Since

$$\frac{d_k^{-(N_k+1)}}{2^{k(N_k+1)}} = C^{N_k}\frac{(k^2)^{N_k+1}}{(2^k)^{N_k+1}},$$

iteration at most $\log_2 N$ times will lead to a product

$$\Pi_{k=1}^{\log_2 N} C^{N_k}\left(\frac{k^2}{2^k}\right)^{\frac{N}{2^k}+1} \leq (C')^N$$

times a constant depending only on the first few derivatives of u in the largest open set encountered.

This yields the analyticity of u in the smallest open set, since all estimates are uniform in N.

11.9 Nonclosed Bicharacteristics with Nontrivial Limit Set

We want to study a model of the form

$$P(t, x, D_t, D_x) = D_t^2 + X_2^2,$$

where

$$\begin{aligned} X_2 &= g_1(x)g_2(x)[x_1 D_2 - x_2 D_1 + \mu(x_1 D_1 + x_2 D_2) + t^k (x_1 D_1 + x_2 D_2)], \\ g_1(x) &= |x|^2 - a^2, \\ g_2(x) &= b^2 - |x|^2, \end{aligned}$$

with $0 < a < b$ in the open set $a < |x| < b$ and $\mu > 0$ is a given constant.

The characteristic set of P in the above-mentioned region is

$$\mathrm{Char(P)} = \{\tau = 0,\ \mathrm{x}_1\xi_2 - \mathrm{x}_2\xi_1 + \mu(\mathrm{x}_1\xi_1 + \mathrm{x}_2\xi_2) + \mathrm{t}^{\mathrm{k}}(\mathrm{x}_1\xi_1 + \mathrm{x}_2\xi_2) = 0\}.$$

As for the Poisson stratification of P, we have

$$\begin{aligned} \Sigma_1 &= \{\tau = 0,\ x_1\xi_2 - x_2\xi_1 + \mu(x_1\xi_1 + x_2\xi_2) + t^k(x_1\xi_1 + x_2\xi_2) = 0, t \neq 0\}; \\ \Sigma_2 &= \{\tau = t = 0,\ x_1\xi_2 - x_2\xi_1 + \mu(x_1\xi_1 + x_2\xi_2) = 0, x_1\xi_1 + x_2\xi_2 \neq 0\}; \\ \Sigma_j &= \Sigma_2, \quad j \leq k; \\ \Sigma_{k+1} &= \{0\}, \end{aligned}$$

i.e. the zero section of $T^*\mathbb{R}^3$ over the above specified region.

Evidently, since codim $\Sigma_2 = 3$, Σ_2 (or rather each of its connected components) is not a symplectic submanifold of $T^*\mathbb{R}^3$.

Let us take a look at the Hamilton foliation of Σ_2. Define

$$A = \begin{bmatrix} \mu & 1 \\ -1 & \mu \end{bmatrix}, \tag{11.44}$$

so that

$$\Sigma_2 = \{\tau = 0 = t,\ \langle x, A\xi\rangle = 0\}, \tag{11.45}$$

where $x = (x_1, x_2)$ and $\xi = (\xi_1, \xi_2)$, $\xi \neq 0$, and $\langle x, \xi\rangle \neq 0$.

Then we know that $\langle x, A\xi\rangle \equiv 0$ on every leaf in Σ_2, i.e., on every integral curve of the Hamilton field of $\langle x, A\xi\rangle$ issued from a point in Σ_2.

The Hamilton system is

$$\begin{aligned} \dot{x} &= g_1(x)g_2(x)\, {}^t\!Ax, \\ \dot{\xi} &= -g_1(x)g_2(x)A\xi. \end{aligned} \tag{11.46}$$

We easily see that because of the structure of the matrix A, we have

$$\frac{1}{2}d_t|x|^2 = g_1(x)g_2(x)\frac{\mu}{2}|x|^2,$$
$$\frac{1}{2}d_t|\xi|^2 = -g_1(x)g_2(x)\frac{\mu}{2}|\xi|^2,$$

so that both the spatial and the covariable projections of the bicharacteristics are logarithmic spirals. Moreover, the spatial projection spirals between the two asymptotic circles $g_i(x) = 0$, $i = 1, 2$, which are stationary orbits of the first two equations in (11.46).

We point out that $d_t\langle x, \xi\rangle \equiv 0$, so that $\langle x, \xi\rangle$ is constant along the orbits and once the first two equations in (11.46) are solved, the second couple—i.e., the covariable projection—is easy:

$$\xi(t) = \exp\left[-\int_0^t g_1(x(s))g_2(x(s))ds\ A\right]\xi_0,$$

where ξ_0 is its initial data.

We may apply Theorem 4.2 in [Sj3] to the operator P and conclude that if γ_0 denotes a segment of a bicharacteristic curve in Σ_2, then either $\gamma_0 \subset WF_a(u)$ or $\gamma_0 \cap WF_a(u) = \varnothing$, where u is a solution of $Pu \in C^\omega$ in some open set.

Let now U be an open set in $\mathbb{R}^3$ projecting onto an annulus of the form $a < |x| < b$ in the x-variables. By iteratively applying the above-mentioned theorem one can prove the following:

Theorem 11.2. *Let u be a distribution such that $Pu \in C^\omega(U)$, U being defined as above. Then if both circles $g_i(x) = 0$, $i = 1, 2$, do not intersect $WF_a(u)$, we have that $u \in C^\omega(U)$.*

Chapter 12
Operators of Kohn Type That Lose Derivatives

In [K4], J.J. Kohn proved hypoellipticity for the operator

$$P = LL^* + (\bar{z}^k L)^*(\bar{z}^k L), \qquad L = \frac{\partial}{\partial z} + i\bar{z}\frac{\partial}{\partial t},$$

for which there is a large loss of derivatives; indeed, in the a priori estimate one bounds only the Sobolev norm of order $-(k-1)/2$, and thus there is a loss of $k-1$ derivatives: $Pu \in H^s_{loc} \implies u \in H^{s-(k-1)}_{loc}$.

In this chapter we will show that all solutions of $Pu = f$ with f real analytic are themselves real analytic in any open set where f is. In so doing we use an a priori estimate that follows easily from that established by Kohn for this operator, namely for test functions v of small support near the origin,

$$\|\overline{L}v\|_0^2 + \|\bar{z}^k Lv\|_0^2 + \|v\|^2_{-\frac{k-1}{2}} \lesssim |(Pv, v)_0|. \tag{12.1}$$

In fact, in [T9] (see also [BDKT]), we give a rapid and direct derivation of (12.1) for this operator and similar estimates for more degenerate operators. In Kohn's work, the difficult part is to establish the estimate; indeed, after this is done, with explicit cut-off functions but applied to functions that do not already have small support, the C^∞ hypoellipticity follows at once.

Our approach uses an estimate on test functions, already of small support, and then carefully localizes the actual solution so that it becomes one of these test functions (once we know that the solution may be differentiated).

The first two terms on the left of the estimate (12.1) exhibit maximal control in $\overline{L}$ and $\bar{z}^k L$, but only these complex directions. Hence in obtaining recursive bounds for derivatives it is essential to keep one of these vector fields available for as long as possible.

For this, we will construct a carefully balanced localization of high powers of $T = -2i\partial/\partial t$ and use the estimate repeatedly, reducing the order of powers of T but accumulating derivatives on the localizing functions. These Ehrenpreis-type

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_12, © Springer Science+Business Media, LLC 2011

localizing functions work “as if analytic” up to a prescribed order, with all constants independent of that order, as in [T4], [T5], but eventually the good derivatives ($\overline{L}$ or $\bar{z}^k L$) are lost and we must use the third term on the left of the estimate, absorb the loss of $\frac{k-1}{2}$ derivatives, introduce a new localizing function of larger support, and start the whole process again, but with only a (fixed) fraction of the original power of T.

12.1 Observations and Simplifications

Our first observation is that we know the analyticity of the solution for z different from 0 from our earlier work [T4], [T5] and Treves’ [Tr4]. Thus, modulo brackets with localizing functions whose derivatives are supported in the known analytic hypoelliptic region, we take all localizing functions independent of z.

Our second observation is that it suffices to bound derivatives measured in terms of high powers of the vector fields L and $\overline{L}$ in L^2 norm, by standard arguments, and indeed, estimating high powers of L can be reduced to bounding high powers of $\overline{L}$ and powers of T of half the order, by repeated integration by parts. Thus our overall scheme will be to start with high powers (order $2p$) of L or $\overline{L}$, and use integration by parts and the a priori estimate repeatedly to reduce to treating $T^p u$ in a slightly larger set.

And to do this, we introduce a new special localization of T^p adapted to this problem.

12.2 The Localization of High Powers of *T*

The new localization of T^p may be written in the form

$$(T^{p_1,p_2})_\varphi = \sum_{\substack{a\le p_1\\ b\le p_2}} \frac{L^a \circ z^a \circ T^{p_1-a} \circ \varphi^{(a+b)} \circ T^{p_2-b} \circ \bar{z}^b \circ \overline{L}^b}{a!b!}. \tag{12.2}$$

Here by $\varphi^{(r)}$ we mean $(-i\,\partial/\partial t)^r \varphi(t)$, since near $z = 0$ we have seen that we may take the localizing function independent of z.

Note that the leading term (with $a + b = 0$) is merely $T^{p_1}\varphi T^{p_2}$, which equals $T^{p_1+p_2}$ on the initial open set Ω_0 where $\varphi \equiv 1$. We observe the close connection between this “splitting” of L and $\overline{L}$ to opposite sides of the localizing function and that in Chapter 8 above, and in particular in Definition 9.1.

We have the commutation relations

$$[L, (T^{p_1,p_2})_\varphi] \equiv L \circ (T^{p_1-1,p_2})_{\varphi'},$$

$$[\overline{L}, (T^{p_1,p_2})_\varphi] \equiv (T^{p_1,p_2-1})_{\varphi'} \circ \overline{L},$$

$$[(T^{p_1,p_2})_\varphi, z] = (T^{p_1-1,p_2})_{\varphi'} \circ z,$$

and

$$[(T^{p_1,p_2})_\varphi, \bar{z}] = \bar{z} \circ (T^{p_1,p_2-1})_{\varphi'},$$

where the $\equiv$ denotes modulo $C^{p_1-p_1'+p_2-p_2'}$ terms of the form

$$\frac{L^{p_1-p_1'} \circ z^{p_1-p_1'} \circ T^{p_1'} \circ \varphi^{(p_1-p_1'+p_2-p_2'+1)} \circ T^{p_2'} \circ \bar{z}^{p_2-p_2'} \circ \overline{L}^{p_2-p_2'}}{(p_1-p_1')!(p_2-p_2')!} \tag{12.3}$$

with either $p_1' = 0$ or $p_2' = 0$, i.e., terms in which all free T derivatives have been eliminated on one side of φ or the other. Thus if we start with $p_1 = p_2 = p/2$, and iteratively apply these commutation relations, the number of T derivatives not necessarily applied to φ is eventually at most $p/2$.

12.3 The Recurrence

We first insert $v = (T^{\frac{p}{2},\frac{p}{2}})_\varphi u$ in the a priori inequality, then bring $(T^{\frac{p}{2},\frac{p}{2}})_\varphi$ to the left of $P = -L\overline{L} - \overline{L}z^k\bar{z}^k L$, since Pu is known and analytic. We have, omitting for now the "subelliptic" term,

$$\begin{aligned}
&\|\overline{L}(T^{\frac{p}{2},\frac{p}{2}})_\varphi u\|_0^2 + \|\bar{z}^k L(T^{\frac{p}{2},\frac{p}{2}})_\varphi u\|_0^2\\
&\quad \lesssim |(P(T^{\frac{p}{2},\frac{p}{2}})_\varphi u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)_0|\\
&\quad \lesssim |((T^{\frac{p}{2},\frac{p}{2}})_\varphi Pu, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)_0| + |([P, (T^{\frac{p}{2},\frac{p}{2}})_\varphi]u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)_0|
\end{aligned}$$

and, by the above bracket relations,

$$\begin{aligned}
&([P, (T^{\frac{p}{2},\frac{p}{2}})_\varphi]u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)\\
&\quad = -([L\overline{L}, (T^{\frac{p}{2},\frac{p}{2}})_\varphi]u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u) - ([\overline{L}z^k\bar{z}^k L, (T^{\frac{p}{2},\frac{p}{2}})_\varphi]u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)\\
&\quad \equiv -(L(T^{\frac{p}{2},\frac{p}{2}-1})_{\varphi'}\overline{L}u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u) - (L(T^{\frac{p}{2}-1,\frac{p}{2}})_{\varphi'}\overline{L}u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)\\
&\qquad -((T^{\frac{p}{2}-1,\frac{p}{2}})_{\varphi'}\overline{L}z^k\bar{z}^k Lu, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)\\
&\qquad -\sum_{k'=1}^{k}(\overline{L}z^{k'}(T^{\frac{p}{2},\frac{p}{2}-1})_{\varphi'}z^{k-k'}\bar{z}^k Lu, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)\\
&\qquad -\sum_{k'=0}^{k-1}(\overline{L}z^k\bar{z}^{k'}(T^{\frac{p}{2}-1,\frac{p}{2}})_{\varphi'}\bar{z}^{k-k'} Lu, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u)\\
&\qquad -(\overline{L}z^k\bar{z}^k L(T^{\frac{p}{2},\frac{p}{2}-1})_{\varphi'}u, (T^{\frac{p}{2},\frac{p}{2}})_\varphi u),
\end{aligned}$$

with the same meaning for $\equiv$ as above.

In every term, no powers of z or $\bar{z}$ have been lost, though some may need to be brought to the left of the $(T^{q_1,q_2})_{\tilde{\varphi}}$ with again no loss of powers of z or $\bar{z}$ and a further reduction in order. Every bracket reduces the order of the sum of the two indices p_1 and p_2 by one (here we started with $p_1 = p_2 = p/2$), picks up one derivative on φ, and leaves the vector fields over which we have maximal control in the estimate intact and in the correct order.

Thus we may bring either $\overline{L}z^k$ or L to the right as $\bar{z}^k L$ or $\overline{L}$, and use a weighted Schwarz inequality on the result to take maximal advantage of the a priori inequality. Iterations of all of this continue until there remain at most $p/2$ free T derivatives (i.e., the T derivatives on at least one side of φ are all "corrected" by good vector fields) and perhaps as many as $p/2$ L or $\overline{L}$ derivatives, and we may continue further until, at worst, these remaining L or $\overline{L}$ derivatives bracket two at a time to produce more T's, with corresponding combinatorial factors.

After all of this, there will be at most $T^{\frac{3p}{4}}$ derivatives remaining, and a factor of $\frac{p}{2}!! \sim \frac{p}{4}!$

It is here that the final term on the left of the a priori inequality is used, in order to bring the localizing function out of the norm after creating another balanced localization of $T^{3p/4}$ with a new localizing function of Ehrenpreis type with slightly larger support, geared, roughly, to $3p/4$ instead of to p.

Recall that such such localizing functions ψ may be constructed for any N and that they satisfy

$$\left|\psi^{(r)}\right| \leq \left(\frac{C}{e}\right)^{r+1} N^r, \quad r \leq 2N,$$

where C is independent of N and $e = \text{dist}(\{\psi \equiv 1\}, (\text{supp}\,\psi)^c)$.

12.4 Conclusion of the Proof

Finally, this entire process, which reduced the order from p to at most $3p/4$, (or more precisely to at most $3p/4 + (k-1)/2$), is repeated, over and over, each time essentially reducing the order by a factor of $3/4$. After at most $\log_{4/3} p$ such iterations we are reduced to a bounded number of derivatives, and, as in [T4] and [T5], all of these nested open sets may be chosen to fit in the one open set Ω' where Pu is known to be analytic, and all constants chosen independent of p (but depending on Pu). The fact that in those works one full iteration reduced the order by half played no essential role; a factor of $3/4$ works just as well.

To be precise, the sequence of open sets $\{\Omega_j\}$, each compactly contained in the next, with $\Omega_{\log_{4/3} p} = \Omega'$, have separations $d_j = \text{dist}(\Omega_j, \Omega^c_{j+1})$, with $\sum d_j = \text{dist}(\Omega_0, \Omega'^c) = d$, which need to be picked carefully. The localizing functions $\{\varphi_j\}$ with $\varphi_j \in C_0^\infty(\Omega_{j+1}) \equiv 1$ on Ω_j satisfy

$$\left|\varphi_j^{(r)}\right| \leq (C/d_j)^{r+1}((3/4)^j p)^r, \qquad r \leq 2(3/4)^j p. \tag{12.4}$$

We shall take $d_j = \frac{1}{(j+1)^2}/d \sum \frac{1}{(j+1)^2}$.

Now at most $(3/4)^j p$ derivatives will fall on φ_j, and most of the effect of the derivatives will be balanced by corresponding factorials in the denominator, as in (12.3), roughly the powers of $(3/4)^j p$ in (12.4) in view of Stirling's formula. In addition, as noted immediately before the last paragraph in the previous section, there will be factorials corresponding to the diminution of powers of T. What will *not* be balanced are the powers of d_j^{-1}, but the product of these factors will contribute

$$\Pi_{j=1}^{\log_{4/3} p} \left(j^2\right)^{(3/4)^j p} = \left(\Pi_{j=1}^{\log_{4/3} p} j^{(3/4)^j}\right)^{2p} = C^p,$$

which, together with the factorials just mentioned, proves the analyticity of the solution in Ω_0.

Remark 12.1. Actually, the d_j could also have been taken equal to $Cd_0/(\frac{4}{3})^j$, which would have been more in line with the previous chapters.

12.5 Gevrey Regularity for Kohn–Oleinik Operators

In this section we merely make two notes. The first is that together with A. Bove in [BT4], we have studied the Gevrey regularity of an operator that blends the characteristics of the Kohn example and the Oleinik–Radkevich operator introduced above. Here we consider the analogue of Kohn's operator but with a point singularity,

$$P = BB^* + B^*(t^{2\ell} + x^{2k})B, \qquad B = \frac{\partial}{\partial x} - i x^{q-1} \frac{\partial}{\partial t},$$

and show that this operator is hypoelliptic and Gevrey hypoelliptic in a certain range, namely $k < \ell q$, with Gevrey index $\frac{\ell q}{\ell q - k} = 1 + \frac{k}{\ell q - k}$.

Work in progress by Bove, Mughetti, and Tartakoff proves that when this range $k \leq \ell q$ is violated, the operator is not even hypoelliptic.

Chapter 13
Nonlinear Problems

13.1 Global Regularity

Our aim in this work is to prove a global analytic regularity result on a compact manifold M for some quasilinear equations.

A model of such results is the following: in C^2 take a bounded domain Ω with strictly pseudoconvex real analytic boundary M.

Then one has two globally defined independent real, real analytic vector fields X_1 and X_2, namely the real and imaginary parts of a (globally defined) holomorphic vector field $L = X_1 - iX_2$ tangent to M.

Take a function u in $C^\infty(M)$ and consider an analytic matrix function $H(x,t)$ defined on a neighborhood of $\{(x, u(x)) : x \in M\}$ in $M \times \mathbb{C}$ and set

$$Y = H(X), \quad \text{i.e., } Y(x) = H(x, u(x))X(x),$$

so that one obtains two C^∞ vector fields, Y_1 and Y_2, on M.

We assume that $H(x,t)$ is invertible, so that one can express $X = H^{-1}(Y)$ with $H^{-1} \in C^\omega$.

Consider the operator

$$P_u = Y_1^2 + Y_2^2 + a_1 Y_1 + a_2 Y_2 + b$$

with a_j, b analytic and assume that $P_u u \in \mathcal{A}(M)$. Can one conclude that u is analytic on M if the associated Levi form is nondegenerate? Note that $P_u u = f$ is a quasilinear equation.

The question is global. There are known local results for more special equations (cf. [TZ]).

In higher dimensions, one generally does not have globally defined vector fields $X_1, \ldots, X_{2n}$ related to a CR structure on M induced by the complex structure on $\mathbb{C}^n$.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_13, © Springer Science+Business Media, LLC 2011

But one can consider a (finite) family of open sets $\{V_\ell\}_{1\le\ell\le p}$ covering M and analytic vector fields $\{X_{k,\ell}\}_{k=1,\dots,2n}$ on V_ℓ. Then we may consider

$$Y_{(\ell)} = H(X_{(\ell)}), \quad \text{where } X_{(\ell)} = \begin{pmatrix} X_{1,\ell} \\ \vdots \\ X_{2n,\ell} \end{pmatrix},$$

and the associated operator

$$P_{\ell,u} = \sum_{j=1}^{2n} Y_{j,\ell}^2 + a_{j,\ell} Y_{j,\ell} + b_\ell, \quad \text{with } a_{j,\ell}, b_\ell \text{ analytic.}$$

Now assume that for all ℓ,

$$P_{\ell,u} u \in \mathcal{A}(V_\ell).$$

Then the question is this: is u analytic on M under a nondegeneracy hypothesis on the associated Levi form?

Theorem 13.1. *Under the above hypotheses, if Pu is real analytic globally on M then so is any (moderately smooth) solution u.*

A more interesting problem (as related to the complex Laplacian on forms) is to consider a system. But here we consider only the scalar case. Note that from results on C^∞ regularity in [Xu1], one need only assume that u is in $C^{2,\alpha}$ in our theorem.

13.2 Some Notation and Definitions

Let M be a compact real analytic manifold of dimension $2n+1 \ge 3$, and let $(V_j)_{j=1,\dots,p}$ be a covering of M such that in each V_j, there are given $2n$ real analytic, real vector fields $X_{1,j}, \dots, X_{2n,j}$ such that:

- On $V_j \cap V_k$ every $X_{\ell,j}$ (resp. $X_{\ell,k}$) is a linear combination of the $(X_{\ell,k})_{\ell=1,\dots,p}$ (resp. of the $(X_{\ell,j})_{\ell=1,\dots,p}$) with real analytic coefficients.
- There exists a globally defined, real analytic real vector field T such that $(X_{1,j}, \dots, X_{2n,j}, T)$ is a basis in V_j and if

$$[X_{\ell,j}, X_{m,j}] \equiv a^j_{\ell m} T \mod (X_{\ell,j}), \tag{13.1}$$

then the matrix $(a^j_{\ell m})$ is nondegenerate.

Remark 13.1. It is a result that goes back to Tanaka [Tan] and used to advantage in the work of the present author in many places that under the nondegeneracy assumption (13.1), one *always* may modify the given vector field T by adding multiples of the $X_{j,\ell}$ in such a way that

$$[X_{j,\ell}, T] \equiv 0 \mod (X_{1,\ell}, \ldots, X_{1,\ell}). \tag{13.2}$$

Definition 13.1. We call such a family $(X_{\ell,j}, T)$ of systems of vector fields a compatible family.

Now we may assume that each (V_j) is the domain of a coordinate chart. So in each V_j and for every $s \geq 0$, we may consider an elliptic pseudodifferential operator of order s, which we denote by Λ_j^s.

Let us fix a family $(\varphi_j)_{j=1,\ldots,p}$ such that

$$\varphi_j \in \mathcal{D}(V_j), \quad 0 \leq \varphi_j \leq 1, \quad \sum \varphi_j \equiv 1 \text{ on } M. \tag{13.3}$$

Let $(\rho_j)_{j=1,\ldots,p}$ be another family such that

$$\rho_j \in \mathcal{D}(V_j), \quad 0 \leq \rho_j \leq 1, \quad \rho_j \equiv 1 \text{ on supp } \varphi_j. \tag{13.4}$$

Now one has, say for $t \geq s \geq 0$,

$$\|\varphi_j v\|_t \lesssim (\|\rho_j \Lambda_j^s \varphi_j v\|_{t-s} + \|\varphi_j v\|_0), \quad \forall v \in C^\infty(M), \tag{13.5}$$

where $\| \quad \|_t$ denotes the Sobolev norm.

So now one has

$$\|v\|_t \lesssim \sum_j \|\varphi_j v\|_t \lesssim \sum_j (\|\rho_j \Lambda_j^s \varphi_j v\|_{t-s} + \|\varphi_j v\|_0), \quad \forall v \in C^\infty(M). \tag{13.6}$$

Now in each V_j, we consider the operator considered in the introduction (depending on the given $u \in C^\infty(M)$):

$$\begin{cases} P_j = \sum_{\ell=1}^{2n} (Y_{\ell,j}^2 + a_{\ell,j} Y_{\ell,j} + b_j), \\ a_{\ell,j}, b_j \in \mathcal{A}(V_j), \end{cases} \tag{13.7}$$

and assume that

$$P_j u \in \mathcal{A}(V_j), \quad \forall j. \tag{13.8}$$

Finally let us denote by $(\quad , \quad)_s$ the s-Sobolev scalar product (in each V_j, when one has functions with compact support in V_j) and remember the following:

$$\forall \delta > 0, \exists C_\delta : \forall w \in C_0^\infty, \|w\|_s^2 \leq \delta \|w\|_{s+1/2}^2 + C_\delta \|w\|_0^2. \tag{13.9}$$

13.3 Maximal Estimates

Our aim in this section is to prove the following:

Theorem 13.2. *We have the following maximal estimates for $s \geq 0$:*

$$\|v\|^2_{s+1/2} + \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell v\|^2_s \lesssim \sum_\ell |(\varphi_\ell P_\ell v, \varphi_\ell v)_s| + \|v\|^2_0 \tag{E1$_s$}$$

for $v \in C^\infty(M)$ and

$$\|v\|^2_{s+1} + \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell v\|^2_{s+1/2} \lesssim \sum_\ell \|\varphi_\ell P_\ell v\|^2_s + \|v\|^2_0 \tag{E2$_s$}$$

for $v \in C^\infty(M)$, and in fact,

$$\|v\|^2_{s+1} + \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell v\|^2_{s+1/2} + \sum_{j,k,\ell} \|X_{j,\ell}X_{k,\ell}\varphi_\ell v\|^2_s \lesssim \sum_\ell \|\varphi_\ell P_\ell v\|^2_s + \|v\|^2_0 \tag{E3$_s$}$$

for $v \in C^\infty(M)$.

Proof. The proof is known when written for functions with compact support in coordinate charts. This is a global version. Let us first verify the statements at level $s = 0$. For simplicity we take ℓ fixed and set $X_{j,\ell} = X_j$, $j = 1, \dots 2n$, and $\varphi = \varphi_\ell$. We have

$$\sum_j \|X_j\varphi v\|^2_0 = \sum (X_j\varphi v, X_j\varphi v) \lesssim \sum (Y_j\varphi v, Y_j\varphi v) \tag{13.10}$$

because H is invertible. Note that $\lesssim$ may indicate a constant that depends on u and its first few derivatives. Now

$$\begin{aligned}(Y_j\varphi v, Y_j\varphi v) &= -(Y_j^2\varphi v, \varphi v) + (\theta_j\varphi v, \varphi v), \quad \theta_j \in C^\infty(V_j),\\ &= -([Y_j^2, \varphi]v, \varphi v) + (\varphi Y_j^2 v, \varphi v) + \mathcal{O}(\|\varphi v\|_0\|Y_j\varphi v\|_0).\end{aligned}$$

Now, using $[Y_j^2, \varphi] = 2Y_j[Y_j, \varphi] - [Y_j, [Y_j, \varphi]]$, we easily obtain

$$|([Y_j^2, \varphi]v, \varphi v)| \lesssim C_1\|v\|^2_0 + \frac{1}{C_0}\sum_j \|X_j\varphi v\|^2_0. \tag{13.11}$$

Thus,

$$\sum_j \|X_j\varphi v\|_0^2 \lesssim |(\varphi P v, \varphi v)| + C_1\|v\|_0^2 + \frac{1}{C_0}\sum_j \|X_j\varphi v\|_0^2. \tag{13.12}$$

Now use

$$\|\varphi v\|_{1/2} \lesssim \sum_j \|X_j\varphi v\|_0^2 + \|\varphi v\|_0^2 \qquad \forall v \in C^\infty(M) \tag{13.13}$$

(see J.J. Kohn [K1]) to obtain (E1_s) in case $s = 0$.

Now we can deduce (E2_0) from (E1_0) in the following way: using (13.5) and (13.6), we have

$$\begin{aligned}\sum_\ell \|\varphi_\ell v\|_1^2 + \sum_{j,\ell}\|X_{j,\ell}\varphi_\ell v\|_{1/2}^2 &\lesssim \sum_\ell \|\rho_\ell\Lambda_\ell^{1/2}\varphi_\ell v\|_{1/2}^2 \\ &+ \sum_{j,\ell}\|\rho_\ell\Lambda_\ell^{1/2}X_{j,\ell}\varphi_\ell v\|_0^2 + \|v\|_0^2 + \sum_{j,\ell}\|X_{j,\ell}\varphi_\ell v\|_0^2 \\ &\lesssim \sum_\ell \|\rho_\ell\Lambda_\ell^{1/2}\varphi_\ell v\|_{1/2}^2 + \sum_{j,\ell}\|X_{j,\ell}\rho_\ell\Lambda_\ell^{1/2}\varphi_\ell v\|_0^2 + \sum_\ell\|\varphi_\ell v\|_{1/2}^2 + \sum_\ell \|X_{j,\ell}\varphi_\ell v\|_0^2.\end{aligned} \tag{13.14}$$

The last two sums are easy to handle. The first two sums are (from the first part of the theorem at level $s = 0$) less than

$$\sum_\ell |(\rho_\ell P_\ell\Lambda_\ell^{1/2}\varphi_\ell v, \rho_\ell\Lambda_\ell^{1/2}\varphi_l v)| + \sum_\ell \|\varphi_l v\|_{1/2}^2. \tag{13.15}$$

Now, for simplicity we forget the index ℓ and consider

$$\rho P\Lambda^{1/2}\varphi v = [\rho P, \Lambda^{1/2}\varphi]v + \Lambda^{1/2}\varphi P v. \tag{13.16}$$

Just as we obtained (13.11), we have

$$|([\rho P, \Lambda^{1/2}\varphi]v, \rho\Lambda^{1/2}\varphi v)| \le C_1\|v\|_{1/2}^2 + \frac{1}{C_0}\sum \|X_j\varphi v\|_{1/2}^2. \tag{13.17}$$

By taking C_0 large enough and using (13.9) we have the desired inequality, because the term $|(\Lambda^{1/2}\varphi P v, \rho\Lambda^{1/2}\varphi v)|$ is less than $C_1\|\varphi P v\|^2 + \frac{1}{C_0}\sum\|\varphi v\|_1^2$ (with C_0 large, C_1 depending on C_0 as usual).

This proves (E2_s) with $s = 0$. To bound also the third term on the left in (E3_s) for $s = 0$, we argue as follows: first the function $X_{j,\ell}v$ is inserted in place of v in (E1_s) with $s = 0$, and an error of the type $C\sum_\ell \|v\|_1^2$ is introduced through a bracket of the form $([X, X]v, X^2 v)$. While this is a new error, it is already controlled by (E2_s)s, which completes the proof for $s = 0$.

Our aim is to prove (E1$_s$) and deduce (E2$_s$) from (E1$_s$) as before.

First observe that (in view of (13.5))

$$\|v\|^2_{s+1/2} \lesssim \sum_\ell \|\rho_\ell \Lambda^s_\ell \varphi_\ell v\|^2_{1/2} + \|v\|^2_s \tag{13.18}$$

and

$$\sum_\ell \|X_{j,\ell}\varphi_\ell v\|^2_s \lesssim \sum_\ell \|X_{j,\ell}\rho_\ell \Lambda^s_\ell \varphi_\ell v\|^2_0 + \|v\|^2_s \lesssim \tag{13.19}$$

$$\lesssim \sum_\ell \|X_{j,\ell}\rho_\ell \Lambda^s_\ell \varphi_\ell v\|^2_0 + \sum_\ell \|\rho_\ell \Lambda^s_\ell \varphi_\ell v\|^2_{1/2} + \|v\|^2_s. \tag{13.20}$$

Now, we use (3.1$_0$) to obtain

$$\|v\|^2_{s+1/2} + \sum_\ell \|X_{j,\ell}\varphi_\ell v\|^2_s \lesssim \sum_\ell |(\rho_\ell P_\ell \Lambda^s_\ell \varphi_\ell v, \rho_\ell \Lambda^s_\ell \varphi_\ell v)|. \tag{13.21}$$

In view of (13.9), the term $\|v\|^2_s$ may be replaced by $\|v\|^2_0$.

Now we consider the first term in the second member of (13.21) and write $\rho_\ell P_\ell \Lambda^s_\ell \varphi_\ell v = \rho_\ell [P_\ell, \Lambda^s_\ell \varphi_\ell] v + \rho_\ell \Lambda^s_\ell \varphi_l P_\ell v$.

So one sees that one is reduced to studying $(\rho_\ell [P_\ell, \Lambda^s_\ell \varphi_\ell] v, \rho_\ell \Lambda^s_\ell \varphi_l v)$, because one has easily

$$|(\rho_\ell \Lambda^s_\ell \varphi_\ell P_\ell v, \rho_\ell \Lambda^s_\ell \varphi_l v)| \le C |(\varphi_\ell P_\ell v, \varphi_\ell v)_s| + \delta \left\{ \sum_\ell \|\varphi_l v\|^2_{s+1/2} + \sum_\ell \|X_{j,\ell}\varphi_l v\|^2_s \right\} + C_\delta \|v\|^2_s. \tag{13.22}$$

Now again forget the index ℓ and consider

$$[P, \Lambda^s \varphi] = \sum [Y_j^2, \Lambda^s \varphi] = \sum 2Y_j [Y_j, \Lambda^s \varphi] - [Y_j, [Y_j, \Lambda^s \varphi]]. \tag{13.23}$$

Then one has, as in (13.11),

$$|(\rho_\ell [P_\ell, \Lambda^2_\ell \varphi_\ell] v, \rho_\ell \Lambda^s_\ell \varphi_\ell v)| \le \frac{1}{C_0} \sum_j \|X_{j,\ell}\rho_\ell \Lambda^s_\ell \varphi_\ell v\|^2_0 + C_1 \|\rho_\ell \Lambda^s_\ell \varphi_\ell v\|^2_0, \tag{13.24}$$

where C_1 depends on C_0, as usual.

Then again using (13.9) and taking C_0 big enough, the first member of (13.21) is less than $C(\sum_\ell |(\varphi_\ell P_\ell v, \varphi_\ell v)_s| + \|v\|^2_0)$ for some $C > 0$.

Now we want to prove (E2$_s$) using (E1$_s$) as we did for $s = 0$.

One has, as in that case,

$$\sum_\ell \|\varphi_\ell v\|_{s+1}^2 + \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell v\|_{s+1/2}^2 \tag{13.25}$$

$$\lesssim \sum_\ell \|\rho_\ell \Lambda_\ell^{1/2}\varphi_\ell v\|_{s+1/2}^2 + \sum_{j,\ell}\|\rho_\ell \Lambda_\ell^{1/2} X_{j,\ell}\varphi_\ell v\|_s^2 + \|v\|_{s+1/2}^2 + \sum_{j,\ell}\|X_{j,\ell}\varphi_\ell v\|_s^2 \tag{13.26}$$

$$\lesssim \sum_\ell \rho_\ell \Lambda_\ell^{1/2}\varphi_\ell v\|_{s+1/2}^2 + \sum_{j,\ell} \|X_{j,\ell}\rho_\ell \Lambda_\ell^{1/2}\varphi_\ell v\|_s^2 + \sum_\ell \|\varphi_\ell v\|_s^2 \tag{13.27}$$

$$+ \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell v\|_s^2. \tag{13.28}$$

The worst terms are the first two, from (E1$_s$). They are less than

$$\sum_\ell |(\rho_\ell P_\ell \Lambda_\ell^{1/2}\varphi_\ell v, \rho_\ell \Lambda_\ell^{1/2}\varphi_\ell v)_s| + \|v\|_s^2. \tag{13.29}$$

The end of the proof of (E2$_s$) now follows the lines of the end of the proof in the case $s = 0$, and the proof of (E3$_s$) is also as before. □

Remark 13.2. This proof of the global version requires nothing more than careful computations.

Corollary 13.1. *Let T be a global, real analytic, nonzero vector field on M complementary to the X and satisfying (13.1). Then*

$$\|Tv\|_s^2 \lesssim \sum_\ell \|\varphi_\ell P_\ell v\|_s^2 + \|v\|_0^2 \qquad \forall v \in C^\infty(M). \tag{13.30}$$

Remark 13.3. The existence of such a global T has been shown [T4], [T5] when M is a compact real CR manifold.

13.4 High Powers of the Vector Field T

The overall strategy is to use the maximal estimates above with v replaced by $T^p u$. Once one has control over high T derivatives of the solution, the other derivatives follow by standard techniques.

Now the vector field T being global, dealing with just a bounded number of vector fields X_j that are only locally defined is not a very delicate issue. For instance, the above corollary may be strengthened to

$$\|Tv\|_s^2 + \|v\|_{s+1}^2 + \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell v\|_{s+1/2}^2 + \sum_{j,k,\ell} \|X_{j,\ell}X_{k,\ell}\varphi_\ell v\|_s^2$$
$$\lesssim \sum_\ell \|\varphi_\ell P_\ell v\|_s^2 + \|v\|_0^2 \qquad \forall v \in C^\infty(M), \tag{13.31}$$

or, in the form we will use it, for any integer r,

$$\|T^{r+1}v\|_s^2 + \sum_{j,\ell} \|X_{j,\ell}\varphi_\ell T^r v\|_{s+1/2}^2 + \sum_{j,k,\ell} \|X_{j,\ell}X_{k,\ell}\varphi_\ell T^r v\|_s^2$$
$$\lesssim \sum_\ell \|\varphi_\ell P_\ell T^r v\|_s^2 + \|T^r v\|_0^2 \qquad \forall v \in C^\infty(M). \tag{13.32}$$

Proposition 13.1. *There exist constants C, C_u, C_h such that for all r,*

$$\frac{1}{r!}\{\|T^{r+1}v\|_s^2 + \|X^2T^r v\|_s^2\} \le 4^r C^r C_u^r C_h^r. \tag{13.33}$$

In view of (13.32), our (only) task is to commute T^r past P_ℓ with errors that can be recursively estimated to grow "analytically," since then after specializing v to u, we know that $P_u u \in \mathcal{A}(M)$.

Now we have the crucial relationship (in V_j)

$$[T, X_{j,\ell}] = \sum_k c_{jk} X_{k,\ell}$$

with c_{jk} real analytic functions. Since P_ℓ is a quadratic polynomial in the $X_{j,\ell}$, with coefficients $h(x,u)$ that are real analytic functions of the spatial variables x *and* the solution $u(x)$, we may write

$$[T, P_\ell] = \sum [T, h(x,u)X_{k,\ell}X_{j,\ell}] = (Th)X^2 + h[T, X^2]$$
$$= (Th)X^2 + h(\tilde{a}X^2 + \tilde{\tilde{a}}X + \tilde{\tilde{\tilde{a}}})$$

with analytic functions $\tilde{a}$, $\tilde{\tilde{a}}$, and $\tilde{\tilde{\tilde{a}}}$. Generically, then, we may write

$$[T, P_\ell] = \sum ((T)h)X^2, \tag{13.34}$$

where $h((T)h)$ denotes (at most) a first derivative (in (x,t)) of $h(x,u(x))$ times one of a finite collection of analytic functions of x, namely the coefficients of the X in the bracket $[T, X]$ mentioned above in (13.1) (and possibly one derivative of this coefficient). There may also be fewer than two X's on the right in (13.34). For the rest of the paper we will assume for simplicity that $h = h(u)$.

Next, we will need an expression for the more complicated bracket

$$[T^r, P_\ell] = \sum_{r'=0}^{r-1} T^{r'}[T, P_\ell]T^{r-r'-1} = \sum_{r'=0}^{r-1} T^{r'} \circ h' X^2\, T^{r-r'-1},$$

only we move the h' to the very left yet leave X^2 for the moment wherever they are, which we denote by enclosing the X^2 in parentheses and placing them on the left. That is,

$$[T^r, P_\ell] = \sum_{r'=1}^{r} \binom{r}{r'} (T^{r'} h)(u(x))(X^2)T^{r-r'}, \tag{13.35}$$

or

$$\frac{[T^r, P_\ell]}{r!} = \sum_{r'=1}^{r} \frac{(T^{r'} h)(u(x))}{r'!} \, \frac{(X^2)T^{r-r'}}{(r-r')!}.$$

Now for the term $(T^{r'} h)(u(x))$, we will use the Faà di Bruno formula or rather crude bounds for the results, writing

$$D_x^k g(u(x)) = (u' D_u + D_x)^{k-1} u' D_u g(u(x)).$$

Writing this crudely as

$$D_x^k g(u(x)) = ((u'\sigma + D_x)^{k-1} u'_{|\sigma = D_g})g,$$

σ becomes a "counter" for the number of derivatives received by g. Then this is *at worst*

$$\sum_{k'} \binom{k}{k'} g^{(k-k')}(D_x^{k'} u'^{k-k'}). \tag{13.36}$$

Finally, we must analyze expressions such as $D^a u'^b$:

$$D^a u'^b = \sum_{a_1+\ldots+a_b=a} \binom{a}{a_1, \ldots, a_b} u'^{(a_1)} \cdots u'^{(a_b)},$$

where $\binom{a}{a_1,\ldots,a_b}$ denotes the multinomial expression

$$\binom{a}{a_1, \ldots, a_b} = \frac{a!}{a_1! \cdots a_b!(a - \sum a_j)!} = \frac{a!}{a_1! \cdots a_b!},$$

since $\sum a_j = a$. We have

$$\frac{D^a u'^b}{a!} = \sum_{a_1+\cdots+a_b=a} \frac{u'^{(a_1)}}{a_1!} \cdots \frac{u'^{(a_b)}}{a_b!}.$$

Thus we have

$$\frac{[T^r, P]}{r!} = \sum_{r'=1}^{r} \frac{(T_{x,t}^{r'} h(u(x,t)))}{r'!} \frac{(X^2)T^{r-r'}}{(r-r')!}, \tag{13.37}$$

and so (cf. (13.36))

$$\begin{aligned}\frac{(T^{r'})h(u(x))}{r'!} &\sim \sum_{r'-r''\geq 1}^{r} \frac{h^{(r'-r'')}}{(r'-r'')!} \frac{(T^{r''}(u'^{r'-r''}))}{r''!}\\ &= \sum_{r'-r''\geq 1}^{r} \frac{h^{(r'-r'')}}{(r'-r'')!} \sum_{\sum_1^{r'-r''} r''_j = r''} \frac{T^{r''_1} u'}{r''_1!} \cdots \frac{T^{r''_{r'-r''}} u'}{r''_{r'-r''}!},\end{aligned}$$

or in all,

$$\frac{[T^r, P]v}{r!} = \sum_{\substack{r\geq r'-r''\geq 1\\ \sum_{j=1}^{r'-r''} r''_j = r''\\ (\sum_{j=1}^{r'-r''}(r''_j+1)=r')}} \frac{h^{(r'-r'')}}{(r'-r'')!} \frac{T^{r''_1} u'}{r''_1!} \cdots \frac{T^{r''_{r'-r''}} u'}{r''_{r'-r''}!} \frac{(X^2)T^{r-r'}v}{(r-r')!}. \tag{13.38}$$

To simplify the argument we have dropped the localizing functions, since for global arguments, in which case these functions always appear on the left of the norms, they may be brought out easily and replaced by another partition of unity, with new X's if needed.

We also have ignored the order of the X's and T's, indicating this by putting the X^2 in parentheses, not to indicate that they may not be present (though they may not) but that they may appear with several T's to the left and more to the right. We have also dropped all subscripts. Schematically, we may then write (13.32) together with (13.38) as

$$\begin{aligned}\|T^{r+1}v\|_s + \|X^2T^rv\|_s &\lesssim \|PT^rv\|_s + \|T^rv\|_s\\ &\lesssim \|T^rPv\|_s + \|T^rv\|_0^2 + \|[T^r, P]v\|_s \quad \forall v \in C^\infty(M),\end{aligned} \tag{13.39}$$

and so

$$\begin{aligned}\frac{1}{r!}\{\|T^{r+1}v\|_s + \|X^2T^rv\|_s\} &\lesssim \frac{1}{r!}\{\|T^rPv\|_s + \|T^rv\|_s\}\\ &+ \sum_{\substack{r\geq r'-r''\geq 1\\ \sum_{j=1}^{r'-r''} r''_j = r''\\ (\sum_{j=1}^{r'-r''}(r''_j+1)=r')}} \left\| \frac{h^{(r'-r'')}}{(r'-r'')!} \frac{T^{r''_1} u'}{r''_1!} \cdots \frac{T^{r''_{r'-r''}} u'}{r''_{r'-r''}!} \frac{(X^2)T^{r-r'}v}{(r-r')!} \right\|_s.\end{aligned} \tag{13.40}$$

Now for our value of s, H^s is an algebra, and so the norm of the product of derivatives of copies of u may be replaced by the product of the norms, each of which will have the form of one of the terms on the left-hand side of (13.40).

Specializing to $v = u$ and bounding the norm of the product by the product of the norms, we observe that except for the term involving derivatives of h, all other terms are of the same form, since one T derivative and two X derivatives carry the same weight on the left-hand side of (13.40):

$$\frac{1}{r!}\{\|T^{r+1}v\|_s + \|X^2T^r v\|_s\} \lesssim \frac{1}{r!}\{\|T^r Pv\|_s + \|T^r v\|_s\}$$
$$+ \sum_{r\geq r'-r''\geq 1} C^{r'-r''+2}\left\|\frac{h^{(r'-r'')}}{(r'-r'')!}\right\|_s \left\|\frac{(X^2)T^{r-r'}v}{(r-r')!}\right\|_s \prod_{\sum_{j=1}^{r'-r''} r_j''=r''} \left\|\frac{T^{r_j''}u'}{r_j''!}\right\|_s . \tag{13.41}$$

The constant includes a power of C for each norm that follows. Note that since the derivatives on h are of that order, this constant will be included with the analyticity constant for h, and in the future, constants with exponents comparable to the number of derivatives on a function known to be analytic will be permitted without comment.

Note that the term in the product with (X^2) is analogous to the extra T derivative on each of the other terms. Hence these terms are similar to the left-hand side, and could be handled at once inductively except for counting the number of them, but it is simpler to iterate (13.41) directly, at least until all terms have order less than $r/2$.

Since there can be at most one term of order larger than $r/2$, after the next "pass" we observe that the product will look just like the right-hand side of (13.40) again, except that there will be one more norm of derivatives of h.

That is, applying (13.41), with r replaced by $r - r'$, to the term in (13.41) with (X^2), we have

$$\frac{1}{(r-r')!}\{\|T^{r-r'+1}v\|_s + \|X^2T^{r-r'}v\|_s\} \lesssim \frac{1}{r-r'!}\{\|T^{r-r'}Pv\|_s + \|T^{r-r'}v\|_0\}$$
$$+ \sum_{r-r'\geq\rho'-\rho''\geq 1} C^{\rho'-\rho''+2}\left\|\frac{h^{(\rho'-\rho'')}}{(\rho'-\rho'')!}\right\|_s \left\|\frac{(X^2)T^{r-r'-\rho'}v}{(r-r'-\rho')!}\right\|_s \prod_{\sum_{j=1}^{\rho'-\rho''} \rho_j''=\rho''} \left\|\frac{T^{\rho_j''}u'}{\rho_j''!}\right\|_s . \tag{13.42}$$

We obtain $r \geq r'+\rho'-r''-\rho''$ and $\sum_{j=1}^{r'-r''}\sum_{k=1}^{\rho'-\rho''}(r_j''+\rho_k'') = r''+\rho''$, so if we set $s' = r'+\rho'$ and $s'' = r''+\rho''$, we have a sum over $s'-s''$ and $\sum_{j+k=2}^{s'-s''} s_{j+k}'' = s''$ subject to the obvious subdivisions.

That is, over $r' + \rho' = s', r'' + \rho'' = s''$,

$$\begin{aligned}
&\frac{1}{r!}\{\|T^{r+1}v\|_s + \|X^2T^r v\|_s\} \\
&\quad\lesssim \sum_{r'\geq 0} \frac{1}{(r-r')!}\{\|T^{r-r'}Pv\|_s + \|T^{r-r'}v\|_s\} \\
&\qquad + \sum_{\substack{s'=r'+\rho', s''=r''+\rho'' \\ r\geq r'-r''\geq 1 \\ r-r'\geq \rho'-\rho''\geq 1 \\ \sum_{j=1}^{s'-s''} s''_j = s'' \\ (\sum_{j=1}^{s'-s''}(s''_j+1)=s')}} \left\| \frac{C^{r'-r''+2}h^{(r'-r'')}}{(r'-r'')!} \right\|_s \left\| \frac{C^{\rho'-\rho''+2}h^{(\rho'-\rho'')}}{(\rho'-\rho'')!} \right\|_s \\
&\qquad \times \left\| \frac{T^{s''_1}u'}{s''_1!} \right\|_s \cdots \left\| \frac{T^{s''_{s'-s''}}u'}{s''_{s'-s''}!} \right\|_s \left\| \frac{(X^2)T^{r-s'}v}{(r-s')!} \right\|_s . \qquad (13.43)
\end{aligned}$$

Note that in using the fact that H^s is an algebra, i.e., $\|fg\|_s \leq B\|f\|_s\|g\|_s$, we have absorbed the algebra constant with the constant C inside the norms of h. We further estimate the norms of derivatives of h (noting that each occurrence contains at least one such derivative) by

$$\|C^{\ell+2}h^{(\ell)}(x, y, u)\|_s \leq C_h^\ell\, \ell!. \qquad (13.44)$$

We are nearly ready to iterate this procedure until even the last term has order less than $r/2$; for except for the sum (the number of terms), each term has a bound that is stable in the number of iterations; namely, the last right-hand side above is bounded by

$$\sum C_h^t \prod \left\{ \left\| \frac{T^p u'}{k!} \right\|_s \text{ or } \left\| \frac{X^2T^p u}{k!} \right\|_s \right\}, \qquad (13.45)$$

where the sum of the k's plus 1 is at most r and $t \leq r$ is the number of terms in the product.

As for the sum, whether after a single full pass or multiple ones, the number of terms corresponds to *at most* the number of ways to partition r derivatives among at most r functions, generally many fewer.

Denoting by D a derivative (r of them) and by u a copy of u (t of them), we are faced with the number of ways to "identify" or select t items (the u's) from among $r + t$ items (the D's and u's) with the understanding that in an expression such as

$$\underbrace{\underbrace{DDDDD}_{r_1}u\underbrace{DDDDD}_{r_2}u\underbrace{DDDDD}_{r_3}u\cdots\underbrace{DDDDD}_{r_t}u}_{r\ D\text{'s and } t(\leq r)\ u\text{'s}}, \qquad (13.46)$$

the D's differentiate only the first u that follows. The answer is that there are certainly not more than $\binom{r+t}{t} \leq 2^{r+t} \leq 2^{2r} = 4^r$ ways.

Finally, since we may iterate this procedure until the maximal order of differentiation on u is 1 or 2, and bound this small number of derivatives by a constant (with at most r such terms, naturally; that is all the derivatives there were), it follows that the left-hand side of (13.40) is bounded by

$$\frac{1}{r!}\{\|T^{r+1}v\|_s^2 + \|X^2T^rv\|_s^2\} \leq 4^r C^r C_u^r C_h^r, \tag{13.47}$$

which clearly yields analytic growth (of T derivatives) of the solution u, since C_u depends only on the first few derivatives of u and s is taken just large enough to ensure that H^s is an algebra. □

13.5 Mixed Derivatives: The Case of Global X

To finish the proof in the case that the vector field(s) X are globally defined, it remains to show that we may estimate mixed derivatives as effectively as we did the high T derivatives.

A result of Helffer and Mattera [HM] shows that it *would* suffice to handle pure powers of the vector field X, but mixed derivatives will invariably enter through brackets of pairs of X's. Thus this we start by using the a priori estimate ($E1_s$)–($E3_s$) with v replaced by φX^r (and later by a mixture of derivatives in X and in T).

What ultimately happens is that brackets of pairs of X's will produce T's, but at most half as many, and we will be led back to (nearly pure) powers of T. The nonlinearity of the problem introduces nothing new in this overall pattern.

When the X's are *globally defined*, for example in $\mathbb{C}^2$, the powers of X are treated like powers of T (e.g., with respect to the use of the Fàa di Bruno formula, especially), with the one change that an additional type of term will appear: starting with $X^{r+2}v$ there will appear as an error $\underline{r}$ copies of X^rTv when two X's bracket to give a T. And this effect, the only new feature, will be repeated until all or nearly all the X's are exhausted. That is, we have the new scheme

$$X^r \to C^{r/2}r!!(X^2)T^{r/2},$$

where we recall the definition $r!! = r(r-2)(r-4)\cdots \sim C^{r/2}(r/2)!$ But this is not a problem, since we have just treated essentially pure powers of T above in (13.47).

Rather than write this case out in more detail, we proceed to the next section, where the *problem* is global but the vector fields X are only locally defined. This case incorporates many of the features of a fully local proof, though fortunately not all! Note that the T vector field is still required to be globally given.

13.6 Locally Defined X_j

When the vector fields X are only locally defined, we cannot afford to change freely from one coordinate patch to another and to another basis of X's each time a localizing function arising from a partition of unity is differentiated, since the constants counting the number of terms and the coefficients would grow far too fast, namely roughly C^r at *each* step. We will need a suitable localization of high powers of the fields X.

We will thus work in a single coordinate patch, drop all subscripts ℓ (and j and k, for simplicity), and in place of v in the a priori estimates substitute $\Psi X^r u$, where the localizing function Ψ will be specified further below. It will not need to be differentiated very often, but the band in which it goes from being identically equal to one to being identically zero will be of a precise width, as will subsequent localizing functions that will be introduced below. The general result on families of localizing functions is given by a result of Ehrenpreis [Eh]:

Proposition 13.2. *For any two open sets $\Omega_0 \Subset \Omega_1$, with separation $d = \mathrm{dist}(\Omega_0, \Omega_1^c)$ and any natural number N, there exist a universal constant C depending only on the dimension and a function $\Psi = \Psi_{\Omega_0,\Omega_1,N} \in C_0^\infty(\Omega_1), \Psi \equiv 1$, on Ω_0 with*

$$|D^\beta \Psi| \le \left(\frac{C}{d}\right)^{|\beta|+1} N^{|\beta|}, \quad |\beta| \le 2N, \tag{13.48}$$

though in this paper we will take $N = 4$ at most; thus the analyticity to be shown in U_0 will be reduced to combining the bounds on $\|T^{r+1}u\|_s$ obtained in a previous section with the bounds

$$\sum \|X^2 \Psi X^r u\|_s + \|T \Psi X^r u\|_s \le C_\Omega^{r+1} r!$$

with $\Psi \equiv 1$ on the set where we want to prove analyticity.

To do this, we start with the a priori estimates as before: for v of compact support where the X_j are defined, and any fixed s, we have ($E3_s$) in the form

$$\|X^2 v\|_s^2 + \|Xv\|_{s-1/2}^2 + \|v\|_{s-1}^2 \lesssim \|Pv\|_s^2 + \|v\|_0^2. \tag{13.49}$$

Note that we have dropped all subscripts but are working with several X's. Naturally we could add a term $\|Tv\|_s$ to the left-hand side using the nonvanishing of the Levi form, but it will not help us here as it did above in handling high powers of T applied to the solution u.

This estimate will be applied to $v = \Psi X^r u$, and then on the right we will write $P \Psi X^r u$ in terms of $\Psi X^r Pu$ modulo an error, namely the commutator of aX^2 with ΨX^r, suitably expanded.

Now the crucial brackets, analogous to (13.35), will be written

$$[P, \Psi X^r]v = a_u [X^2, \Psi] X^r v + a_u \Psi [X^2, X^r] v + \Psi [a_u, X^r] X^2 v, \tag{13.50}$$

where coefficients depending on the solution u (those arising in P and here denoted by a_u) are subscripted with u, while those that depend only on the spatial variables are not subscripted. Now

$$[X^2, \Psi] = 2\Psi' X + \Psi''$$

(and notice that at most two derivatives appear on Ψ and that these will fall to the left of all other terms in the bracket and will be changed with each iteration) and

$$[X, X] = aT, \tag{13.51}$$

and so

$$[X^2, X^r] = \underline{C\,r} \text{ terms } [X, X]X^r + \cdots$$

independent of u. Here underlining a coefficient indicates the number of terms of the given type that occur, and the $\cdots$ denotes terms arising from bringing at least the coefficient in aT to the left of X^r, incurring additional derivatives of course on the coefficient a. However, all of this is linear.

The nonlinear phenomena occur in the last term, where $a_u = a(x, u)$ is differentiated. But this proceeds as before (cf. (13.38)): letting, for instance, $b^{(s)}(x, u)$ denote derivatives of the function b in *its* arguments, all derivatives of the solution u being split off to the right, we obtain

$$\frac{[a_u, X^r]w}{r!} = \sum_{\substack{r \geq r'-r'' \geq 1 \\ \sum_{j=1}^{r'-r''} r''_j = r'' \\ (\sum_{j=1}^{r'-r''} (r''_j + 1) = r')}} \frac{a_u^{(r'-r'')}}{(r'-r'')!} \, \frac{X^{r''_1} u'}{r''_1!} \cdots \frac{X^{r''_{r'-r''}} u'}{r''_{r'-r''}!} \, \frac{X^{r-r'} w}{(r-r')!}. \tag{13.52}$$

So, altogether, (13.50) becomes

$$\begin{aligned}[P, \Psi X^r]v \sim{} & 2\, a_u\, \Psi' X X^r v + a_u\, \Psi'' X^r v + \underline{r}\; a_u\, \Psi\, a\, T X^r v + \cdots \\ & + \Psi\, r! \sum_{\substack{r \geq r'-r'' \geq 1 \\ \sum_{j=1}^{r'-r''} r''_j = r'' \\ (\sum_{j=1}^{r'-r''} (r''_j + 1) = r')}} \frac{a_u^{(r'-r'')}}{(r'-r'')!} \, \frac{X^{r''_1} u'}{r''_1!} \cdots \frac{X^{r''_{r'-r''}} u'}{r''_{r'-r''}!} \, \frac{X^{r-r'} X^2 v}{(r-r')!}.\end{aligned} \tag{13.53}$$

Now once we specialize v to u, we will take the H^s norm of everything and use the property that this space is an algebra for our choice of s. The function Ψ in the product on the right has served its purpose, and we will eventually introduce a new localizing function for each term in the product (except the coefficient, which will just be estimated), though at most one of these terms can have "order" even half of r, and the rest will be handled inductively.

13.7 High Powers of X; New Localizing Functions

More precisely, we restate (13.50) after specialization and introduction of the H^s norm:

$$\frac{\|[P, \Psi X^r]u\|_s}{r!} \le C_u \left\{ \frac{\|\Psi' X^{r+1} u\|_s}{r!} + \frac{\|\Psi'' X^r u\|_s}{r!} + \underline{r} \frac{\|\Psi T X^r u\|_s}{r!} + \cdots \right\}$$
$$+ \sum_{\substack{r \ge r'-r'' \ge 1 \\ \sum_{j=1}^{r'-r''} r''_j = r'' \\ (\sum_{j=1}^{r'-r''} (r''_j+1) = r')}} \left\| \Psi \frac{a_u^{(r'-r'')}}{(r'-r'')!} \frac{X^{r''_1} u'}{r''_1!} \cdots \frac{X^{r''_{r'-r''}} u'}{r''_{r'-r''}!} \frac{X^{r-r'} X^2 u}{(r-r')!} \right\|_{H^s}.$$

We treat the functions $X^2 u$ and u' similarly—they are equivalently handled by the a priori estimate—and *for convenience only* we suppose that the term with $X^2 u$ is of highest order, i.e., $r - r' \ge r''_j \quad \forall j$. Noting that $\operatorname{supp}\Psi \Subset \mathcal{U}_{1/r}$ and bounding the norm of the coefficients by $C^{r'-r''}$ yields

$$\frac{\|[P, \Psi X^r]u\|_s}{r!} \le C_u \left\{ \frac{\|\Psi' X^{r+1} u\|_s}{r!} + \frac{\|\Psi'' X^r u\|_s}{r!} + \underline{r} \frac{\|\Psi T X^r u\|_s}{r!} + \cdots \right\}$$
$$+ \sum_{\substack{r \ge r'-r'' \ge 1 \\ \sum_{j=1}^{r'-r''} r''_j = r'' \\ (\sum_{j=1}^{r'-r''} (r''_j+1) = r')}} C^{r'-r''+2} \left\| \prod_{j=1}^{r'-r''} \frac{X^{r''_j} u'}{r''_j!} \right\|_{H^s(\mathcal{U}_{1/r})} \left\| \frac{\Psi X^{r-r'} X^2 u}{(r-r')!} \right\|_{H^s}. \tag{13.54}$$

Again, we note that the number of terms in the product is $r' - r''$, and hence the constant arising from the algebraicity of H^s will be absorbed with the analyticity constant for the coefficients a_u. Thus we restate (13.54) with this observation, writing ΨX^2 in place of $X^2 \Psi$ on the left, modulo terms on the right, and associating powers of r with derivatives of Ψ or with powers of T, and taking $Pu = 0$:

$$\frac{\|\Psi X^2 X^r u\|_s}{r!} \le C_u \left\{ \sum_{j=1}^{2} \frac{\frac{1}{r^j} \|\Psi^{(j)} X^{r+2-j} u\|_s}{(r-j)!} + \frac{\frac{1}{r} \|\Psi T X^r u\|_s}{(r-2)!} + \cdots \right\}$$
$$+ \sum_{\substack{r \ge r'-r'' \ge 1 \\ \sum_{j=1}^{r'-r''} r''_j = r'' \\ (\sum_{j=1}^{r'-r''} (r''_j+1) = r')}} C_h \left\| \prod_{j=1}^{r'-r''} \frac{X^{r''_j} u'}{r''_j!} \right\|_{H^s(K)} \left\| \frac{\Psi X^2 X^{r-r'} u}{(r-r')!} \right\|_{H^s}. \tag{13.55}$$

As we iterate the terms on the right without T, the order will drop, and we will control the coefficients and the sum below. The term with T is slightly different, but we may always write

$$\frac{\frac{1}{r}\|\Psi T X^r u\|_s}{(r-2)!} = \frac{\frac{1}{r}\|\Psi X^2 X^{r-2} T u\|_s}{(r-2)!}$$

and reapply (13.55) with Tu in place of u but with r decreased by two. Thus we gradually increase the number of T vector fields, with T^σ being balanced by $\frac{1}{\sigma!!}$ before the norm, where

$$\sigma!! = \sigma(\sigma-2)(\sigma-4)\cdots,$$

preserving the balance between remaining powers of X and the large factorial in the denominator, and using up two X's for each T until there are essentially only powers of T, a situation we have treated above. (Of course there will be a "zigzag" effect: sometimes pairs of X's will generate a T and other times the X's will differentiate the coefficients and produce the terms at the end of (13.55) above, so both effects will be combined.)

And as with the estimates of pure T derivatives above, iterating the "principal" term (here the last one, the one with $X^2 X^{r-r'} u$) will lead to a sum with the same bounds for the number of its terms (cf. (13.46)), and with one new norm of derivatives of a coefficient a_u. Even when Ψ has not been differentiated, it will be prudent to change to a new localizing function, one better geared to the number of derivatives appearing under the norm. For there are fewer derivatives now, and it would create significant difficulties to have Ψ' contribute a factor of r when the denominator contains only $(r-r')!$ for rather general r'.

13.8 The Localizing Functions

The first localizing function, $\Psi = \Psi_r$, satisfies

$$\Psi_r \equiv 1 \text{ on } \mathcal{U}_0, \quad \Psi_r \in C_0^\infty(\mathcal{U}_{1/r}), \quad |\Psi_r^{(k)}| \le c^k r^k, \quad k \le p(s) \tag{13.56}$$

(cf. (13.60)), where we have set, for $a \ge 0$,

$$\mathcal{U}_a = \{(x,t) \in \mathcal{U}_1 \ : \ \operatorname{dist}((x,t),\mathcal{U}_0) < a(\operatorname{dist}(\mathcal{U}_0,\mathcal{U}_1^c))\}. \tag{13.57}$$

When the first localizing function needs to be replaced but, say, $\tilde{r}$ derivatives of u remain to be estimated, we shall localize it with a function identically equal to one on $\mathcal{U}_{1/r}$, the support of Ψ_r, but dropping to zero in a band of width $\frac{1}{\tilde{r}} \times (1-\frac{1}{r})(\mathrm{dist}(\mathcal{U}_0, \mathcal{U}_1^c)) = \frac{1}{\tilde{r}}$ times the remaining distance to the complement of $\mathcal{U}_1$, i.e., supported in

$$\mathcal{U}_{\frac{1}{r}+(\frac{1}{\tilde{r}})(1-\frac{1}{r})} = \mathcal{U}_{\frac{1}{r}+\frac{r-1}{r\tilde{r}}} = \mathcal{U}_{1-(1-\frac{1}{r})(1-\frac{1}{\tilde{r}})}. \tag{13.58}$$

We shall denote such a function by ${}_{\frac{1}{r}}\Psi_{\tilde{r}}$. That is, ${}_{\rho}\Psi_{\sigma}$ satisfies

$${}_{\rho}\Psi_{\sigma} \equiv 1 \text{ on } \mathcal{U}_{\rho}, \quad {}_{\rho}\Psi_{\sigma} \in C_0^{\infty}(\mathcal{U}_{\rho+\frac{1}{\sigma}(1-\rho)} \Subset \mathcal{U}_1). \tag{13.59}$$

Derivatives of ${}_{\rho}\Psi_{\sigma}$ satisfy, with universal constant C,

$$|D^k\left({}_{\rho}\Psi_{\sigma}\right)| \leq C^k \left(\frac{\sigma}{1-\rho}\right)^k, \quad k \leq p(s), \tag{13.60}$$

uniformly in ρ, σ, where $p(s)$ will be a small number depending on the s necessary to make H^s an algebra in the given dimension. Of course any other (fixed) bound for k would do.

13.9 Taking a Localizing Function out of the Norm

While it is true that we could just write $\|\Psi w\|_s \leq c\|\Psi\|_s\|w\|_s$, for $s > 1$, to do so would incur at least two derivatives on Ψ with no gain on w. To avoid this difficulty, we use the following finer estimates of the H^s norm of product of functions.

Proposition 13.3. *If* $\Psi, \tilde{\Psi}$ *are two smooth, compactly supported functions with* $\tilde{\Psi} \equiv 1$ *on* $\operatorname{supp}\Psi$ *then for every* $s \leq p \in \mathbb{Z}^+$,

$$\|\Psi D^p u\|_s \leq C^2_{s,\operatorname{supp}\Psi} \sup_{q\leq s} \|D^q\Psi\|_{L^\infty} \|\tilde{\Psi} D^{p-q} u\|_s \tag{13.61}$$

and

$$\|\Psi D^p u\|_s \leq C^2_{s,\operatorname{supp}\Psi} \sup_{q\leq s} \|D^q\Psi\|_{L^\infty} \|D^{p-q} u\|_{H^s(\operatorname{supp}\Psi)}. \tag{13.62}$$

Thus removing a localizing function from an H^s norm, while incurring up to two derivatives on it, does not increase the total number of derivatives being measured, and thus should have minimal impact on the estimates.

Next, we need to confront the effect of these few derivatives on a localizing function Ψ, which may have been chosen with a high number of derivatives (r of them) on u in mind, and which hence adds a factor of r each time a derivative lands on it, when the factorial in the denominator of the corresponding term may be far smaller, e.g., $(r-r')!$ for relatively large r'.

There are in fact several ways to handle this; one is to emphasize that at the level of (13.37) one could endeavor to keep two derivatives to the left of the big bracket whenever possible, so that using Proposition 13.3, those derivatives would serve to bring the Sobolev norms on the right up to H^s again, or one can proceed as follows, the method used in the author's earlier work [TZ]: since the number of terms in the product in (13.54) is $r' - r''$, with

$$r = (r - r') + \sum_{j=1}^{r'-r''} (r_j'' + 1), \text{ with } r - r' \geq \max\{r_j'' + 1\}, \tag{13.63}$$

it follows that

$$(r - r')(r' - r'' + 1) \geq r \tag{13.64}$$

or

$$\left(\frac{r}{r - r'}\right)^k \leq (r' - r'' + 1)^k, \ \forall k, \tag{13.65}$$

a relationship we will use only for small values of k, but note that this factor, $(r' - r'' + 1)^k$, can be absorbed in the bound of derivatives of the coefficients a_u in (13.54).

The first time we remove a localizing function from an H^s norm, in (13.55) for instance, the couple of derivatives that will fall on Ψ will produce powers of r in view of (13.60), since initially $\rho = 0$. These will be balanced against $(r-1)!$, thanks to (13.65), with small powers of $(r' - r'' + 1)$, increasing C_h slightly in (13.55). We will see at the very end that the slightly different denominators in (13.60) will make little difference in the bounds.

Furthermore, upon the next iteration of (13.55), the new right-hand side *will have the same form*. That is, there will again be a product of lower-order terms (the same ones plus new ones), a second factor of a_u with derivatives that will give possibly another constant, C, in front of the supremum and another copy of C_h to its appropriate power, though these constants pass into the norms of the corresponding terms, just as in the treatment of powers of T above. But notice that the number of terms in the product increases at each pass (to at most r) and that the order of the top-order term decreases. Thus this sequence of constants will not contribute in the end more than C^r, which is also to be expected.

Handling the sum is as before as well, and we will not comment on it further except to recall (13.46).

When the last term on the right no longer has maximal order, we turn our attention to any of the other terms of highest order and proceed as before. The factorials have been adjusted so that the behavior that will in the end guarantee analyticity is that

$$\frac{||\Psi X^2 X^r u||_{H^s}}{(r-1)!} \leq C^{r+1} + C^{r/2} \frac{||X^2 T^{r/2} u||_{H^s(\mathcal{U}_1)}}{(r/2)!},$$

which will be the evident outcome of the repeated iterations of (13.55), taking the precise localizing functions into account, and which, together with the previous results on (nearly) pure T derivatives, will complete the proof.

We should remark at the end that what was true for the first localizing function, namely (13.65), will be a little different on the next pass, since the next localizing function may bring not a factor of $r - r'$ with each derivative it receives but rather the factor (cf. (13.60))

$$\frac{r - r'}{1 - \frac{1}{r}} = (r - r')\left(\frac{r}{r - 1}\right),$$

so that passing from $r - r'$ to $r - r' - t'$, we encounter instead of just

$$\frac{r}{r - r'} \leq r' - r'' + 1,$$

an extra factor of $r/r-1$, possibly to the $p(s)$th power; and this may keep occurring as the order of the leading term keeps decreasing. For instance, after a few iterations, the analogous 'extra' factors from (13.60) will be

$$\left(\frac{r}{r - 1}\right)\left(\frac{r - r_1}{r - r_1 - 1}\right)\left(\frac{r - r_1 - r_2}{r - r_1 - r_2 - 1}\right)\cdots$$

or even the $p(s)$th power of such a product. But there cannot be more than r terms in the product, and each factor is far less than 2, leading to an easily acceptable constant C^r in the end.

The same procedure works at any stage. We have already seen that expanding the term of highest order leads to a new product, but of the *same form* with one new norm of derivatives of a coefficient a_u, and the total number of terms, as with the T derivatives, never exceeds 4^r, which is certainly acceptable; and the factor $(r' - r'' + 1)^k$ just above is immediately attached to the $a_u^{(r'-r'')}$ that occurs with that product (cf.(13.53)).

This means that we may remove localizing functions from the H^s norms easily and replace them with new localizing functions, identically one on the support of the old one and supported in a larger open set such that a derivative of the new function is proportional to the number of derivatives still to be estimated in that term in the sense of (13.60). And while there will appear a number of copies of the (analytic) coefficients a, the sum of the number of derivatives they receive, and the powers of the corresponding constants arising from the algebraicity of H^s, is equal to the total decrease in derivatives on the terms of highest order taken step by step, which is also reflected in the number of norms in the product.

Thus the total number of derivatives appearing on coefficients will be, in the end, equal to the total number of terms in the product of norms, and since each contains a copy of u with one or two derivatives, this number is comparable to r. Thus this product will be bounded by $C_{u,h}^r$ for suitable $C_{u,h}$ depending only on the first couple of derivatives of u and on the coefficients a_u (and the dimension and the initial open sets).

We are not quite home. For high X derivatives, in addition to being "used up" as above, will also flow to half as many T derivatives, though in $\mathcal{U}_1$, due to the bracketing $[X, X] = T$ (cf. (13.51)), and the number of terms and the sums proceed exactly as in the estimation of T derivatives above, in ways that have nothing to do with the local versus global behavior. There appear new norms of derivatives of the coefficient functions, exactly as before, and one slightly new feature, which is the mixture between X and T derivatives, which is inevitable but has been seen before in many of the author's earlier works.

13.10 Local Regularity

We consider sums of squares of nonlinear vector fields, that is, equations such as

$$P(x, u, D)u = \sum_1^q X_j^2 u = f,$$

with the new feature that the $\{X_j\}$ may depend in their "coefficients" on the solution u. As a prime example of this class we consider the following case (for $r > 0$):

$$P_u(D)v := \left((D_x)^2 + (x^r D_t)^2 + (x^r \tilde{h}(x, t, u)D_t)^2\right) v, \quad (t, x) \in \mathbb{R}^2, \quad (13.66)$$

with $\tilde{h}$ real-valued and real analytic in its arguments.

We shall assume our solution u to be C^∞, since smoothness (starting from C^3) follows from the arguments of Xu [Xu1], which are based on the subelliptic estimate clearly satisfied by P_u and the paradifferential calculus of Bony [Bo].

13.11 Results

Theorem 13.3. *If f is real analytic near (x_0, t_0), then so is any smooth solution to (13.66).*

We remark that the problem is significant in its own right and also because it bears the same resemblance to general quasilinear subelliptic partial differential equations that the sums of squares of linear vector fields did to the subelliptic complexes and "boundary Laplacians" arising from the $\overline{\partial}_b$ operator in several complex variables.

In particular, the local real analytic hypoellipticity of those (in the linear case) with symplectic characteristic variety (roughly corresponding to $r = 1$ here), proved independently by Treves and Tartakoff in 1978 [T4], [T5], [Tr4], propels one quite reasonably to ask the same question in the quasilinear setting, of which the type of operator under study here is a simple prototype. (NB: the vector fields arising from $\overline{\partial}_{(b)}$ correspond more directly to $\partial_x - y\partial_t$, and $\partial_y + x\partial_t$ than to $\partial_x, \partial_y, x\partial_t$, and $y\partial_t$ as separate vector fields; nonetheless, the "Grushin-type" operators have always provided the most tractable models.)

13.12 Proof

Using standard arguments it is easy to prove the following a priori estimates: $\forall s \geq 0$, $u \in C^\infty$, and compact $\mathcal{U}_1$, $\exists C = C_{s,u,\mathcal{U}_1} : \forall\, v \in C_0^\infty(\mathcal{U}_1)$,

$$\sum_1^3 \|X_i v\|_s^2 + \|v\|_{s+\frac{1}{r+1}}^2 \leq C\left\{|(P_u v, v)_s| + \|v\|_s^2\right\}$$

and

$$\sum_{i,j=1}^3 \|X_i X_j v\|_s^2 + \sum_1^3 \|X_i v\|_{s+\frac{1}{r+1}}^2 + \|v\|_{s+\frac{2}{r+1}}^2 \leq C\left\{\|P_u v\|_s^2 + \|v\|_s^2\right\}, \quad (13.67)$$

where $\|\cdot\|_s = \|\cdot\|_{H^s}$, $P_u \equiv P(x, u, D)$, $X_1 = D_x$, $X_2 = x^r D_t$, $X_3 = X_3^{(u)} = x^r \tilde{h}(t, x, u) D_t$, and C depends only on the first $s + 3$ derivatives of u.

However, the estimate we will need uses the maximality of (13.67) rather than its subellipticity: with $X^I = X^{I_1} X^{I_2} \cdots X^{I_{|I|}}$ and $|||v|||_s$ defined by

$$|||v|||_s \equiv \sum_{|I|\leq 2} \|X^I v\|_s \quad (\text{and } |||v|||_{H^s(\mathcal{U})} \equiv \sum_{|I|\leq 2} \|X^I v\|_{H^s(\mathcal{U})}, \text{for } \mathcal{U} \text{ open}), \tag{13.68}$$

then for K arbitrary, $\exists C_K : \forall v \in C_0^\infty(\mathcal{U}_1)$,

$$|||v|||_2 + K \sum_{|I|\leq 1} \|X^I v\|_2 \leq C\|P_u v\|_2 + C_K \|v\|_0. \tag{13.69}$$

The general scheme, as always, will be to use the a priori estimate applied to functions $v = \varphi D^m u$ and then to bring φD^m to the left of P_u modulo errors that are handled inductively. Noting that the a priori estimate provides for maximal control (i.e. no loss Here derivatives) in the D_x direction, we limit ourselves to estimating $\varphi D_t^m u$. φ will be a smooth localizing function, namely identically equal to one in a fixed open set $\mathcal{U}_0$, where we wish to prove that the solution u is analytic, and supported in $\mathcal{U}_1$, the open set where the data are assumed to be real analytic. The localizing function $\varphi(x, t)$ may be taken to be of the form $\tilde{\varphi}(t)\tilde{\varphi}(x)$, and terms with derivatives on $\tilde{\varphi}(x)$ may be disregarded since the operator is elliptic for $x \neq 0$ where derivatives of $\tilde{\varphi}(x)$ would be supported.

We work in the *algebra* H^2 in order to handle large products of norms below. Taking $Pu = 0$ without loss of generality, we have from (13.69),

$$
\begin{aligned}
&|||\varphi D_t^m u|||_2 \\
&\quad = \|X_j^2 \varphi D_t^m u\|_2 + \cdots \le \|P\varphi D_t^m u\|_2 + \cdots \le \sum \|[X_j^2, \varphi D_t^m]u\|_2 + \cdots \\
&\quad \lesssim \sum_{k=1}^{2} \|g_k(u, D_t u)\varphi^{(k)} x^{2r} D_t^{m+2-k} u\|_2 + C\|\varphi x^{2r}[h(u), D_t^m u] D_t^2 u\|_2 + \cdots \\
&\quad \lesssim C\left(\sum_{k=1}^{2} \|\varphi^{(k)} X^2 D_t^{m-k} u\|_2 + \|\varphi x^{2r}[h(u), D_t^m u] D_t^2 u\|_2 + \cdots\right),
\end{aligned}
\tag{13.70}
$$

where $h(\cdot) \equiv \tilde{h}^2(\cdot)$ and we have estimated $\|g_k(u, D_t u)\|_{H^2(\mathcal{U}_1)}$ by a constant. Here the $g_k(u, D_t u)$ stand for the coefficients, aside from x^{2r}, that enter when φ is differentiated once or twice, and the dots "$\cdots$" denote terms arising from lower-order terms in the operator P, terms containing fewer X's.

For now we disregard the sum on the right and pay attention to the bracket in the last norm, the crucial one. To expand it, we need to use a different, and to us new, version of the classical formula of Faà di Bruno for derivatives of composite functions:

$$
D_t^\ell h(u(t)) = \sum_{\substack{b, r_j \ge 1 \\ \sum_1^b r_j = \ell}} (\ell - 1)! H_{b,\ell,\{r_j\}}(u) \prod_{j=1}^{b} \frac{u^{(r_j)}}{(r_j - 1)!},
\tag{13.71}
$$

where

$$
H_{b,\ell,\{r_j\}}(u) = \frac{h^{(b)}(u)}{(\ell - r_1)(\ell - r_1 - r_2)\cdots(\ell - r_1 - \cdots - r_{b-1})},
\tag{13.72}
$$

so that when $b = 1$, $H_{b,\ell,\{r_j\}}(u) = h'(u)$. See the previous section for the (elementary) derivation.

We don't use formula (13.71) in this nice form, but in the following (not so nice) one: for any $\varrho > 0$,

$$
D_t^\ell h(u) = \sum_{\substack{b, r_j \ge 1 \\ \sum_1^b r_j = \ell}} \frac{H_{b,\ell,\{r_j\}}(u)}{\ell} \binom{\ell}{r_1 \cdots r_b} \prod_{j,k=1}^{b} \left[\frac{\varrho r_j!}{(r_j + 1)^2}\right]\left[\frac{u^{(r_k)}(r_k + 1)^2}{\varrho (r_k - 1)!}\right].
\tag{13.73}
$$

The reason for introducing the squares in the denominators is that it is possible to choose ϱ so that

$$
\sum_{0 \le d \le \ell} \binom{\ell}{d} \frac{\varrho d!}{(d + 1)^2} \frac{\varrho (\ell - d)!}{(\ell - d + 1)^2} \le \frac{\varrho \ell!}{(\ell + 1)!},
\tag{13.74}
$$

hence such that

$$\sum_{\sum_1^b r_j(\geq 1)=\ell} \binom{\ell}{r_1 \cdots r_b} \prod \frac{\varrho r_j!}{(r_j+1)^2} \leq \frac{\varrho \ell!}{(\ell+1)!}. \tag{13.75}$$

The estimate (13.73) will be especially useful when we use derivatives of composite functions in conjunction with Leibniz's formula.

We fix ϱ such that (13.74), and hence (13.75), is valid. But before continuing to estimate (13.70), we remark that since h is analytic, it satisfies

$$|h^{(k)}(y)| \leq C_1 C_2^k k!, \quad y \in \overline{u(\mathcal{U}_1)}, \quad \forall k. \tag{13.76}$$

However, we claim that in our setting, C_2 may be taken arbitrarily small. To see this, we take

$$h_\delta(y) = h(\delta y), \quad u_\delta = \delta^{-1} u. \tag{13.77}$$

Then $P(u) = \delta P_\delta(u_\delta)$ with P_δ exactly like P but with h_δ instead of h. Obviously, the analyticity of u is equivalent to that of u_δ. Thus, with δ as small as needed (to be chosen at the end of the proof),

$$|h^{(k)}(y)| \leq C_1' \delta^k k!, \quad y \in \overline{u(\mathcal{U}_1)}, \quad \forall\, k. \tag{13.78}$$

Hence

$$\|h^{(k)}(u) w\|_2 \leq C_1'' \delta^k k! \|w\|_2, \quad \forall k, \quad w \in C_0^\infty(\mathcal{U}_1), \tag{13.79}$$

with C_1'' depending on $\|u\|_{H^2(\mathcal{U}_1)}$ but not on δ. Using this and (13.72), we have

$$\ell^{-1} \|H_{b,\ell,\{k_j\}}(u) w\|_2 \leq C_1'' \delta^b \|w\|_2 \quad \forall\, w \in C_0^\infty(\mathcal{U}_1), \tag{13.80}$$

because $\forall\, b \geq 1,\ \ell(\ell - r_1) \cdots (\ell - r_1 - \cdots - r_{b-1}) \geq b!$

Coming back to the bracket, using (13.73), and omitting now terms designated "$\cdots$" above, namely those arising from terms with fewer X's in the original operator, we have

$$\begin{aligned} \|\varphi x^{2r} [h(u), D_t^m u] D_t^2 u\|_2 &\leq \sum_{\ell=1}^m \binom{m}{\ell} \|\varphi x^2 (D_t^\ell h(u)) D_t^{m-\ell+2} u\|_2 \\ &\leq \sum_{\ell=1}^m \binom{m}{\ell} \sum_{\substack{b, r_j \geq 1 \\ \sum_1^b r_j = \ell}} \binom{\ell}{r_1 \cdots r_b} \prod_{j,k=1}^b \frac{\varrho r_j!}{(r_j+1)^2} \\ &\qquad \times \left\| \ell^{-1} H_{b,\ell,r_j}(u) \varphi x^{2r} \frac{\varrho^{-1} u^{(r_k)} (r_k+1)^2}{(r_k-1)!} D_t^{m-\ell+2} u \right\|_2. \end{aligned} \tag{13.81}$$

From here on, C may change from line to line but will be independent of m and u, and δ may also change into a fixed multiple of itself, but still may be chosen as small as necessary at the end. Thus, using (13.80) we may continue the inequality:

$$\leq \sum_{\ell=1}^{m} \binom{m}{\ell} \sum_{\sum_1^b r_j(\geq 1)=\ell} \binom{\ell}{r_1 \cdots r_b} \prod_1^b \frac{\varrho r_j!}{(r_j+1)^2} \frac{\varrho(m-\ell)!}{(m-\ell+1)^2}$$

$$\times C_1'' \delta^b \left\| \varphi x^{2r} \prod_1^b \frac{\varrho^{-1} u^{(r_k)}(r_k+1)^2}{(r_k-1)!} D_t^{m-\ell+2} u \frac{\varrho^{-1}(m-\ell+1)^2}{(m-\ell)!} \right\|_2 . \qquad (13.82)$$

Thanks to (13.75), the first line of the right-hand side is estimated by $\varrho \frac{m!}{(m+1)^2}$.

Concerning the last norm, we assume, renaming the indices if necessary, that the largest number of derivatives landing on a copy of u is $m-\ell+2$ and associate φx^{2r} with those derivatives. We read $\varphi x^{2r} D_t^{m-\ell+2} u$ as $\varphi X_2^2 D_t^{m-\ell} u$ and apply the algebra property of $H^2(\operatorname{supp} \varphi)$. So the last line above is estimated by

$$C''' \delta^b \frac{(m-\ell+1)^2}{(m-\ell)!} \|\varphi X^2 D_t^{m-\ell} u\|_2 \cdot \prod_1^b \frac{(r_j+1)^2}{(r_j-1)!} \|D_t^{r_j} u\|_{H^2(\operatorname{supp} \varphi)} . \qquad (13.83)$$

We note that $r_j < \frac{m}{2}, \ j = 1, \ldots, b$.

Summing up, from (13.69) and (13.81) we have (with X^2 denoting any product of *at most* two X's)

$$\begin{aligned} \frac{(m+1)^2}{m!} |||\varphi D_t^m u|||_2 \lesssim & \frac{(m+1)^2}{m!} \sum_{k=1}^{2} \|\varphi^{(k)} X^2 D_t^{m-k} u\|_2 \\ & + \sup_{\substack{b, r_j \geq 1 \\ \sum r_j = \ell \\ 0<\ell\leq m}} \frac{(m-\ell+1)^2}{(m-\ell)!} \|\varphi X^2 D_t^{m-\ell} u\|_2 \, \delta^b \\ & \times \prod_1^b \frac{(r_j+1)^2}{(r_j-1)!} \|D_t^{r_j} u\|_{H^2(\operatorname{supp} \varphi)}, \end{aligned} \qquad (13.84)$$

or equivalently,

$$\begin{aligned} \frac{(m+1)^2}{(m-1)!} |||\varphi D_t^m u|||_2 \lesssim & \frac{(m+1)^2}{(m-1)!} \sum_{k=1}^{2} \|\varphi^{(k)} X^2 D_t^{m-k} u\|_2 \\ & + \sup_{\substack{b, r_j \geq 1 \\ \sum r_j = \ell \\ 0<\ell\leq m}} \frac{(m-\ell+1)^2}{(m-\ell-1)!} \|\varphi X^2 D_t^{m-\ell} u\|_2 \, \frac{m\, \delta^b}{m-\ell} \\ & \times \prod_1^b \frac{(r_j+1)^2}{(r_j-1)!} \|D_t^{r_j} u\|_{H^2(\operatorname{supp} \varphi)} \end{aligned} \qquad (13.85)$$

13.13 Passing to Another Localizing Function

Canonically, the first localizing function, $\varphi \equiv \varphi_m$, satisfies

$$\varphi_m \equiv 1 \ \text{ on } \mathcal{U}_0, \quad \varphi_m \in C_0^\infty(\mathcal{U}_{1/m}), \quad |\varphi_m^{(k)}| \le c^k m^k, \quad k \le 4, \tag{13.86}$$

where we have set, for $a \ge 0$,

$$\mathcal{U}_a = \{(x,t) \in \mathcal{U}_1 \ : \ \operatorname{dist}((x,t), \mathcal{U}_0) < a(\operatorname{dist}(\mathcal{U}_0, \mathcal{U}_1^c))\}. \tag{13.87}$$

When a particular term in the estimate for $\varphi_m D_t^m u$ still has $D_t^{m-\ell} u$, with $m - \ell \ge m/2$, then we shall localize $D_t^{m-\ell} u$ by a function $\varphi_{m,m-\ell}$ satisfying

$$\varphi_{m,m-\ell} \equiv 1 \quad \text{on } \mathcal{U}_{1/m}, \quad \varphi_{m,m-\ell} \in C_0^\infty(\mathcal{U}_1). \tag{13.88}$$

Now we are almost in position to take $\varphi = \varphi_m$ out of the norm in (13.85) and specify the new function $\varphi_{m,m-\ell}$. In fact, we have to make two remarks first, and will treat the case $m - \ell < m/2$ below.

Remark. While it is true that we could just write $\|\varphi w\|_s \le c\|\varphi\|_s \|w\|_s$, for $s \ge 2$, to do so would incur at least two derivatives on φ with no gain on w. To avoid this difficulty, we use the following finer estimates of the H^2 norm of a product of functions.

If $\varphi, \tilde{\varphi}$ are two smooth, compactly supported functions with $\tilde{\varphi} \equiv 1$ on supp φ, then for every $p \ge 2$,

$$\|\varphi D^p u\|_2 \le C^2 \sup_{q \le 2} \|D^q \varphi\|_{L^\infty} \|\tilde{\varphi} D^{p-q} u\|_2 \tag{13.89}$$

and

$$\|\varphi D^p u\|_2 \le C^2 \sup_{q \le 2} \|D^q \varphi\|_{L^\infty} \|D^{p-q} u\|_{H^2(\operatorname{supp}\,\varphi)}. \tag{13.90}$$

Remark 2. Since $m = (m - \ell) + r_1 + \cdots + r_b$ with $m - \ell \ge r_j, \ \forall j$, implies $b + 1 \ge \frac{m}{m-\ell}$, i.e., $1 \le (b+1) \cdot \frac{m-\ell}{m}$, we have:

$$\delta^b \le c_0\, \tilde{\delta}^b \left(\frac{m-\ell}{m}\right)^k, \quad 0 \le k \le 2, \tag{13.91}$$

where $\tilde{\delta} = 2\delta, \ c_0 = \sup_{b \ge 1} 2^{-b}(b+1)^2$.

Claim: There exist two constants C and A such that

$$\sup_{r \le \frac{m}{2}} \frac{(r+1)^2}{(r-1)!} \frac{|||D_t^r u|||_{H^2(\mathcal{U}_{1/2})}}{C^r} \le A \tag{13.92}$$

implies

$$\sup_{r\le m} \frac{(r+1)^2}{(r-1)!}\,\frac{|||D_t^r u|||_{H^2(\mathcal{U}_0)}}{C^r} \le A. \tag{13.93}$$

Proof. Using (13.91) with $k = 1$, we at first rewrite (13.85) with $\tilde{\delta}^b$ in place of $\delta^b \frac{m}{m-\ell}$ and with a new constant C_0. Then assuming (13.92) for constants $C, A > 1$ to be chosen, we have (writing X^2 for any product of *at most* two X's)

$$\frac{(m+1)^2}{(m-1)!}\,\frac{|||\varphi_m D_t^m u|||_2}{C^m} \le \frac{C_0}{C}\Bigg(\frac{(m+1)^2}{(m-1)!}\sum_{1\le k\le 2}\frac{\|\varphi_m^{(k)} X^2 D_t^{m-k} u\|_2}{C^{m-k}} + \sup_{\substack{b, r_j\ge 1\\ \sum_1^b r_j=\ell\\ 0<\ell<m}} \frac{(m-\ell+1)^2}{(m-\ell-1)!}\,\frac{\|\varphi_m X^2 D_t^{m-\ell} u\|_2}{C^{m-\ell}}\,\tilde{\delta}^b A^b\Bigg). \tag{13.94}$$

Now we use (13.89) to estimate the first term on the right-hand side in (13.94):

$$\begin{aligned}
&\frac{(m+1)^2}{(m-1)!}\sum_{1\le k\le 2}\frac{\|\varphi_m^{(k)} X^2 D_t^{m-k} u\|_2}{C^{m-k}}\\
&\le 2\sup_{\substack{1\le k\le 2\\ 0\le q\le 2}} \frac{(m+1)^2}{(m-1)!}(cm)^{k+q}\,\frac{\|X^2\varphi_{m,m-k-q} D_t^{m-k-q} u\|_2}{C^{m-k-q}}\\
&\le C_1 \sup_{1\le k\le 4}\frac{(m-k+1)^2}{(m-k-1)!}\,\frac{|||\varphi_{m,m-k} D_t^{m-k} u|||_2}{C^{m-k}}.
\end{aligned} \tag{13.95}$$

For the second term on the right-hand side in (13.94), we distinguish two cases. If $m-\ell > \frac{m}{2}$, we argue as above (which is permissible, since for $m \ge 5$ we have $1 \le \frac{m}{m-\ell-q} \le 10$, for every $1 < \ell \le \frac{m}{2}$, $q < 2$), so we obtain

$$\begin{aligned}
&\frac{(m-\ell+1)^2}{(m-\ell-1)!}\,\frac{\|\varphi_m X^2 D_t^{m-\ell} u\|_2}{C^{m-\ell}}\tilde{\delta}^b A^b\\
&\le C_2 \sup_{k\le 2}\frac{(m-\ell-k+1)^2}{(m-\ell-k-1)!}\,\frac{|||\varphi_{m,m-\ell-k} D_t^{m-\ell-k} u|||_2}{C^{m-\ell-k}}\,\tilde{\delta}^b A^b.
\end{aligned} \tag{13.96}$$

If, on the other hand, $m-\ell \le \frac{m}{2}$, we have, from (13.90),

$$\|\varphi_m X^2 D_t^{m-\ell} u\|_2 \le \sup_{q\le 2}(Cm)^q\,||X^2 D_t^{m-\ell-q} u||_{H^2(\mathcal{U}_{1/2})}, \tag{13.97}$$

and now invoking the (simple but very useful) estimate (13.91) yields

$$\frac{(m-\ell+1)^2}{(m-\ell-1)!}\,\frac{\|\varphi_m X^2 D_t^{m-\ell} u\|_2}{C^{m-\ell}}\tilde{\delta}^b A^b$$
$$\le c_0\,\tilde{\tilde{\delta}}^b \sup_{q\le 2}(m-\ell)^q\,\frac{(m-\ell+1)^2}{(m-\ell-1)!}\,\frac{||X^2 D_t^{m-\ell-q} u||_{H^2(\mathcal{U}_{1/2})}}{C^{m-\ell-q}}A^b$$
$$\le \left(\text{arguing as above, since } \frac{(m-\ell-q+1)^2}{(m-\ell-q-1)!}\sim (m-\ell)^q\,\frac{(m-\ell+1)^2}{(m-\ell-1)!}\right)$$
$$\le C_3\tilde{\tilde{\delta}}^b A^{b+1},\quad \tilde{\tilde{\delta}} = 2\tilde{\delta} = 4\delta. \tag{13.98}$$

Summing up, from (13.94), (13.95), (13.96), and (13.98) we have

$$\frac{(m+1)^2}{(m-1)!}\,\frac{|||\varphi_m D_t^m u|||_2}{C^m}$$
$$\le \frac{C_0}{C}C_3\tilde{\tilde{\delta}}^b A^{b+1} + \frac{C_0}{C}\big(C_1 + C_2 A^b\tilde{\delta}^b\big)\sup_{1\le\ell<\frac{m}{2}}\frac{(m-\ell+1)^2}{(m-\ell-1)!}\,\frac{|||\varphi_{m,m-\ell} D_t^{m-\ell}u|||_2}{C^{m-\ell}}. \tag{13.99}$$

Now we choose

$$\tilde{\tilde{\delta}} = A^{-2},\ \text{i.e.,}\ \delta = (2A)^{-2} \equiv \delta(A),\quad C = C_0(C_1 + C_2 + C_3), \tag{13.100}$$

so we obtain

$$\frac{(m+1)^2}{(m-1)!}\,\frac{|||\varphi_m D_t^m u|||_2}{C^m} \le \max\left\{\sup_{1\le\ell<\frac{m}{2}}\frac{(m-\ell+1)^2}{(m-\ell-1)!}\,\frac{|||\varphi_{m,m-\ell} D_t^{m-\ell}u|||_2}{C^{m-\ell}}, 1\right\}.$$

13.14 Reducing the Order by Half; the End of the Proof

We have reduced the order by at least one. If we are able to start again and argue for $\varphi_{m,m-\ell}D_t^{m-\ell}u$ as above for $\varphi_m D_t^m u$, and continue with localizing functions, all with their supports in $\mathcal{U}_{1/2}$, until the order is reduced to $\frac{m}{2}$, we shall obtain (again X^2 standing for *at most* two X's)

$$\frac{(m+1)^2}{(m-1)!}\,\frac{||X^2 D_t^m u||_{H^2(\mathcal{U}_0)}}{C^m} \le \max\left\{\sup_{r\le\frac{m}{2}}\frac{(r+1)^2}{(r-1)!}\,\frac{||X^2 D_t^r u||_{H^2(\mathcal{U}_{1/2})}}{C^r}, 1\right\}; \tag{13.101}$$

that is, our claim will be proved, with $A \ge 1$.

Now to do this, we take $\varphi_{m,m-\ell} \equiv 1$ on $\mathcal{U}_{1/m}$ and $C_0^\infty\left(\mathcal{U}_{\frac{1}{m}+\frac{1}{m-\ell}(1-\frac{1}{m})}\right)$,

$$|\varphi^{(k)}_{m,m-\ell}| \leq c\left((m-\ell)\left(1+\frac{1}{m+1}\right)\right)^k, \quad k \leq 4, \tag{13.102}$$

and then

$$\varphi_{m,m-\ell,m-\ell-\ell_1} \equiv 1 \ \text{ on } \ \mathcal{U}_{\frac{1}{m}+\frac{1}{m-\ell}(1-\frac{1}{m})} \tag{13.103}$$

and

$$\varphi_{m,m-\ell,m-\ell-\ell_1} \in C_0^\infty\left(\mathcal{U}_{\frac{1}{m}+\frac{1}{m-\ell}(1-\frac{1}{m})+\frac{1}{m-\ell-\ell_1}(1-(\frac{1}{m}+\frac{1}{m-\ell}(1-\frac{1}{m})))}\right) \tag{13.104}$$

and so on until the order of derivatives becomes $\leq \frac{m}{2}$.

Denote by p the number of iterations required for this and call $\ell_0 = \ell$,

$$m(+2) - \sum_0^p \ell_j \leq \frac{m}{2}, \tag{13.105}$$

and let $W_p = a$ if $\mathcal{U}_a$ is the support of the pth localizing function, and so $1 - W_p$ is the remaining width (of the complement of $\mathcal{U}_1$). Computing W_p, we obtain

$$W_p = 1 - \prod_{p' \leq p}\left(1 - \frac{1}{m - \sum_0^{p'} \ell_j}\right), \tag{13.106}$$

and it is not too hard to prove (see the next section) that

$$W_p \leq \frac{1}{2} \quad \text{for } \ m \geq 6, \tag{13.107}$$

so our claim is proved with

$$A = \max\left\{\sup_{r \leq 6} \frac{(r+1)^2}{(r-1)!}||X^2 D_t^r u||_{H^2(\mathcal{U}_1)}, 1\right\}, \tag{13.108}$$

and the analytic regularity of u in $\mathcal{U}_0$ is proved.

13.15 The Width of the Critical Band

We want to compute the maximum of $1 - \Pi_{p' \leq p}\big(1 - \frac{1}{m - \sum_0^{p'} \ell_j}\big)$, i.e., the minimum of $\Pi_{p' \leq p}\big(1 - \frac{1}{m - \sum_0^{p'} \ell_j}\big)$. This happens when the individual fractions in the product

$$\left(\frac{m-1}{m}\right)\left(\frac{m-\ell_1-1}{m-\ell_1}\right)\cdots\left(\frac{m-\sum_0^p \ell_j-1}{m-\sum_0^p \ell_j}\right)$$

are as small as possible. If $\sum_0^p \ell_j = m/2$, then for a given p, we should take $\ell_p = \ell_{p-1} = \cdots = \ell_2 = 1$ and $\ell_1 = \frac{m}{2} - (p-1)$, giving the product

$$\left(\frac{m-1}{m}\right)\left(\frac{m-(\frac{m}{2}-(p-1))-1}{m-\frac{m}{2}-(p-1)}\right)\left(\frac{m-(\frac{m}{2}-(p-1))-2}{m-(\frac{m}{2}-(p-1))-1}\right)\cdots\left(\frac{\frac{m}{2}-1}{m-\frac{m}{2}}\right).$$

Here many, many things cancel, leaving only

$$\left(\frac{m-1}{m}\right)\left(\frac{\frac{m}{2}-1}{m-\frac{m}{2}-(p-1)}\right). \tag{13.109}$$

Consider the minimum when $p = 1$: The minimum of $\frac{(m-1)(m-2)}{m^2}$ occurs for $m = 4/3$; this function is increasing after that, and for $m \geq 5.3$,

$$\frac{(m-1)(m-2)}{m^2} \geq \frac{1}{2}.$$

For other values of p, it is clear that $W_p < W_1$, and hence at least half of the total width always remains. Further, for the new localizing function, $\varphi_{m,m-\ell_1,\ldots,m-\sum_0^p}$, one computes without difficulty that each derivative is essentially $m - \sum_0^p \ell_j \sim m/2$.

13.16 The Faà di Bruno Formula

The derivation of our version of the formula of Faà di Bruno (of which there are many) is completely natural (to us!):

$$\begin{aligned}
D_t^m h(u(t)) &= D_t^{m-1}(u'(t)h'(u(t))) \\
&= \sum_{0\leq m_1\leq m-1} \binom{m-1}{m_1} u^{(m_1+1)} D_t^{(m-1-m_1)} h'(u(t)) \\
&= \sum_{1\leq \ell_1\leq m} \binom{m-1}{\ell_1-1} u^{(\ell_1)} D_t^{(m-\ell_1)} h'(u(t)) \\
&= \sum_{1\leq \ell_1\leq m} \binom{m-1}{\ell_1-1} u^{(\ell_1)} D_t^{m-\ell_1-1}(u'(t)h''(u(t))) \\
&= \sum_{1\leq \ell_1\leq m} \binom{m-1}{\ell_1-1} \sum_{0\leq m_2\leq m-\ell_1-1} \binom{m-\ell_1-1}{m_2} \\
&\quad \times u^{(\ell_1)} u^{(m_2+1)} D_t^{m-\ell_1-1-m_2} h''(u(t))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{1\le \ell_1\le m} \binom{m-1}{\ell_1-1} \sum_{1\le \ell_2\le m-\ell_1} \binom{m-\ell_1-1}{\ell_2-1} u^{(\ell_1)} u^{(\ell_2)} D_t^{m-\ell_1-\ell_2} h''(u(t)) \\
&= \cdots = \\
&= \sum_{b,\ell_j\ge 1, \sum_1^b \ell_j = m} (m-1)!\tilde{H}_{b,m,\{\ell_j\}} \prod_1^b \frac{u^{(\ell_j)}}{(\ell_j-1)!},
\end{aligned}
$$

where

$$
\tilde{H}_{b,m,\{\ell_j\}} = \frac{h^{(b)}}{(m-\ell_1)(m-\ell_1-\ell_2)\cdots(m-\ell_1-\cdots-\ell_{b-1})}.
$$

Chapter 14
Treves' Approach

In 1978, simultaneously with our result, F. Treves published a very different proof of analytic hypoellipticity in the nondegenerate case.

The outline of the proof is natural enough: using well-known methods, the problem is reduced to a pseudodifferential one that lives in the boundary, and then a more general result is proved for a class of systems of pseudodifferential operators with scalar principal part $P = I_d P_2 + B_1$, where P_2 and B_1 are (classical, analytic) pseudodifferential operators of order 2 and 1, respectively, P_2 is a scalar, and I_d is the $d \times d$ identity matrix; the principal symbol $\sigma(P_2)$ of P is assumed to satisfy:

- Its principal symbol is nonnegative and vanishes to exactly order two on its characteristic variety Char P.
- Char P is a symplectic analytic submanifold of the cotangent bundle.
- The whole operator P is hypoelliptic with loss of one derivative.

The principal example of such systems in their own right is the complex boundary Laplacian $\Box_b$, for which the hypotheses are rapidly shown to reduce to Kohn's condition $Y(q)$ (on (p, q)-forms) and the hypothesis that the Levi form is nondegenerate.

Treves proves the following results:

Theorem 14.1. *Let the system P satisfy (1) and (2) and be hypoelliptic with loss of one derivative; then P is microlocally analytic hypoelliptic.*

Corollary 14.1. *The $\bar{\partial}$-Neumann problem is microlocally analytic hypoelliptic (at the boundary) provided the Levi form satisfies Kohn's condition $Z(q)$ and is nondegenerate.*

Corollary 14.2. *The complex boundary Laplacian $\Box_b$ is microlocally analytic hypoelliptic provided the Levi form satisfies Kohn's condition $Y(q)$ and is nondegenerate.*

In slightly more detail, the proof proceeds as follows.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_14,

After the reduction to the boundary, of dimension $n + 1$, the proof begins by using Sato's theorem to reduce the problem to consideration of the operator

$$P_S = I_d \sum_{i,j=1}^{n} a_{ij} Y_i Y_j^* + B,$$

where the a_{ij} form an $n \times n$ matrix of pseudodifferential operators whose principal symbol is self-adjoint positive definite and of order zero, while B is a matrix of firs-order pseudodifferential operators and the Y_j are of the particularly simple form $Y_j = \partial/\partial y_j - y_j \partial/\partial x$, where x is written for y_{n+1}. Thus the Y_j and their adjoints should be thought of as coming from the portion of the holomorphic structure of the ambient space that is tangent to the boundary, and y_{n+1} from the "normal" variable.

After this reduction, the author studies the model "Grušin" operator

$$P_G = I_d \sum_{i,j=1}^{n} \tilde{a}_{ij} L_i L_j^* + \tilde{B},$$

where now the $\tilde{a}_{ij}$ are complex scalars and $\tilde{B}$ a complex matrix with A positive definite and self-adjoint, and

$$L_j = \partial/\partial y_j - y_j, \quad j \leq n.$$

The assumption that -2 times no positive integral linear combination of the eigenvalues of a_{ij} should be an eigenvalue of B at a reference point will ensure that P_G is invertible and that in fact the inverse has an integral kernel. And this condition is clearly satisfied whenever the self-adjoint part of B is positive definite, since the eigenvalues of A are all positive definite. While the bulk of the paper makes this assumption of positivity for the self-adjoint part of B, in the end it is relaxed to the previous condition.

This operator P_G is next inverted by means of an integral representation arising from a heat kernel and analyzed very carefully in the major part of the paper, with approximation theorems that finally yield a parametrix to the original problem which lives on the boundary and hence to the original boundary value problem.

However daunting, this method has the advantage that it produces a true parametrix, which yields, potentially, more information about the solution than a proof that estimates higher and higher derivatives of the solution and shows that they force the solution itself to be a real analytic function.

The method has the weakness that it may not be particularly flexible, and if any of the hypotheses is weakened it will fail.

At some level the two approaches may well be equivalent, that of Treves being perhaps more geometric but also more delicate in treating asymptotic series in the analytic category, while ours requires successive corrections in the spirit of a (noncommutative) Taylor series that has its own symplectic formulation (cf. [T5]).

It is true that when the coefficients are pseudodifferential, our method increases in complexity. Even nonrigid, variable, but not pseudodifferential coefficients pose real problems, since the proper balance in our Taylor-like series is not immediately maintained, and only by invoking an identity in binomial coefficients can it be seen that the poor-looking unbalanced terms can be written as sums of balanced terms and the proof can proceed. This binomial identity seems to be standard, however, and appears in Feller's book [Fel].

As mentioned earlier, Métivier has followed Treves' approach to prove a similar result for pseudodifferential operators P of order m whose characteristic set Σ is a real analytic submanifold on which the principal symbol vanishes exactly to order k (and lower-order symbols vanishing appropriately). Such an operator P must be subelliptic with loss of $k/2$ derivatives. Then P is AHE, provided Σ is symplectic.

We note that in a previous chapter we have applied our method to this situation as well, obtaining an "elementary" proof of Métivier's result.

Chapter 15
Appendix

15.1 A Discussion of the Localizing Functions

We have already discussed the family of localizing functions we use. Here we present some remarks that may elucidate their usefulness and the subtlety of the refinements we introduced in [T4].

We recall the definition:

Given two open sets $\Omega_1 \Subset \Omega_2$ in $\mathbb{R}^n$ and an integer N, we will construct a function $\Psi_N \in C_0^\infty(\Omega_2)$ with $\Psi_N \equiv 1$ on Ω_1 and such that for some constant C independent of the separation d between Ω_1 and Ω_2,

$$|D^\alpha \Psi_N| \leq C^{|\alpha|+1} d^{-|\alpha|} N^{|\alpha|}, \quad |\alpha| \leq 3N.$$

By scaling, it suffices to ignore the distance d, and then Ψ_N is obtained by taking an open set intermediate between the Ω_j and convolving its characteristic function with N copies of a standard (nonnegative) bump function of integral one, but of support proportional to $1/N$. In differentiating such a convolution up to $3N$ times, at most three derivatives will fall on any one bump function, and the support properties of Ψ_N are easily verified.

Finally, in view of Stirling's formula, when $|\alpha| = N$,

$$C^{N+1} d^{-N} N^{|\alpha|} \sim C^{N+1} d^{-N} N! \sim C' C'^N N!$$

with C' independent of N.

In the body of the text, in fact in each chapter, we found it necessary to nest many of these open sets (denoted by $\omega_j, 1 \leq j \leq \log_2 N$ for N given once and for all, with their corresponding functions $\varphi_j \equiv 1$ on $\overline{\omega_j}$ but in $C_0^\infty(\omega_{j+1})$, and we denote by d_j the distance from $\overline{\omega_j}$ to the complement of ω_{j+1}).

We have seen that this can be done, by taking a suitable convolution of $N_j = N/2^j$ copies of a standard "bump" function with the characteristic function of an

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22,
DOI 10.1007/978-1-4419-9813-2_15, © Springer Science+Business Media, LLC 2011

open set containing ω_j but contained in ω_{j+1} and whose boundary lies halfway between the boundary of ω_j and the complement of ω_{j+1}, and such that

$$|D^\alpha \varphi_j| \le (N_j/d_j)^{|\alpha|}, \quad |\alpha| \le 3N_j.$$

The factor of 3 is merely for convenience.

The reason that each N_j is one-half of the previous one is that each full set of iterations of the a priori estimate reduces the power of T by at least one-half; hence the next time at most half the number of derivatives is even present in the estimates, which means that the next φ need accept only half the number of derivatives of the previous one.

However, it is clearly necesary that the sum of the d_j be finite. In fact the sum will be equal to the (finite) separation d_0 between the initial open set where we will prove that the solution is analytic to the complement of the open set where the data are assume to be real analytic functions.

That is, we require that

$$\sum_j d_j = d_0$$

and that in the extreme case,

$$\prod_{j=0}^{\log_2 N} (N_j/d_j)^{N_j} = \prod_{j=0}^{\log_2 N} \left(\frac{N/2^j}{d_j}\right)^{N/2^j} \le C^N N^N$$

for some constant C (this is the extreme case in which all derivatives land on the localizing functions).

But there is one additional condition, buried in the estimates in places such as the definition of d_j above, which is to ensure that at each stage, long after N has been replaced by $N_j = N/2^j$, we may estimate

$$N_j^{N_j} \le C^{N_j} N_j!,$$

so that after we have done this $\log_2 N$ times, the resulting product satisfies

$$\prod C^{N_j} \sim C^N,$$

If the corresponding left hand side is N^{N_j}, estimating

$$N^{N_j} \le C^N N_j!,$$

The tempting choice $d_j = d_0/2^j$ will actually work here, as observed earlier; even though at each level the equivalence does not seem sufficiently uniform, in the ultimate product it will be (cf. (4.20), ..., (5.21)).

15.2 The Analytical Material Used

15.2.1 Some Fourier Analysis and Sobolev Spaces

For a smooth function $f(x)$, the Fourier transform is defined by

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x)\, dx.$$

The L^2 Sobolev spaces H^s may be defined as the completion of C_0^∞ with respect to the norm

$$\|v\|_s = \left(\int_{\mathbb{R}^n} (1 + |\xi|^2)^s |\hat{v}(\xi)|^2\, d\xi \right)^{1/2}$$

or the space of tempered distributions v for which $\|v\|_s$ is finite.

It is not hard to conclude the celebrated Sobolev embedding theorem, that locally, $f \in C^r$, provided $f \in H^{r+s}$ for some $s > n/2$, using the Schwarz inequality, and thus that locally, $C^\infty = \cap_s H^s$.

15.2.2 The Heisenberg Group

As mentioned earlier, we do not actually need the Heisenberg group in our work. What suffices is the particular vector fields mentioned above. However, for completeness we mention that the group law on $\mathbb{R}^{2n+1}$ given by

$$(x_1, y_1, t_1) \cdot (x_2, y_2, t_2) = \left(x_1 + x_2, y_1 + y_2, t_1 + t_2 + \frac{1}{2}(x_1 y_2 - x_2 y_1) \right).$$

yields what is known as the Heisenberg group, and a basis of vector fields that are invariant under translation by the group action is the set of vector fields cited above.

15.2.3 Pseudodifferential Operators

Just as a partial differential operator

$$P(x, D) = \sum_{|\alpha| \le m} a_\alpha D^\alpha, \quad D = \frac{1}{i} \frac{\partial}{\partial x},$$

may be defined in terms of its "symbol" $p(x,\xi) = \sum_{|\alpha|\le m} a_\alpha \xi^\alpha$,

$$\widehat{P(x,D)u}(\xi) = p(x,\xi)\hat{u}(\xi),$$

so more general "symbols" $p(x,\xi)$, which may be sums of terms homogeneous of decreasing degrees in ξ (or even asymptotic sums of terms of orders going to $-\infty$), define operators by the same formula and obey a calculus similar to that of partial differential operators.

By the principal symbol $p_m(x,\xi)$ of a pseudodifferential operator $P_m(x,D)$ of degree m we mean a smooth function $p_m(x,\xi)$, homogeneous of degree m in ξ for $|\xi| \ge 1$ and such that locally in x,

$$|D_x^\alpha D_\xi^\beta p_m(x,\xi)| \le C^{|\alpha|+|\beta|+1}(1+|\xi|)^{m-|\beta|}, \quad |\xi| \ge 1.$$

The calculus yields then that if the p_{m_j} have degree m_j, $j = 1,2$, then $P_{m_1}P_{m_2}$ has degree $m_1 + m_2$ with leading symbol $p_{m1}p_{m_2}$ and hence that the bracket $P_{m_1}P_{m_2} - P_{m_2}P_{m_1}$ has "leading symbol" identically equal to zero, and hence $[P_{m_1}, P_{m_2}]$ has order at most $m_1 + m_2 - 1$.

We will mostly use constant-coefficient pseudodifferential operators, generally powers of a Laplacian in one or more variables. In particular, in $\mathbb{R}^{n+1}$ with coordinates (x,t), the operator Λ_t with symbol $(1+|\tau|^2)^{1/2}$ will commute with functions of the other variables, while the operator Λ with symbol $\lambda = (1+|\xi|^2 + |\tau|^2)^{1/2}$ in obvious notation provides a natural bijection from H^s to H^{s-1} whose inverse has symbol $1/\lambda$.

References

[A] K.G. ANDERSSON. *Propagation of Analyticity of Solutions of Partial Differential Equations with Constant Coefficients.* Ark. för Matematik, **8** (1970), 277–302.

[Ash] M.E. ASH. *The Basic Estimate of the $\overline{\partial}$-Neumann Problem in the Non-Kählerian Case*, Amer. J. Math., **86** (1964), 247–254.

[BG] M.S. BAOUENDI AND C. GOULAOUIC. *Nonanalytic-Hypoellipticity for Some Degenerate Elliptic Operators*, Bull. Amer. Math. Soc., **78** (1972), 483–486.

[BM] D. BELL AND S. MOHAMMED. *An extension of Hörmander's Theorem for Infinitely Degenerate Differential Operators,* Duke Math. J. **78** (1995), 453–475.

[BCR] C. BOLLEY, J. CAMUS, AND L. RODINO. *Hypoellipticité Analytique-Gevrey et Itérés d'Opérateurs.* Rend. Sem. Mat. Univ. Politec. Torino **45** (3) (1987), 1–61.

[Bo] J.-M. BONY. *Calcul Symbolique et Propagation des Singularités pour les Équations aux Dérivées Partielles Non Linéaires*, Ann. Scientifiques de l'École Normale Supérieure (4) **14** (1981), 209–246.

[B1] L. BOUTET DE MONVEL. *Hypoelliptic Operators with Double Characteristics and Related Pseudo-differential Operators*, Comm. Pure Appl. Math., **27** (1974), 585–639.

[B2] L. BOUTET DE MONVEL. *Intégration des Équations de Cauchy–Riemann Induites Formelles.* Séminaire Goulaouic–Lions–Schwartz 1974–1975; Équations aux dérivées partielles linéaires et non linéaires, Exp. No. 9, 14 pp. Centre Math., École Polytech., Paris, 1975.

[BGH] L. BOUTET DE MONVEL, A. GRIGIS, AND B. HELFFER. *Paramétrixes d'Opérateurs Pseudo-différentiels à Caractéristiques Multiples*, Astérisque, **34-35** (1976), 93–121.

[BK] L. BOUTET DE MONVEL AND P. KRÉE. *Pseudodifferential Operators and Gevrey Classes.* Ann. Inst. Fourier Grenoble, **17** (1967), 295–323.

[BDKT] A. BOVE, M. DERRIDJ, J.J. KOHN, AND D.S. TARTAKOFF. *Sums of Squares of Complex Vector Fields and (Analytic-) Hypoellipticity,* Math. Res. Lett. **13** no. 5-6 (2006), 683–701.

[BDT] A. BOVE, M. DERRIDJ, AND D.S. TARTAKOFF. *Analyticity in the Presence of Non Symplectic Characteristic Points*, J. Funct. Anal., **234** (2006), 464–472.

[BT1] A. BOVE AND D.S. TARTAKOFF. *Propagation of Gevrey Regularity for a Class of Hypoelliptic Equations,* Transactions of the A.M.S. **348** (7) (1996), 2533–2575.

[BT2] A. BOVE AND D.S. TARTAKOFF. *Optimal Non-isotropic Gevrey Exponents for Sums of Squares of Vector Fields*, Comm. Partial Differential Equations, **22** (7-8) (1997), 1263–1282.

[BT3] A. BOVE AND D.S. TARTAKOFF. *Analytic Hypo-ellipticity at Non-symplectic Characteristics When the Symplectic Form Changes Its Rank,* in preparation.

[BT4] A. BOVE AND D.S. TARTAKOFF. *Gevrey Hypoellipticity for Non-subelliptic Operators*, Pure and Applied Mathematics Quarterly, **6** (3) (2010), 663–675.

D.S. Tartakoff, *Nonelliptic Partial Differential Equations: Analytic Hypoellipticity and the Courage to Localize High Powers of T*, Developments in Mathematics 22, DOI 10.1007/978-1-4419-9813-2, © Springer Science+Business Media, LLC 2011

[BTr] A. BOVE AND F. TREVES, *On the Gevrey Hypo-ellipticity of Sums of Squares of Vector Fields*, Ann. Inst. Fourier Grenoble, **54** (2004), 1443–1475.

[Che1] S.C. CHEN. *Global Real Analyticity of Solutions to the $\overline{\partial}$-Neumann Problem on Reinhardt Domains,* Indiana U. Math. J. **37** (2) (1988), 421–430.

[Che2] S.C. CHEN. *Global Analytic Hypoellipticity of $\Box_b$ on Circular Domains,* Pac. J. Math. **148** (2) (1991) 225–235.

[Chr1] M. CHRIST. *Certain Sums of Squares of Vector Fields Fail to Be Analytic Hypoelliptic,* Comm. in P.D.E. **10** (1991), 1695–1707.

[Chr2] M. CHRIST. *Intermediate Gevrey Exponents Occur,* Comm. in P.D.E. **22** (1997).

[Chr3] M. CHRIST. *The Szegő Projection Need Not Preserve Global Analyticity,* Annals of Math. **143** (1996), 301–330.

[Chr4] M. CHRIST. *A Remark on Sums of Squares of Complex Vector Fields,* preprint, arXiv:math.CV/0503506.

[Chr5] M. CHRIST. *Hypoellipticity in the Infinitely Degenerate Regime,* Complex analysis and geometry, de Gruyter.

[Chr6] M. CHRIST. *Global C^∞ Irregularity of the $\overline{\partial}$-Neumann Problem for Worm Domains,* J. Amer. Math. Soc. **9** (4) (1996), 1171–1185.

[CG] M. CHRIST AND D. GELLER. *Counterexamples to Analytic Hypoellipticity for Domains of Finite Type*, Annals of Math. **135** (1992), 551–566.

[CH] PAULO D. CORDARO AND N. HANGES. *Impact of Lower Order Terms on a model PDE in Two Variables*, Contemp. Math, **368** (2005), 157–176.

[D1] M. DERRIDJ. *Sur une Classe d'Opérateurs Différentiels Hypoelliptiques à Coefficients Analytiques,* Sem. Goulaouic–Schwartz, 1970–1971, Équations aux dérivées partielles et analyse fonctionelle, Exp. no. 12, 6 pp., Centre de Math., École Polytechnique, Paris (1971).

[D2] M. DERRIDJ. *Sur la Régularité Gevrey jusqu'au Bord des Solutions du Problème de Neumann pour $\overline{\partial}$*, Proc. Symp in Pure Math **30** (1) (1977), 123–126.

[D3] M. DERRIDJ. *Régularité pour $\overline{\partial}$ dans Quelques Domaines Faiblement Pseudoconvexes*, Journal of Differential Geometry **13** (4) (1978), 559–576.

[DT1] M. DERRIDJ AND D.S. TARTAKOFF. *Local Analyticity for $\Box_b$ and the $\overline{\partial}$-Neumann Problem at Certain Weakly Pseudo-convex Points,* Comm. P. D. E. **13** (12) (1988), 1521–1600.

[DT2] M. DERRIDJ AND D.S. TARTAKOFF. *Local Analyticity in the $\overline{\partial}$-Neumann Problem and for $\Box_b$ in Some Model Domains without Maximal Estimates,* Duke Mathematical Journal **64** (2) (1991), 377–402.

[DT3] M. DERRIDJ AND D.S. TARTAKOFF. *Global Analyticity for $\Box_b$ on Three Dimensional Pseudoconvex CR Manifolds,* Comm. P. D. E. **18** (11) (1993), 1847–1868.

[DT4] M. DERRIDJ AND D.S. TARTAKOFF. *Local Analyticity in the $\overline{\partial}$-Neumann Problem for a Class of Totally Decoupled Weakly Pseudoconvex Domains,* Journal of Geometric Analysis **3** (2) (1993), 141–151.

[DT5] M. DERRIDJ AND D.S. TARTAKOFF. *Microlocal Analyticity for the Canonical Solution to $\overline{\partial}_b$ on Strictly Pseudoconvex CR Manifolds of Real Dimension Three,* Comm. P. D. E. **20** (9&10) (1994), 1647–1667.

[DT6] M. DERRIDJ AND D.S.TARTAKOFF. *Global Analytic Hypoellipticity for a Class of Quasilinear Sums of Squares of Vector Fields*, Contemporary Mathematics **368** (2005), 177–200.

[DT7] M. DERRIDJ AND D.S. TARTAKOFF. *On the Global Real Analyticity for the $\overline{\partial}$-Neumann Problem*, Journal Comm. in P.D.E. **5** (1976), 401–435.

[DT8] M. DERRIDJ AND D.S. TARTAKOFF. *Sur la Régularité Locale des Solutions du problème de Neumann pour $\overline{\partial}$*, Sem. Anal. P. Lelong (1977), in Lecture Notes in Mathematics (Springer-Verlag) v. 578, 207–216.

[DT9] M. DERRIDJ AND D.S. TARTAKOFF. *Analyticité Locale pour le problème de $\overline{\partial}$-Neumann en des Points de Faible Pseudoconvexité*, C. R. Acad. Sci. Paris Sr. I Math. **306** (10) (1988), 429–432.

[DT10] M. Derridj and D.S. Tartakoff. *Analyticity and Loss of Derivatives,* Annals of Mathematics **162** (2) (2005), 982–986.

[DZ1] M. Derridj and C. Zuily. *Sur la Régularité Gevrey des Opérateurs de Hörmander,* J. math. pures et appl. **52** (1973), 309–336.

[DZ2] M. Derridj and C. Zuily. *Régularité Analytique et Gevrey pour des Classes d'Opérateurs Élliptiques Paraboliques Dégénérées du Second Ordre,* Astérisque **2,3** (1973), 371–381.

[DF] K. Diederich and J.E. Fornaess. *Pseudoconvex Domains with Real-Analytic Boundaries*, Annals of Math, **107** (1978) 371–384.

[Eh] L. Ehrenpreis. *Solutions of Some Problems of Division IV,* Amer. J. Math. **82** (1960), 522–588.

[Fed] V. S. Fediĭ. *On a Criterion for Hypoellipticity,* Math. USSR Sb. **14** (1971), 14–45.

[Fel] W. Feller. *Introduction to Probability Theory and its Applications*, John Wiley & Sons, New York, 1957.

[FP] C. Fefferman and X. Phong. *Subelliptic Eigenvalue Problems*, Proceedings of Conference on Harmonic Analysis in Honor of Antoni Zygmund, 1981, 590–606.

[FK] G.B. Folland and J.J. Kohn. *The Neumann Problem for the Cauchy–Riemann Complex*, Annals of Math Studies # 75, Princeton 1972.

[Fr] K.O. Friedrichs. *On the Differentiability of the Solutions of Linear Elliptic Differential Equations*, Communications on Pure and Applied Mathematics **VI** (3) (1953) 299–326.

[GR] A.Grigis and L.P. Rothschild. *A Criterion for Analytic Hypoellipticity of a Class of Differential Operators with Polynomial Coefficients*, Ann. of Math., **118** (1983), 443–460.

[GS] A. Grigis and J. Sjöstrand. *Front d'Onde Analytique et Somme de Carrés de Champs de Vecteurs*, Duke Math. J., **52** (1985), 35–51.

[Gru] V.V. Grušin. *A Certain Class of Elliptic Pseudodifferential Operators That Are Degenerate on a Submanifold,* Mat. Sbornik **84** (13) (1971), 163–195.

[Han] N. Hanges. *Analytic Regularity for an Operator with Treves Curves*, J. Functional Analysis, **210** (2004), 295–320.

[HH] N. Hanges and A. Himonas. *Singular Solutions for Sums of Squares of Vector Fields,* Comm. in P.D.E. **16** (8,9) (1991), 1503–1511.

[Har] Ph. Hartman. *Ordinary Differential Equations*, second edition, Birkhäuser, Boston–Basel–Stuttgart, 1982.

[He] B. Helffer. *Sur l'Hypoellipticité des Opérateurs Pseudodifférentiels à Caractéristiques Multiples* (*Perte de $\frac{3}{2}$ Dérivées*), Bull. Soc. Math. France, Mémoire 51–52, 1977, 13–61.

[HM] B. Helffer and C. Mattera. *Analyticité et Itérés Réduits d'un Système de Champs de Vecteurs*, Commun. in P.D.E. **5 (10)** (1980), 1065–1072.

[Hi] F. Hirzebruch. *Topological Methods in Algebraic Geometry*, Springer-Verlag, New York, 1966.

[Hö1] L. Hörmander. *Hypoelliptic Second Order Differential Equations,* Acta Math. **119** (1967), 147–171.

[Hö2] L. Hörmander. *Uniqueness Theorems and Wave Front Sets for Solution of Linear Differential Equations with Analytic Coefficients,* Comm. Pure Appl. Math., **24** (1971), 671–704.

[Hö3] L. Hörmander. *A Class of Hypoelliptic Pseudo-differential Operators with Double Characteristics.* Math. Ann., **217** (1975), 165–188.

[Hö4] L. Hörmander. *The Analysis of Linear Partial Differential Operators I–IV.* Springer-Verlag, Berlin, 1983–1985.

[KW] K. Kajitan and S. Wakabayashi. *Hypoelliptic Operators in Gevrey Classes,* Recent Developments in Hyperbolic Equations, L. Cattabriga et al., eds, Pitman Research Notes in Math. **183** (1988), 115–134.

[K1] J.J. KOHN. *Harmonic Integrals on Strongly Pseudo-convex Manifolds, I & II*, Annals of Mathematics **78** (1963), 112–148 and **79** (1964), 450–472.

[K2] J.J. KOHN. *Boundary Behaviour of $\overline{\partial}$ on Weakly Pseudo-convex Manifolds of Dimension two*, J. of Diff. Geom. **6** (1972), 532–542.

[K3] J.J. KOHN. *Subellipticity of the $\overline{\partial}$-Neumann Problem on Pseudo-convex Domains: Sufficient Conditions*, Acta Math. **142** (1979), 79–122.

[K4] J.J. KOHN. *Hypoellipticity and Loss of Derivatives,* Annals of Mathematics, **162** (2) (2005), 943–986.

[Kom] G. KOMATSU. *Global Analytic Hypoellipticity of the $\overline{\partial}$-Neumann Problem*, Tohoku Math. J. **28** (1976), 145–156.

[KN] J.J. KOHN AND L. NIRENBERG. *Non-coercive Boundary Value Problems,* Comm. Pure Appl. Math. **18** (1967), 443–492.

[KS] S. KUSUOKA AND D. STROOK. *Applications of the Malliavin Calculus II,* J. Fac. Sci. Univ. Tokyo **32** (1985), 1–76.

[Kw1] K.H. KWON. *Concatenations Applied to Analytic Hypoellipticity of Operators with Double Characteristics*, Trans. Amer. Math. Soc., **283** (1984), 753–763.

[Mé1] G. MÉTIVIER. *Une Classe d'Opérateurs Non Hypoelliptiques Analytiques*, Indiana Math. J., **29** (1980), 823–860.

[Mé2] G. MÉTIVIER. *Analytic Hypoellipticity for Operators with Multiple Characteristics*, Comm. Partial Differential Equations, **6** (1981), 1–90.

[Mé3] G. MÉTIVIER. *Non Hypoellipticité Analytique pour $D_x^2 + (x^2 + y^2)D_y^2$.* Comptes Rendus Acad. Sci. Paris, **292** (1981), 401–404.

[Mé4] G. MÉTIVIER. *Non Hypoellipticité Analytique pour des Opérateurs à Caractéristiques Doubles.* Séminaire Goulaouic–Meyer–Schwartz, École Polytechnique, 1981–1982, Exposé XII, 1–12.

[Mo] Y. MORIMOTO. *Hypoellipticity for Infinitely Degenerate Elliptic Operators,* Osaka J. Math. **24** (1987), 13–35.

[Ok] T. OKAJI. *Analytic Hypoellipticity for Operators with Symplectic Characteristics,* J. Math. Kyoto Univ. **25** (3) (1985), 489–514.

[O1] O. OLEINIK. *On the Analyticity of Solutions to Partial Differential Equations and Systems,* Soc. Math. de France, Astérisque **2-3** (1973), 272–285.

[OR] O. OLEINIK AND R. RADKEVICH. *Conditions for the Analyticity of all Solutions of a Second Order Linear Equation* (Russian) Uspehi Mat. Nauk **177** (3) (1974), 221–222.

[PP] C. PARENTI AND A. PARMEGGIANI. *On the Hypoellipticity with a Big Loss of Derivatives,* Kyushu J. Math. **59** (2005), 155–230.

[RS] L.P. ROTHSCHILD, E.M. STEIN. *Hypoelliptic Differential Operators and Nilpotent Groups,* Acta Mathematica, **137** (1976), 247–320.

[S1] T. SAITO. *Global Analytic Hypoellipticity for the $\overline{\partial}$ Problem for Some Weakly Pseudo-convex Domains in $\mathbb{C}^n$*, Math. Rep. Toyama Univ. **6** (1983), 105–120.

[Si] Y.-T. SIU. *Extension of Twisted Pluricanonical Sections with Plurisubharmonic Weight and Invariance of Semipositively Twisted Plurigenera for Manifolds not Necessarily of General Type,* in *Complex Geometry: Collection of Papers Dedicated to Professor Hans Grauert*, pp. 223–277, Springer-Verlag, New York, 2002.

[Sj1] J. SJÖSTRAND. *Parametrices for Pseudodifferential Operators with Multiple Characteristics*, Ark. för Mat., **12** (1974), 85–130.

[Sj2] J. SJÖSTRAND. *Singularités Analytiques Microlocales*, Astérisque, **95** (1982).

[Sj3] J. SJÖSTRAND. *Analytic Wavefront Set and Operators with Multiple Characteristics*, Hokkaido Math. J., **12** (1983), 392–433.

[ST1] N.S. STANTON AND D.S. TARTAKOFF. *The Real Analytic and Gevrey Regularity of the Heat Kernel for $\Box_b$*, in Pseudodifferential Operators and Applications, Proc. Sympos. Pure Math., 43 (1984), Amer. Math. Soc., Providence, RI, 1985, 247–259.

[ST2] N.S. STANTON AND D.S. TARTAKOFF. *The Heat Equation for the $\overline{\partial}_b$-Laplacian*, Comm. Partial Differential Equations **9** (7) (1984), 597–686.

[St] E.M. STEIN. *An Example on the Heisenberg Group Related to the Lewy Operator,* Invent. Math. **69** (2) (1982), 209–216.

[Tan] N. TANAKA. *On Generalized Graded Lie Algebras and Geometric Structures I*, J. Math. Soc. Japan, **19** (1967), 215–254.

[T1] D.S. TARTAKOFF. *Gevrey Hypoellipticity for Subelliptic Boundary Value Problems.* Communications on Pure and Applied Math. **26** (1973), 251–312.

[T2] D.S. TARTAKOFF. *On the Global Real Analyticity of Solutions to $\Box_b$ on Compact Manifolds,* Comm. in P.D.E. **1** (1976), 283–311.

[T3] D.S. TARTAKOFF. *On the Local Gevrey and Quasianalytic Hypoellipticity for $\Box_b$,* Comm. in P.D.E. **2** (1977), 699–712.

[T4] D.S. TARTAKOFF. *Local Analytic Hypoellipticity for $\Box_b$ on Non-Degenerate Cauchy Riemann Manifolds*, Proc. Nat. Acad. Sci. U.S.A., **75** (1978), 3027–3028.

[T5] D.S. TARTAKOFF. *On the Local Real Analyticity of Solutions to Box-b and the $\bar\partial$-Neumann Problem*, Acta Mathematica, **145** (1980), 117–204.

[T6] D.S. TARTAKOFF. *Operators with Multiple Characteristics—an L^2 Proof of Analytic Hypoellipticity,* Conference on linear partial and pseudodifferential operators (Torino, 1982). Rend. Sem. Mat. Univ. Politec. Torino 1983, Special Issue, 251–282.

[T7] D.S. TARTAKOFF. *Global (and Local) Analyticity for Second Order Operators Constructed from Rigid Vector Fields on Products of Tori,* Transactions of the A.M.S. **348** (7) (1996), 2577–2583.

[T8] D.S. TARTAKOFF. *Analytic Hypoellipticity for $\Box_b + c$ on the Heisenberg Group: An L^2 Approach*, Far East Journal of Applied Mathematics, **15** (3) (2004), 353–363.

[T9] D.S. TARTAKOFF. *Analyticity for Singular Sums of Squares of Degenerate Vector Fields,* Proc. Amer. Math. Soc. **134** (2006), 3343–3352.

[T10] D.S. TARTAKOFF. *Local Gevrey and Quasi-analytic Hypoellipticity for $\Box_b$,* Bull. Amer. Math. Soc. **82** (5) (1976), 740–742.

[TZ] D.S. TARTAKOFF AND L. ZANGHIRATI. *Local Real Analyticity of Solutions for Sums of Squares of Non-linear Vector Fields*, J. Differential Equations, **213** (2) (2005), 341–351.

[Tr1] F. TREVES. *Symplectic Geometry and Analytic Hypo-ellipticity, in Differential Equations: La Pietra 1996 (Florence)*, Proc. Sympos. Pure Math., **65**, Amer. Math. Soc., Providence, RI, 1999, 201–219.

[Tr2] F. TREVES. *On the Analyticity of Solutions of Sums of Squares of Vector Fields*, to appear in the Proceedings of the Pienza meeting on "Phase space analysis and PDEs."

[Tr3] F. TREVES. *Hypo-analytic Structures, Local Theory,* Princeton University Press, Princeton, 1992.

[Tr4] F. TREVES. *Analytic Hypo-ellipticity of a Class of Pseudo-differential Operators with Double Characteristics and Application to the d-bar Neumann Problem*, Commun. in P.D.E., **3** (6-7), (1978), 475–642.

[Xu1] CHAO-QIANG XU. *Hypoellipticité pour les Équations aux Dérivées Partielles Non Linéaires Associées à un Système de Champs de Vecteurs*, C. R., Acad. Sc. Paris Série I, **300** (8) (1985), 235–237.

[Xu2] CHAO-QIANG XU. *Régularité des Solutions pour les Équations aux Dérivées Partielles Quasi Linéaires Non Elliptiques du Second Ordre*, C. R., Acad. Sc. Paris Série I, **300** (9) (1985), 7267–7270.