

A CLASS OF SUMS OF SQUARES WITH A GIVEN POISSON-TREVES STRATIFICATION

ANTONIO BOVE AND DAVID TARTAKOFF

ABSTRACT. We study a class of sum of squares exhibiting the same Poisson-Treves stratification as the Oleinik-Radkevič operator. We find three types of operators having distinct microlocal structures. For one of these we prove a Gevrey hypoellipticity theorem analogous to our recent result for the corresponding Oleinik-Radkevič operator.

1. INTRODUCTION

The problem of analytic hypoellipticity for second order operators which are sums of squares of vector fields with analytic coefficients has been widely studied and has been around since the paper of L. Hörmander [19] on C^∞ hypoellipticity for this type of operator. In particular D. S. Tartakoff [23], [24], for second order, and F. Treves [26] for general order, gave general analytic hypoellipticity theorems for the case when the characteristic manifold is symplectic and the operator degenerates on it to an exact order. In the case of non exact and higher order degeneracy O. Oleinik [21] and O. Oleinik and R. Radkevič [22] (see also Christ [8]) showed that in general one cannot have analytic hypoellipticity (see also the paper [3] by the present authors for a precise and optimal partial regularity result in the case of the operator studied by Oleinik and Radkevič), but only certain degrees of Gevrey hypoellipticity.

Into this scenario there appeared in 1999 the well known paper by F. Treves [28] introducing the notion of Poisson stratification for a set of vector fields satisfying Hörmander's Lie algebra condition and having analytic coefficients. Basically, crudely simplifying Treves' setting, the conjecture states that an operator which is a sum of squares of vector fields is analytic hypoelliptic if and only if every layer in its Poisson

Date: February 10, 2003.

stratification is symplectic. To our knowledge the conjecture has been neither proved nor disproved up to now.

In this paper, the first of a series, we study an operator which is the sum of the squares of three vector fields with analytic coefficients in three variables. For such an operator we assume that its Poisson-Treves stratification is given in such a way that its Hörmander numbers are the same as those of the Oleinik-Radkevič model operator. Here by Hörmander numbers we mean both the number and the relative codimensions of the stratification's layers. Our purpose is to classify such kind of operators and obtain their Gevrey (or possibly analytic) hypoellipticity threshold.

In the present paper we classify the operators having the required properties and, for one of the classes, we obtain the same Gevrey hypoellipticity threshold as that of the Oleinik-Radkevič model operator. We are unable to deduce these (optimal) results for every class of operators sharing the Poisson-Treves stratification with the Oleinik-Radkevič model, but we shall come back to this subject in a forthcoming paper [5].

Before stating our assumptions precisely, we want to make a couple of remarks.

1. Our vector fields are linearly independent outside of the characteristic manifold. This essentially implies that the characteristic manifold is cylindrical with respect to a two dimensional subspace of the fibers of the cotangent bundle, or in other words, it is the zero set of one covariable and one function of the variables in the base. This restriction eliminates cases where two of the vector fields can become colinear outside of the characteristic set. On the other hand, many results are known for the case of the sums of two squares.

2. The Oleinik-Radkevič model operator has a codimension 2 symplectic characteristic manifold. In three dimensions one might consider also cases where the characteristic manifold is symplectic and of dimension 4 or has symplectic layers of codimension 2 and symplectic layers of codimension 4. Even for the sums of two squares, though, this situation faces difficulties of the same kind as those appearing in Christ's example [9].

In the first part of the paper we deduce some standard forms (cf. Theorem 6.3.1) below) that can be useful in proving *a priori* estimates. Then we proceed to prove a Gevrey hypoellipticity threshold for one of these standard forms (cf. Theorem 9.1)

Essentially the operators verifying our assumptions fall into three classes, depending on how the vector fields vanish on the characteristic set. For the first case, called Case 1, we make a finer analysis of the extent to which the vector fields under consideration are linearly independent outside of the characteristic manifold. This is accomplished by looking at each of the “characteristic” vector fields and computing it on the null bicharacteristic curve of the only non characteristic vector field.

This operation does not affect the covariables (i.e. affects only the coefficients of the base), since the null bicharacteristic curve is a curve in the base variables. Then one focusses on the zero set of the resulting vector fields. Computing the symbol of one of the vector fields on the zero set of the other allows us to define a sort of degeneracy rate which turns out to be useful in the *a priori* estimates. The last section of the present paper is concerned with the case when the above mentioned degeneracy rate is zero. Then we obtain the same (optimal) result as for the Oleinik-Radkevič model.

If the degeneracy rate is larger than zero, the estimates are deduced in a very different way and yield a different result. This is the subject of a forthcoming paper.

The second and third classes (called Case 2a and Case 2b below) will also be studied in a subsequent paper.

2. ASSUMPTIONS

We now specify the assumptions. Our operators have the general form

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

where $x \in \mathbb{R}^3$ and $D_j = \frac{1}{i} \frac{\partial}{\partial x_j}$, $j = 1, 2, 3$. Here X_j denotes a vector field with real analytic coefficients defined in a neighborhood of the origin in the x variable.

The following assumptions try to mimic the fact that P has the same Poisson-Treves stratification as the operator $D_1^2 + x_1^{2(p-1)} D_2^2 + x_1^{2(q-1)} D_3^2$, where p and q are integers and $1 \leq p \leq q$.

- (A1) The operator P satisfies the Hörmander Lie algebra condition and hence is C^∞ hypoelliptic. As a consequence not all the vector fields are characteristic (i.e. have vanishing coefficients) on the characteristic manifold. Hence we may suppose without loss of generality that

$$X_1(x, D) = D_1.$$

- (A2) We may always assume that the point $(0; e_3)$ is a characteristic point for P (using a translation and a rotation if necessary). We assume then that near $(0; e_3)$, the characteristic set of P is an analytic symplectic submanifold of $T^*\mathbb{R}^3 \setminus 0$ of codimension two which we denote by Σ_1 . We explicitly note that this is a microlocal assumption.

- (A3) Let $\Omega = U \times \Gamma$ a conic neighborhood of the point $(0, e_3)$. And let $\pi_1: U \times \Gamma \rightarrow U$ be the projection onto the space variables. We assume that the vector fields

$$X_j|_{U \setminus \pi_1 \Sigma_1}$$

are linearly independent; the above notation means that restricting the coefficients of the fields X_j to the space projection of Σ_1 yields linearly independent vectors in \mathbb{R}^3 .

Note that, because of (A2), the coefficients of the vector fields depend non trivially on the x variable. This assumption has strong implications on the structure of Σ_1 and, to avoid technical details at this point, we refer to Section 3.

- (A4) Denote by

$$\Sigma_2 = \{(x, \xi) \in T^*\mathbb{R}^3 \setminus 0 \mid (x, \xi) \in \Sigma_1, \{X_i, X_j\}(x, \xi) = 0, \\ i, j \in \{1, 2, 3\}\},$$

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

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

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

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

Let $p \leq q$ be two positive integers. Then we make the following assumptions:

- (i) $\Sigma_1 \cap \Omega = \dots = \Sigma_{p-1} \cap \Omega$.
- (ii) $\Sigma_p \cap \Omega$ is a non-empty analytic submanifold of $\Sigma_1 \cap \Omega$ of codimension one.
- (iii) $\Sigma_p \cap \Omega = \Sigma_{p+1} \cap \Omega = \dots = \Sigma_{q-1} \cap \Omega$.
- (iv) $\Sigma_q \cap \Omega$ is empty in $T^*\mathbb{R}^3 \setminus 0$ (i.e. $\Sigma_q \cap \Omega$ is contained in the zero section of the cotangent bundle over Ω .)

3. STANDARD FORMS: THE EQUATIONS OF Σ_1

Due to the above assumptions we may suppose that the vector fields have the following form:

$$\begin{aligned} X_1(x, \xi) &= \xi_1 \\ X_2(x, \xi) &= a_{21}(x)\xi_1 + a_{22}(x)\xi_2 + a_{23}(x)\xi_3 \\ X_3(x, \xi) &= a_{31}(x)\xi_1 + a_{32}(x)\xi_2 + a_{33}(x)\xi_3. \end{aligned} \tag{3.1}$$

Hence $\xi_1 = 0$ is one of the two equations defining Σ_1 ; letting

$$A(x) = \begin{bmatrix} a_{22}(x) & a_{23}(x) \\ a_{32}(x) & a_{33}(x) \end{bmatrix},$$

(the a_{jk} being analytic), the other equation is given by

$$A(x)\xi' = 0, \tag{3.2}$$

where $\xi' = (\xi_2, \xi_3)$.

We claim that this can only be the second defining condition of Σ_1 if $A(x) \equiv 0$ on Σ_1 (locally). For suppose (x_0, ξ'_0) , $\xi'_0 \neq 0$, is such that

$$A(x_0)\xi'_0 = 0 \tag{3.3}$$

and assume that to the contrary, for x near x_0 on Σ_1 ,

$$A(x) \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (\text{which we will write as } A(x) \neq 0).$$

This implies that at x_0 the rank of A is equal to 1 since $0 \neq \xi'_0 \in \ker A(x_0)$. It follows that, in a conic neighborhood of $(x_0; 0, \xi'_0)$, the characteristic manifold Σ_1 is defined by

$$\Sigma_1 = \{(x, \xi) \mid \xi_1 = 0, \det A(x) = 0, \xi' \in \ker A(x)\},$$

because we may always assume that $\text{rank } A(x) \geq 1$ near x_0 .

Since $A(x_0) \neq 0$, the latter two equations in the definition of Σ_1 are certainly independent (the second of them has non-zero ξ' -gradient, while the first of them must have a non-zero x -gradient). As a consequence one of them must be identically satisfied in order to accomplish the codimension 2 condition. Since $\text{rank } A(x_0) = 1$, the condition $\xi' \in \ker A(x)$ cannot be identically satisfied. Hence the only possibility is that

$$\det A(x) \equiv 0$$

in a full neighborhood of x_0 . However this fact would imply that there exist points (x, ξ) , $\xi_1 = 0$, $(x, \xi) \notin \Sigma_1$, such that the vector fields X_1, X_2, X_3 are not linearly independent.

Consequently the only possible case left is that A is the zero matrix:

$$(3.4) \quad A(x) = 0,$$

if $(x, \xi) \in \Sigma_1$. This means that

$$(3.5) \quad \Sigma_1 = \{(x, \xi) \mid \xi_1 = 0, A(x) = 0\}.$$

Hence the matrix condition $A(x) = 0$ must be (locally) equivalent to $\varphi(x) = 0$, where φ is a real analytic scalar function and such that $d_x \varphi(x) \neq 0$.

By Assumption (A2), $\{\xi_1, \varphi(x)\} \neq 0$ at Σ_1 . Hence by the implicit function theorem the equation $\varphi(x) = 0$ is equivalent to the equation

$$(3.6) \quad x_1 - g(x') = 0,$$

where g is a suitable real analytic function, $x' = (x_2, x_3)$, and g is defined locally. We conclude then that

$$(3.7) \quad \Sigma_1 = \{(x, \xi) \mid \xi_1 = 0, x_1 - g(x') = 0\},$$

and that

$$(3.8) \quad A(x) = (x_1 - g(x'))\tilde{A}(x),$$

for a suitable 2×2 matrix \tilde{A} with real analytic entries \tilde{a}_{ij} , $i, j \in \{2, 3\}$.

Next we perform a change of variables (and hence a canonical transformation) which is linear in ξ , so that vector fields are mapped to vector fields in the new coordinates, allowing us to make the function g identically zero.

Define:

$$(3.9) \quad \begin{aligned} y_1 &= x_1 + g(x') & \eta_1 &= \xi_1 \\ y_2 &= x_2 & \eta_2 &= \xi_2 - \xi_1 \frac{\partial g}{\partial x_2} \\ y_3 &= x_3 & \eta_3 &= \xi_3 - \xi_1 \frac{\partial g}{\partial x_3}. \end{aligned}$$

The three vector fields become:

$$\begin{aligned} X_1(y, \eta) &= \eta_1 \\ X_2(y, \eta) &= \left(a_{21}(y_1 - g(y'), y') + y_1 \tilde{a}_{22}(y_1 - g(y'), y') \frac{\partial g(y')}{\partial y_2} \right. \\ &\quad \left. + y_1 \tilde{a}_{23}(y_1 - g(y'), y') \frac{\partial g(y')}{\partial y_3} \right) \eta_1 \\ &\quad + y_1 [\tilde{a}_{22}(y_1 - g(y'), y') \eta_2 + \tilde{a}_{23}(y_1 - g(y'), y') \eta_3] \\ X_3(y, \eta) &= \left(a_{31}(y_1 - g(y'), y') + y_1 \tilde{a}_{32}(y_1 - g(y'), y') \frac{\partial g(y')}{\partial y_2} \right. \\ &\quad \left. + y_1 \tilde{a}_{33}(y_1 - g(y'), y') \frac{\partial g(y')}{\partial y_3} \right) \eta_1 \\ &\quad + y_1 [\tilde{a}_{32}(y_1 - g(y'), y') \eta_2 + \tilde{a}_{33}(y_1 - g(y'), y') \eta_3]. \end{aligned}$$

The above fields can be rewritten, with obvious notation, in the following way:

$$(3.10) \quad \begin{aligned} X_1(x, \xi) &= \xi_1 \\ X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1 [a_{22}(x)\xi_2 + a_{23}(x)\xi_3] \\ X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1 [a_{32}(x)\xi_2 + a_{33}(x)\xi_3] \end{aligned}$$

with suitable real analytic functions a_{ij} defined in a neighborhood of the origin.

4. STANDARD FORMS: THE EQUATIONS OF $\Sigma_2, \dots, \Sigma_{p-1}$

Let us now turn to Assumption (A4) concerning Σ_2 . We have

$$\{X_1(x, \xi), X_j(x, \xi)\} = \frac{\partial}{\partial x_1} X_j(x, \xi),$$

for $j = 2, 3$ and the latter quantity is equal to

$$\frac{\partial a_{j1}(x)}{\partial x_1} \xi_1 + [a_{j2}(x)\xi_2 + a_{j3}(x)\xi_3] + O(|x_1|),$$

for $j = 2, 3$, and

$$\begin{aligned} & \{X_2, X_3\}(x, \xi) \\ &= \{a_{21}\xi_1 + x_1 [a_{22}(x)\xi_2 + a_{23}(x)\xi_3], a_{31}\xi_1 + x_1 [a_{22}(x)\xi_2 + a_{23}(x)\xi_3]\}, \end{aligned}$$

which gives

$$(4.1) \quad \{X_2, X_3\}(x, \xi) = a_{21}(x)\{X_1, X_2\}(x, \xi) - a_{31}(x)\{X_1, X_3\}(x, \xi) + O(|x_1| + |\xi_1|),$$

where $O(|x_1| + |\xi_1|)$ stands for a vector field with principal symbol vanishing on Σ_1 . Hence we obtain $\Sigma_2 = \Sigma_1 \cap \{(x, \xi) \mid \{X_1, X_j\} = 0\}$, $j = 2, 3$.

Let us again denote by $A(x)$ the 2×2 matrix

$$A(x) = \begin{bmatrix} a_{22}(x) & a_{23}(x) \\ a_{32}(x) & a_{33}(x) \end{bmatrix};$$

then Assumption (A4) means that

$$(4.2) \quad A(x)\xi' = 0$$

if and only if $x_1 = 0$. This implies that

$$(4.3) \quad A(x) = x_1 \tilde{A}(x)$$

for a suitable 2×2 matrix \tilde{A} with analytic entries.

Iterating the above argument we can conclude that the vector fields can be written in the form

$$(4.4) \quad \begin{aligned} X_1(x, \xi) &= \xi_1 \\ X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [a_{22}(x)\xi_2 + a_{23}(x)\xi_3] \\ X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [a_{32}(x)\xi_2 + a_{33}(x)\xi_3]. \end{aligned}$$

We summarize what has been proved up to this point in the

Proposition 4.1. *Suppose that (A1)–(A3) and (A4)(i) hold. Then the vector fields X_1, X_2, X_3 can be written, in a suitable system of local coordinates, in the form (4.4).*

5. THE EQUATION DEFINING Σ_p WITH RESPECT TO Σ_1

Let us denote by $\varphi(x', \xi')$ a real analytic function defined on a (conic) neighborhood of $(0, e_3)$ in Σ_1 and such that $d_{(x', \xi')} \varphi(0, e_3) \neq 0$ and the equation $\varphi(x', \xi') = 0$ is equivalent to $A(0, x')\xi' = 0$.

We have either

$$(5.1) \quad \frac{\partial \varphi}{\partial \xi'}(0, e_3) \neq 0 \quad (\text{Case I})$$

or

$$(5.2) \quad \frac{\partial \varphi}{\partial x'}(0, e_3) \neq 0 \quad (\text{Case II}).$$

5.1. Case I. For the case of non-zero ξ gradient, assume that it is the ξ_2 derivative of φ that is non-zero at $(0, e_3)$ (we will see below that the case of a non-zero ξ_3 derivative cannot occur). Then we may write

$$(5.3) \quad \varphi(x', \xi') = (\xi_2 - \chi(x', \xi_3))e(x', \xi'),$$

where e and χ are analytic and $e(0, e_3) \neq 0$ and thus

$$\varphi(x', \xi') = 0 \iff \xi_2 - \chi(x', \xi_3) = 0 \iff A(x')\xi' = 0.$$

We claim that $\chi(x', \xi')$ has the simpler form $\tilde{\chi}(x')\xi_3$, and to see this let t denote a non-zero real number; if $A(x')\xi' = 0$ then obviously $A(x')t\xi' = 0$. Thus $(x', \xi') \in \Sigma_2 \implies (x', t\xi') \in \Sigma_2$, so that $t\xi_2 - \chi(x', t\xi_3) = 0$. Since $\xi_2 = \chi(x', \xi_3)$, we have $\chi(x', t\xi_3) = t\chi(x', \xi_3)$ for every non-zero real number t . But now $\xi_3 \neq 0$ in a conic neighborhood of $(0, e_3)$, so that $\chi(x', \xi_3) = \xi_3\chi(x', 1) = \xi_3\tilde{\chi}(x')$, for a suitable analytic function $\tilde{\chi}$ of the space variable only, and so finally we obtain

$$(5.4) \quad A(x')\xi' = 0 \iff \xi_2 - \chi(x')\xi_3 = 0,$$

where we have written χ again for the function $\tilde{\chi}$.

The above formula has been derived in the case that it is the ξ_2 derivative of φ that is non-zero at $(0, e_3)$. Now suppose that the ξ_3 derivative of φ is non-zero at $(0, e_3)$ instead. Then arguing as above we find that the equation $\varphi(x', \xi') = 0$ is equivalent to $\xi_3 - \chi(x', \xi_2) = 0$. As before, let t be a non-zero real number; now, since if (x', ξ')

belongs to Σ_p then also $(x', t\xi')$ belongs to Σ_p , keeping in mind that, by assumption, the point $(0, e_3)$ belongs to Σ_p , we find that $t = \chi(0, 0)$ for any $t \in \mathbb{R} \setminus 0$, which is absurd.

We thus have proved that if (5.1) is true then the equation defining Σ_p relatively to Σ_1 is given by (5.4).

5.2. Case II. We now turn to the case where $\varphi'_\xi(0, e_3) = 0$ but

$$(5.5) \quad \frac{\partial \varphi}{\partial x'}(0, e_3) \neq 0 \quad (\text{Case II})$$

and we assume here that

$$(5.6) \quad \frac{\partial \varphi}{\partial x_2}(0, e_3) \neq 0 \quad (\text{Case II}_{x_2}).$$

The case $\varphi_{x_3}(0, e_3) \neq 0$ (Case II $_{x_3}$) has some obvious but non-trivial differences that we shall stress later.

Arguing along the same lines as above we obtain that there is a function $\chi(x_3, \xi')$ such that the equation $\varphi(x', \xi') = 0$ is equivalent to $x_2 - \chi(x_3, \xi') = 0$. Here χ is analytic and defined on a conic neighborhood of $(0, e_3)$ in $\mathbb{R}_{x_3} \times (\mathbb{R}_{\xi'}^2 \setminus 0)$. Again we may assume that on that neighborhood ξ_3 is not zero. Moreover if t denotes a non-zero real number we obtain that $\chi(x_3, t\xi') = \chi(x_3, \xi')$, so that

$$(5.7) \quad A(x')\xi' = 0 \iff x_2 - \chi(x_3, \frac{\xi_2}{\xi_3}) = 0,$$

where we have denoted by $\chi(x_3, \sigma)$ the function $\chi(x_3, \sigma, 1)$.

We point out that the function χ in (5.7) is an analytic function defined in a neighborhood of the origin in $\mathbb{R}_{x_3} \times \mathbb{R}_\sigma$.

From (5.7) we obtain that there is a positive integer k such that

$$(5.8) \quad A(x')\xi' = \left(x_2 - \chi(x_3, \frac{\xi_2}{\xi_3}) \right)^k B(x', \xi'),$$

where $B(x', \xi')$ denotes an analytic 2-vector defined and non-zero in a conic neighborhood of $(0, e_3)$. The existence of such an integer k is a consequence of our analyticity assumption.

Our aim is to draw some consequences from the linearity of the left hand side of Equation (5.8) with respect to ξ' .

Assume first that $k > 1$ in (5.8). Then taking the ξ -gradient, we get

$$A(x') = O\left(\left(x_2 - \chi\left(x_3, \frac{\xi_2}{\xi_3}\right)\right)^{k-1}\right),$$

which implies that $\chi(x_3, \xi_2/\xi_3)$ actually depends only on x_3 . Thus $B(x', \xi')$ is linear with respect to ξ' , so that we obtain

$$(5.9) \quad A(x')\xi' = (x_2 - \chi(x_3))^k \tilde{A}(x')\xi',$$

where $\tilde{A}(x')$ denotes another 2×2 matrix with real analytic entries.

Let us now assume that $k = 1$. Equation (5.8) becomes

$$(5.10) \quad A(x')\xi' = \varphi(x', \xi')B(x', \xi'),$$

where B is a vector-valued symbol of order 0. Recall that we are assuming that

$$\varphi(0, e_3) = 0, \quad \frac{\partial \varphi}{\partial \xi'}(0, e_3) = 0, \quad \text{and} \quad \frac{\partial \varphi}{\partial x_2}(0, e_3) \neq 0.$$

Since the vanishing in (5.10) is of the first order, we have that $B(0, e_3) \neq 0$; in particular we may assume that

$$B(0, e_3) = (b_2(0, e_3), b_3(0, e_3))$$

and

$$(5.11) \quad b_3(0, e_3) \neq 0.$$

This is no restriction since we can always interchange the second and the third vector fields. Taking the ξ' -gradient of (5.10) and computing everything at $(0, e_3)$, we easily see that $A(0) = 0$. Hence

$$(5.12) \quad A(x') = x_2 A^{(2)}(x') + x_3 A^{(3)}(x'),$$

where the $A^{(j)}$ are real analytic 2×2 matrices, $j = 2, 3$. From this equation we obtain

$$\frac{\partial}{\partial x_2} A(x')\xi' = \frac{\partial \varphi}{\partial x_2} B(x', \xi') + \varphi(x', \xi') \frac{\partial B}{\partial x_2}(x', \xi'),$$

which, when computed at $(0, e_3)$, yields

$$\frac{\partial A(0)}{\partial x_2} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\partial \varphi}{\partial x_2}(0, e_3) B(0, e_3).$$

Let us now consider the second component of the above equation: we have, from (5.11),

$$\begin{aligned} \frac{\partial}{\partial x_2} \left[\left(x_2 a_{32}^{(2)}(x') + x_3 a_{32}^{(3)}(x') \right) \xi_2 \right. \\ \left. + \left(x_2 a_{33}^{(2)}(x') + x_3 a_{33}^{(3)}(x') \right) \xi_3 \right] \Big|_{x'=0, \xi'=e_3} \neq 0, \end{aligned}$$

from which we deduce that

$$(5.13) \quad a_{33}^{(2)}(0) \neq 0.$$

The second line of the equation $A(x')\xi' = 0$ then reads:

$$(5.14) \quad \left(x_2 a_{32}^{(2)}(x') + x_3 a_{32}^{(3)}(x') \right) \xi_2 + \left(x_2 a_{33}^{(2)}(x') + x_3 a_{33}^{(3)}(x') \right) \xi_3 = 0.$$

Because of (5.13) this is the equation of an analytic submanifold of codimension one containing the point $(0, e_3)$ and, since $b_3(0, e_3) \neq 0$, (5.10) implies that (5.14) is equivalent to $\varphi(x', \xi') = 0$.

Thus we are allowed to change notation and write

$$(5.15) \quad \varphi(x', \xi') = \left(x_2 a_{32}^{(2)}(x') + x_3 a_{32}^{(3)}(x') \right) \xi_2 \\ + \left(x_2 a_{33}^{(2)}(x') + x_3 a_{33}^{(3)}(x') \right) \xi_3,$$

where

$$(5.16) \quad a_{33}^{(2)}(0) \neq 0,$$

(recall that we are in Case II _{x_2} where $\varphi_{x_2} \neq 0$ at $(0, e_3)$).

The following Lemma will help distinguish between two very different types of families of vector fields. They are both in Case II and will be denoted Case IIa and Case IIb, which of course will be further subscripted according to whether $\varphi_{x_2} \neq 0$ or $\varphi_{x_3} \neq 0$.

Lemma 5.1. *Let λ and μ be real analytic functions defined in a neighborhood of the origin and consider the vector field*

$$Y = \lambda(x') \frac{\partial}{\partial x_2} + \mu(x') \frac{\partial}{\partial x_3}.$$

Assume that the symbol of Y , $\lambda(x')\xi_2 + \mu(x')\xi_3$, vanishes where φ vanishes, φ being defined in Equation (5.15). Then two cases may occur:

- (a) *The set $\varphi^{-1}(0)$ is cylindrical in the ξ' -fibers. Then $\varphi(x', \xi') = 0$ if and only if $g(x') = 0$, for a suitable analytic function g defined in a neighborhood of the origin and having a non-zero x' -gradient. In this case*

$$Y(x', \xi') = g(x')\tilde{Y}(x', \xi'),$$

for a suitable vector field \tilde{Y} .

- (b) *The set $\varphi^{-1}(0)$ is not cylindrical in the ξ' -fibers. Then there exists an analytic function of x' , $h(x')$, defined near 0, such that*

$$Y(x', \xi') = h(x')\varphi(x', \xi').$$

Proof. Let us write the function φ in (5.15) as

$$\varphi(x', \xi') = \alpha(x')\xi_2 + \beta(x')\xi_3,$$

where, by (5.16), $\partial\beta(0)/\partial x_2 \neq 0$. The vanishing of the symbol of Y where φ vanishes can be expressed by the following equation:

$$\lambda(x')\xi_2 + \mu(x')\xi_3 = a(x', \xi')(\alpha(x')\xi_2 + \beta(x')\xi_3),$$

where a is a suitable analytic symbol of order 0—actually homogenous of degree zero—defined near the point $(0, e_3)$.

Dividing by ξ_3 , which is non-zero near e_3 , and writing $\sigma = \xi_2/\xi_3$, we have

$$a(x', \sigma, 1) = \frac{\lambda(x')\sigma + \mu(x')}{\alpha(x')\sigma + \beta(x')}, \quad |\sigma| \leq C,$$

for a suitable positive constant C . Thus

$$a(x', 0, 1) = \frac{\mu(x')}{\beta(x')},$$

which is also analytic with respect to the variable x' near the origin. Since $\beta(0) = 0$ and $\partial\beta(0)/\partial x_2 \neq 0$, we have that $\beta^{-1}(0)$ is a regular analytic curve in \mathbb{R}^2 near the origin. Hence there exists an analytic function $\gamma(x')$ defined near 0 such that

$$\mu(x') = \gamma(x')\beta(x'),$$

and thus

$$a(x', \sigma, 1) = \frac{\lambda(x')\sigma + \gamma(x')\beta(x')}{\alpha(x')\sigma + \beta(x')}.$$

Now we have

$$\frac{\partial}{\partial \sigma} \frac{\lambda \sigma + \gamma \beta}{\alpha \sigma + \beta} = \frac{\beta(\lambda - \alpha \gamma)}{(\alpha \sigma + \beta)^2}$$

$$\left(\frac{\partial}{\partial \sigma} \right)^{h+1} \frac{\lambda \sigma + \gamma \beta}{\alpha \sigma + \beta} = (-1)^h (h+1)! \beta (\lambda - \alpha \gamma) \frac{\alpha^h}{(\alpha \sigma + \beta)^{h+2}}.$$

Setting $\sigma = 0$ in the first line and taking into account the analyticity of the left hand side, we obtain that

$$\lambda(x') = \alpha(x')\gamma(x') + \delta(x')\beta(x'),$$

for a suitable analytic function δ defined near the origin. On the other hand, for $\sigma = 0$, the second line gives

$$\left(\frac{\partial}{\partial \sigma} \right)^{h+1} a(x', \sigma, 1) \Big|_{\sigma=0} = (-1)^h (h+1)! \delta(x') \left(\frac{\alpha(x')}{\beta(x')} \right)^h.$$

Now two cases may occur:

- i) Assume that β is a factor of α , i.e. that $\alpha(x') = \eta(x')\beta(x')^k$, for a suitable positive integer k and a suitable analytic function η . In this case $\varphi(x', \xi') = (\eta\beta^{k-1}\xi_2 + \xi_3)\beta(x')$, with $\beta(0) = 0$, $\partial\beta(0)/\partial x_2 \neq 0$ and $(\eta\beta^{k-1}\xi_2 + \xi_3)|_{x'=0, \xi'=e_3} \neq 0$. We conclude that $\varphi^{-1}(0) = \beta^{-1}(0)$, or that $\varphi^{-1}(0)$ is the zero set of a function of x' only. Moreover in this case we have that $\lambda = (\eta\beta^{k-1} + \delta)\beta$, so that

$$\lambda(x')\xi_2 + \mu(x')\xi_3 = \beta(x') [(\eta(x')\beta(x')^{k-1} + \delta(x'))\xi_2 + \gamma(x')\xi_3],$$

which is the conclusion in part (a) of the statement of the Lemma.

- ii) The function β is not a factor of α , i.e. the quotient α/β is not analytic near 0. Then necessarily we must have that $\delta = 0$ if $\beta = 0$. But then it is easy to see that there exists a positive integer h such that δ/β^h is not analytic near the origin, unless $\delta \equiv 0$ in a neighborhood of the origin. Thus

$$\lambda(x') = \alpha(x')\gamma(x'),$$

at least in a possibly smaller neighborhood of the origin. The above equation implies that

$$\lambda(x')\sigma + \mu(x') = \gamma(x') (\alpha(x')\sigma + \beta(x')),$$

which is the desired conclusion for part (b) of the Lemma.

□

Summing up we can state the following

Proposition 5.1. *Assume that the quantity $A(0, x')\xi'$ vanishes exactly on an analytic submanifold Σ_p of codimension one inside Σ_1 . Let us denote by $\varphi(x', \xi') = 0$ a (microlocal) equation of Σ_p near the point $(0, e_3) \in \Sigma_p$. Then the following cases may occur:*

I) *If*

$$\frac{\partial \varphi}{\partial \xi'}(0, e_3) \neq 0,$$

then necessarily $\partial \varphi / \partial \xi_2(0, e_3) \neq 0$ and the equation $\varphi(x', \xi') = 0$ is equivalent to

$$(5.17) \quad \xi_2 = 0,$$

provided a suitable change of coordinates is performed near the x' -origin.

In particular we deduce that in this case $\text{rank } A(0) = 1$ so that, on Σ_p we also have $\text{rank } A(x') = 1$ near the origin.

II) *Assume that*

$$\frac{\partial \varphi}{\partial \xi'}(0, e_3) = 0,$$

and

$$\frac{\partial \varphi}{\partial x'}(0, e_3) \neq 0.$$

Then the following cases may occur:

(a) *The equation of Σ_p relatively to Σ_1 does not depend on ξ' , i.e. Σ_p is cylindrical with respect to the ξ' -fibers. Then we may change coordinates near the origin in such a way that, in Σ_1 , Σ_p is defined by the equation*

$$(5.18) \quad x_j = 0,$$

where $j \in \{2, 3\}$.

(b) *Denote by $\varphi(x', \xi') = 0$ the equation of Σ_p in Σ_1 . Then if $\partial \varphi / \partial x_2(0, e_3) \neq 0$ in a suitable system of coordinates near the origin φ is equivalent to*

$$(5.19) \quad Y(x', \xi') \equiv \lambda(x')\xi_2 + x_2\xi_3 = 0.$$

Here λ denotes a real analytic function such that $\lambda(0) = 0$.

On the other hand assume that $\partial\varphi/\partial x_2(0, e_3) = 0$ and that $\partial\varphi/\partial x_3(0, e_3) \neq 0$. Then the equation $\varphi = 0$ is equivalent to

$$(5.20) \quad Y(x', \xi') \equiv \lambda(x')\xi_2 + \mu(x')\xi_3 + x_3\xi_3 = 0,$$

where $\lambda(0) = 0$, $\mu(0) = 0$, $d_{x'}\mu(0) = 0$.

Proof. To prove the above statement we need only remark that in Case I any equation of the form $\xi_2 - \chi(x')\xi_3 = 0$ may be written as $\xi_2 = 0$, performing a change of coordinates that leaves x_1 unchanged.

As for Case IIa it suffices to notice that Σ_p is given, by what has been shown previously, by the equation $\beta(x') = 0$ with $d_{x'}\beta(0) \neq 0$. Thus we can always change coordinates in the (x_2, x_3) -plane in such a way that $\beta(x') = 0$ becomes $x_2 = 0$ if $\partial\beta/\partial x_2(0) \neq 0$, or $x_3 = 0$ otherwise.

Let us consider the Case IIb. If $\partial\varphi/\partial x_2(0, e_3) \neq 0$, we have

$$\varphi(x', \xi') = \left(x_2 a_{k2}^{(2)}(x') + x_3 a_{k2}^{(3)}(x') \right) \xi_2 + \left(x_2 a_{k3}^{(2)}(x') + x_3 a_{k3}^{(3)}(x') \right) \xi_3$$

where $k = 2$ or $k = 3$ depending on which component of the 2-vector B in (5.10) is elliptic at $(0, e_3)$; moreover $a_{k3}^{(2)}(0) \neq 0$. Then we conclude that the equation $x_2 a_{k3}^{(2)}(x') + x_3 a_{k3}^{(3)}(x') = 0$ is equivalent to $x_2 - \chi(x_3) = 0$, for a suitable analytic function χ defined near the origin. Let us perform the following change of variables in the (x_2, x_3) -plane:

$$\begin{cases} y_2 = x_2 - \chi(x_3) \\ y_3 = x_3 \end{cases} \quad \begin{cases} \eta_2 = \xi_2 \\ \eta_3 = \xi_3 + \frac{\partial\chi(x_2)}{\partial x_3} \xi_2 \end{cases} .$$

Then in the new coordinates, modulo a non-vanishing factor, we have

$$\varphi(x', \xi') = e(x')(\lambda(x')\xi_2 + x_2\xi_3),$$

which gives (5.19).

Assume now that $\partial\varphi/\partial x_2(0, e_3) = 0$ and that $\partial\varphi/\partial x_3(0, e_3) \neq 0$. In the above expression of φ we then have $a_{k3}^{(3)}(0) \neq 0$ and $a_{k3}^{(2)}(0) = 0$, otherwise we would be in the same situation as above.

Thus

$$\varphi(x', \xi') = a_{k3}^{(3)}(x') [\lambda(x')\xi_2 + \mu(x')\xi_3 + x_3\xi_3],$$

with $\lambda(0) = 0$ and $\mu(x') = O(|x'|^2)$ which yields Equation (5.20). This completes the proof of the proposition. \square

Remark. The seemingly pedantic distinction between the x_2 and x_3 variable in the proof above will be useful in subsequent work, where we

shall be concerned with the Gevrey (analytic) hypoellipticity properties of our operators. The basic tool for us are microlocal a priori estimates and we shall see that, from a microlocal point of view, the Gevrey hypoellipticity thresholds for cases (5.19) and (5.20), near the same point $(0, e_3)$, are very different. Naturally, near different base points, both Cases IIa and IIb may occur for the same operator, yielding different microlocal hypoellipticity results and the expected local result.

Using Proposition 5.1 we can write the vector fields in a simpler way.

Proposition 5.2. *The vector fields X_1 , X_2 and X_3 satisfying hypotheses (A1)-(A4) can be written in the following way:*

Case I:

$$\begin{aligned}
 X_1(x, \xi) &= \xi_1 \\
 X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [\alpha(x')\xi_2 \\
 (5.21) \quad &\quad + x_1 \{ \tilde{a}_{22}(x)\xi_2 + \tilde{a}_{23}(x)\xi_3 \}] \\
 X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [\lambda(x')\alpha(x')\xi_2 \\
 &\quad + x_1 \{ \tilde{a}_{32}(x)\xi_2 + \tilde{a}_{33}(x)\xi_3 \}],
 \end{aligned}$$

for suitable functions $\alpha(x')$, with $\alpha(0) \neq 0$, and $\lambda(x')$.

Case IIa:

$$\begin{aligned}
 X_1(x, \xi) &= \xi_1 \\
 X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [x_j (\tilde{a}_{22}(0, x')\xi_2 + \tilde{a}_{23}(0, x')\xi_3) \\
 (5.22) \quad &\quad + x_1 \{ \hat{a}_{22}(x)\xi_2 + \hat{a}_{23}(x)\xi_3 \}] \\
 X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [x_j (\tilde{a}_{32}(0, x')\xi_2 + \tilde{a}_{33}(0, x')\xi_3) \\
 &\quad + x_1 \{ \hat{a}_{32}(x)\xi_2 + \hat{a}_{33}(x)\xi_3 \}].
 \end{aligned}$$

where j is equal to 2 or 3.

Case IIb:

$$\begin{aligned}
 X_1(x, \xi) &= \xi_1 \\
 X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [\alpha(x')Y(x', \xi') \\
 (5.23) \quad &\quad + x_1 \{ \hat{a}_{22}(x)\xi_2 + \hat{a}_{23}(x)\xi_3 \}] \\
 X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [\beta(x')Y(x', \xi') \\
 &\quad + x_1 \{ \hat{a}_{32}(x)\xi_2 + \hat{a}_{33}(x)\xi_3 \}].
 \end{aligned}$$

where α is a non-vanishing analytic function defined in a neighborhood of the origin, β is analytic and $Y(x', \xi')$ is a vector field of the form (5.19) or (5.20).

Proof. Case I is straightforward, due to Proposition 5.1. The same proposition also implies Case IIa. Case IIb follows from Proposition 5.1 and Lemma 5.1 (b). \square

Remark. We point out that, since $\Sigma_1 = \{x_1 = 0, \xi_1 = 0\}$, the forms (5.21) - (5.23) for our vector fields actually have some further properties, which will turn out to be important for the regularity estimates. Basically these properties state that the fields are linearly independent outside of the characteristic manifold and that the number of layers of the Poisson stratification is finite. We postpone a precise statement of this fact until the final step in order not to burden the exposition too much.

The next step consists in using Assumption (A3) and the remaining part of (A4) to make the form of the vector fields more precise.

6. FINER FORMS FOR THE VECTOR FIELDS

6.1. **Case I.** By Proposition 5.2 we are dealing with the fields:

$$\begin{aligned} X_1(x, \xi) &= \xi_1 \\ \begin{bmatrix} X_2(x, \xi) \\ X_3(x, \xi) \end{bmatrix} &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 + x_1^{p-1} \left\{ \begin{bmatrix} \alpha(x') & 0 \\ \lambda(x')\alpha(x') & 0 \end{bmatrix} \xi' + x_1 \tilde{A}(x) \xi' \right\}, \end{aligned}$$

with obvious notation. We can see at once that the only brackets that matter are

$$\text{ad}^j(X_1)X_k, \quad k = 2, 3, \quad j = p, p+1, \dots, q-2.$$

The above quantity vanishes on $\Sigma_p = \{x_1 = \xi_1 = 0, \xi_2 = 0\}$, so that, taking $j = p$, we conclude that

$$\tilde{A}(x') \begin{bmatrix} 0 \\ \xi_3 \end{bmatrix} = 0,$$

which implies that

$$\tilde{a}_{23}(x) \Big|_{x_1=0} = \tilde{a}_{33}(x) \Big|_{x_1=0} = 0.$$

Thus we may write the fields as

$$\begin{aligned} X_1(x, \xi) &= \xi_1 \\ \begin{bmatrix} X_2(x, \xi) \\ X_3(x, \xi) \end{bmatrix} &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 + x_1^{p-1} \left\{ \begin{bmatrix} \alpha(x') + x_1 \tilde{a}_{22}(x) & 0 \\ \lambda(x')\alpha(x') + x_1 \tilde{a}_{32}(x) & 0 \end{bmatrix} \xi' \right. \\ &\quad \left. + x_1 \begin{bmatrix} \tilde{a}_{23}(x) \\ \tilde{a}_{33}(x) \end{bmatrix} \xi_3 \right\}, \end{aligned}$$

for suitable analytic coefficients \tilde{a}_{i3} , $i = 1, 2$.

Proceeding analogously and using the remaining brackets, we conclude that

$$\begin{aligned} X_1(x, \xi) &= \xi_1 \\ \begin{bmatrix} X_2(x, \xi) \\ X_3(x, \xi) \end{bmatrix} &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 + x_1^{p-1} \left\{ \begin{bmatrix} \alpha(x') + x_1 \tilde{a}_{22}(x) & 0 \\ \lambda(x')\alpha(x') + x_1 \tilde{a}_{32}(x) & 0 \end{bmatrix} \xi' \right. \\ (6.1.1) \quad &\quad \left. + x_1^{q-p} \begin{bmatrix} \tilde{a}_{23}(x) \\ \tilde{a}_{33}(x) \end{bmatrix} \xi_3 \right\}, \end{aligned}$$

for suitable analytic functions $\alpha(x') \neq 0$ (as always in Case I - cf.(5.21)), \tilde{a}_{i3} , $i = 1, 2$, and $\lambda(x')$.

Furthermore the ellipticity of the Poisson brackets of length q tells us that

$$(6.1.2) \quad \tilde{A}(0, x') \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq 0.$$

On the other hand, the fields X_2, X_3 in (6.1.1) are linearly independent for $x_1 \neq 0$ if and only if

$$x_1^{q-p} \det \begin{bmatrix} \alpha(x') + x_1 \tilde{a}_{22}(x) & \tilde{a}_{23}(x) \\ \lambda(x')\alpha(x') + x_1 \tilde{a}_{32}(x) & \tilde{a}_{33}(x) \end{bmatrix} \neq 0,$$

i.e.

$$-\lambda(x')\tilde{a}_{23}(x) + \tilde{a}_{33}(x) + \frac{x_1}{\alpha(x')} \det \tilde{A}(x) \neq 0,$$

if $x_1 \neq 0$, or

$$(6.1.3) \quad \det \left(\begin{bmatrix} 1 \\ \lambda(x') \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{x_1}{\alpha(x')} \tilde{A}(x) \right) \neq 0,$$

if $x_1 \neq 0$. Another way of stating the above condition is

$$(6.1.4) \quad \left\langle \begin{bmatrix} -\lambda(x') \\ 1 \end{bmatrix}, \tilde{A}(x) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle + \frac{x_1}{\alpha(x')} \det \tilde{A}(x) \neq 0,$$

if $x_1 \neq 0$.

6.2. **Case IIa.** We begin by considering the fields in (5.22) and again use Assumption (A4) and (A3). Thanks to the remarks made above, we can see that, taking p derivatives with respect to x_1 , we have

$$\hat{A}(x)\xi' = 0 \quad \text{if} \quad x_1 = x_j = 0, \quad j = 2, 3,$$

i.e.

$$(6.2.1) \quad \hat{A}(x) = x_1 \hat{A}_1(x) + x_j \hat{A}_2(x).$$

Hence X_2 and X_3 can be written:

$$\begin{aligned} \begin{bmatrix} X_2(x, \xi) \\ X_3(x, \xi) \end{bmatrix} &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 + x_1^{p-1} \left\{ x_j \tilde{A}(x') \xi' \right. \\ &\quad \left. + \left(x_1^2 \hat{A}_1(x) + x_1 x_j \hat{A}_j(x) \right) \xi' \right\} \\ &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 + x_1^{p-1} \left\{ x_j \tilde{A}(x) \xi' + x_1^2 \hat{A}(x) \xi' \right\}, \end{aligned}$$

the meaning of the symbols being obvious.

Iterating this argument we reach the following form for the vector fields:

$$(6.2.2) \quad \begin{aligned} X_1(x, \xi) &= \xi_1 \\ X_2(x, \xi) &= a_{21}(x) \xi_1 + x_1^{p-1} [x_j (\tilde{a}_{22}(x) \xi_2 + \tilde{a}_{23}(x) \xi_3) \\ &\quad + x_1^{q-p} (\hat{a}_{22}(x) \xi_2 + \hat{a}_{23}(x) \xi_3)] \\ X_3(x, \xi) &= a_{31}(x) \xi_1 + x_1^{p-1} [x_j (\tilde{a}_{32}(x) \xi_2 + \tilde{a}_{33}(x) \xi_3) \\ &\quad + x_1^{q-p} (\hat{a}_{32}(x) \xi_2 + \hat{a}_{33}(x) \xi_3)], \end{aligned}$$

$j \in \{2, 3\}$.

Proceeding as in Case I we see that the ellipticity of the last Poisson layer means that

$$(6.2.3) \quad \det \hat{A}(x) \Big|_{x_1=0, x_j=0} \neq 0.$$

On the other hand, Assumption (A3) together with (6.2.3) means that

$$(6.2.4) \quad \det \left(x_j \tilde{A}(x) + x_1^{q-p} \hat{A}(x) \right) \neq 0,$$

if $x_1 \neq 0$.

6.3. Case IIb. Let us consider the fields in (5.23) and use Assumptions (A3) and (A4). Taking the p -th derivative with respect to x_1 we obtain that

$$\hat{A}(x)\xi' = 0 \quad \text{if } x_1 = 0 \quad \text{and} \quad Y(x', \xi') = 0.$$

By Lemma 5.1 (b), this implies that there is an analytic 2-vector, $h^{(1)}(x')$, defined near the origin, such that

$$\hat{A}(0, x')\xi' = h^{(1)}(x')Y(x', \xi'), \quad h^{(1)}(x') = \begin{bmatrix} h_2^{(1)}(x') \\ h_3^{(1)}(x') \end{bmatrix},$$

so that

$$\hat{A}(x)\xi' = h^{(1)}(x')Y(x', \xi') + x_1\hat{A}^{(1)}(x)\xi'$$

and hence

$$\begin{aligned} \begin{bmatrix} X_2(x, \xi) \\ X_3(x, \xi) \end{bmatrix} &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 + \\ &x_1^{p-1} \left\{ \left(\begin{bmatrix} \alpha(x') \\ \beta(x') \end{bmatrix} + x_1 h^{(1)}(x') \right) Y(x', \xi') \right. \\ &\left. + x_1^2 \hat{A}^{(1)}(x)\xi' \right\} \end{aligned}$$

Iterating this argument we obtain that

$$\begin{aligned} X_1(x, \xi) &= \xi_1 \\ (6.3.1) \quad \begin{bmatrix} X_2(x, \xi) \\ X_3(x, \xi) \end{bmatrix} &= \begin{bmatrix} a_{21}(x) \\ a_{31}(x) \end{bmatrix} \xi_1 \\ &+ x_1^{p-1} \left\{ h(x)Y(x', \xi') + x_1^{q-p} \hat{A}(x)\xi' \right\}, \end{aligned}$$

where $h(x)$ is a 2-vector function, $h(x) = (h_2(x), h_3(x))$, such that $h_2(0) \neq 0$, and \hat{A} is a 2×2 matrix with real analytic entries defined near the origin.

Assumption (A4) then implies that $\hat{A}(x)\xi'$ cannot be zero if $x_1 = 0$, $\xi_1 = 0$ and $Y(x', \xi') = 0$; but, since $Y(0, \xi') \equiv 0$ for every $\xi' \in \mathbb{R}^2$, we easily get that

$$(6.3.2) \quad \det \hat{A}(0) \neq 0,$$

while the linear independence of the vector fields outside of Σ_1 yields

$$(6.3.3) \quad \det(h(x) \otimes Y(x') + x_1^{q-p} \hat{A}(x)) \neq 0,$$

if $x_1 \neq 0$. Here $Y(x')$ denotes the 2-vector whose components are the coefficients of the vector field Y .

We summarize the above argument in

Theorem 6.3.1. *Let X_1, X_2, X_3 satisfy Assumptions (A1) - (A4). Then there is a suitable system of coordinates defined in a neighborhood of the point $(0, e_3)$, such that the field can be written in one of the following ways:*

Case I)

$$\begin{aligned}
 X_1(x, \xi) &= \xi_1 \\
 (6.3.4) \quad X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [(\alpha(x') + x_1\tilde{a}_{22}(x))\xi_2 \\
 &\quad + x_1^{q-p}\tilde{a}_{23}(x)\xi_3] \\
 X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [(\lambda(x')\alpha(x') + x_1\tilde{a}_{32}(x))\xi_2 \\
 &\quad + x_1^{q-p}\tilde{a}_{33}(x)\xi_3],
 \end{aligned}$$

for suitable analytic functions \tilde{a}_{ij} , $i, j = 2, 3$, $\lambda(x')$, and $\alpha(x') \neq 0$.
Moreover we have

$$(6.3.5) \quad \begin{bmatrix} \tilde{a}_{23}(0, x') \\ \tilde{a}_{33}(0, x') \end{bmatrix} \neq 0,$$

and

$$(6.3.6) \quad -\lambda(x')\tilde{a}_{23}(x) + \tilde{a}_{33}(x) + \frac{x_1}{\alpha(x')} \det \tilde{A}(x) \neq 0,$$

if $x_1 \neq 0$.

Case IIa)

$$\begin{aligned}
 X_1(x, \xi) &= \xi_1 \\
 X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [x_j (\tilde{a}_{22}(x)\xi_2 + \tilde{a}_{23}(x)\xi_3) \\
 (6.3.7) \quad &\quad + x_1^{q-p} (\hat{a}_{22}(x)\xi_2 + \hat{a}_{23}(x)\xi_3)] \\
 X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [x_j (\tilde{a}_{32}(x)\xi_2 + \tilde{a}_{33}(x)\xi_3) \\
 &\quad + x_1^{q-p} (\hat{a}_{32}(x)\xi_2 + \hat{a}_{33}(x)\xi_3)],
 \end{aligned}$$

where $j \in \{2, 3\}$, \tilde{a}_{ij} , \hat{a}_{ij} are analytic functions, $i, j = 2, 3$, such that

$$(6.3.8) \quad \det \tilde{A}(x) \Big|_{x_1=0} \neq 0, \quad \det \hat{A}(x) \Big|_{x_1=x_j=0} \neq 0.$$

Moreover

$$(6.3.9) \quad \det \left(x_j \tilde{A}(x) + x_1^{q-p} \hat{A}(x) \right) \neq 0,$$

if $x_1 \neq 0$.

Case IIb)

$$(6.3.10) \quad \begin{aligned} X_1(x, \xi) &= \xi_1 \\ X_2(x, \xi) &= a_{21}(x)\xi_1 + x_1^{p-1} [h_2(x) (\alpha(x')\xi_2 + \beta(x')\xi_3) \\ &\quad + x_1^{q-p} (\hat{a}_{22}(x)\xi_2 + \hat{a}_{23}(x)\xi_3)] \\ X_3(x, \xi) &= a_{31}(x)\xi_1 + x_1^{p-1} [h_3(x) (\alpha(x')\xi_2 + \beta(x')\xi_3) \\ &\quad + x_1^{q-p} (\hat{a}_{32}(x)\xi_2 + \hat{a}_{33}(x)\xi_3)] \end{aligned}$$

where we may assume that $h_2(0) \neq 0$, h_j and \hat{a}_{ij} are suitable analytic functions, and the field $\alpha(x')\xi_2 + \beta(x')\xi_3$ has the form in (5.19) or (5.20). Moreover

$$(6.3.11) \quad \det \hat{A}(0) \neq 0$$

and

$$(6.3.12) \quad \det \left(h(x) \otimes \begin{bmatrix} \alpha(x') \\ \beta(x') \end{bmatrix} + x_1^{q-p} \hat{A}(x) \right) \neq 0,$$

if $x_1 \neq 0$.

7. EXAMPLES

We collect in this section a few examples of the fields obtained in Theorem 6.3.1. The Case I examples all have the following stratification:

$$\begin{aligned} \Sigma_1 &= \{x_1 = \xi_1 = 0\} \\ \Sigma_2 &= \Sigma_1 \\ &\vdots \\ \Sigma_p &= \{x_1 = \xi_1 = 0, \xi_2 = 0\} \\ \Sigma_{p+1} &= \Sigma_p \\ &\vdots \\ \Sigma_q &= \{0\}, \end{aligned}$$

where $\{0\}$ denotes the zero section of the cotangent bundle.

7.1. Case I.

- Let $\alpha \equiv 1$, $\lambda = 0$, $a_{21} = a_{31} = 0$ and

$$\tilde{A} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Then we have the fields

$$\xi_1, \quad x_1^{p-1} [\xi_2 + x_1^{q-p} \xi_3], \quad x_1^{q-1} \xi_3.$$

- Let $\alpha \equiv 1$, $\lambda = 0$, $a_{21} = a_{31} = 0$ and

$$\tilde{A} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then we have the fields

$$\xi_1, \quad x_1^{p-1} \xi_2, \quad x_1^{q-1} \xi_3,$$

which is the Oleinik-Radkevič operator.

- Let $\alpha \equiv 1$, $\lambda = 0$, $a_{21} = a_{31} = 0$ and

$$\tilde{A} = \begin{bmatrix} 0 & 1 \\ x_1^{q-p-1} & 0 \end{bmatrix}.$$

Then we have the fields

$$\xi_1, \quad x_1^{p-1} [\xi_2 + x_1^{q-p} \xi_3], \quad x_1^{q-1} \xi_2.$$

Concerning the conditions of Theorem 6.3.1 we see that the vector $(\tilde{a}_{23}, \tilde{a}_{33})$ is equal to $(1, 1)$ in the first case, $(0, 1)$ in the second case and to $(1, 0)$ in the third case. Moreover (6.3.6) reads as $1 + x_1 \cdot 0 \neq 0$ in the first and second cases, $x_1 \det \tilde{A} = -x_1^{q-p} \neq 0$ if $x_1 \neq 0$ in the third case.

7.2. Case IIa. For the Case IIa, the stratification is as for Case I except that Σ_p is now given by:

$$\Sigma_p = \{x_1 = \xi_1 = 0, x_2 = 0\}.$$

Let us take $j = 2$, $a_{21} = a_{31} = 0$ and $\tilde{A} = Id$. Then from the condition $\det(x_2 Id + x_1^{q-p} \hat{A}) \neq 0$ if $x_1 \neq 0$ we easily deduce that the matrix \hat{A} must have non-zero strictly complex eigenvalues. Set

$$\hat{A} = \begin{bmatrix} \lambda & \mu \\ -\mu & \lambda \end{bmatrix}, \quad \mu \neq 0.$$

Then our conditions are satisfied and we obtain the fields

$$\xi_1, \quad x_1^{p-1} [x_2 \xi_2 + x_1^{q-p} (\lambda \xi_2 + \mu \xi_3)], \\ x_1^{p-1} [x_2 \xi_3 + x_1^{q-p} (-\mu \xi_2 + \lambda \xi_3)].$$

7.3. Case IIb. Here the non-symplectic layer Σ_p is given near $(0, \xi_3)$ by:

$$\Sigma_p = x_3^2 \xi_2 + x_2 \xi_3 = 0.$$

Let $\varphi(x', \xi') = \alpha(x') \xi_2 + \beta(x') \xi_3 = \lambda(x') \xi_2 + x_2 \xi_3$, with $\lambda \not\equiv 0$, $\lambda(0) = 0$, as e.g. in (5.19); we may assume that $\lambda(x')/x_2$ is not an analytic function near the origin.

Moreover let $a_{21} = a_{31} = 0$, $h_2 = 1$ and $h_3 = 0$. Then we have the fields

$$\xi_1, \quad x_1^{p-1} [\lambda \xi_2 + x_2 \xi_3 + x_1^{q-p} (\hat{a}_{22} \xi_2 + \hat{a}_{23} \xi_3)], \quad x_1^{q-1} [\hat{a}_{32} \xi_2 + \hat{a}_{33} \xi_3].$$

Conditions (6.3.11) and (6.3.12) become $\det \hat{A} \neq 0$ and $\lambda \hat{a}_{33} - \hat{a}_{32} x_2 + x_1^{q-p} \det \hat{A} \neq 0$ if $x_1 \neq 0$. If $q - p$ is e.g. even we may choose $\lambda = x_3^2$, $\hat{a}_{33} = \text{sign det } \hat{A}$, $\hat{a}_{32} = 0$ to write a particular case of the above fields.

8. THE BEHAVIOR OF THE BICHARACTERISTIC CURVES AND A FINER CLASSIFICATION

In this Section we present a classification of the various instances of the ‘‘sums of squares operators’’ in which we get in Case I.

Consider (6.3.4); X_1 actually denotes the only non characteristic vector field. Let us consider the null bicharacteristic curves of X_1 , $\gamma_{(\bar{x}, \bar{\xi})}(t) = (\bar{x}, \bar{\xi}) + t(e_1, 0)$, where $\bar{\xi}_1 = 0$. If $\bar{x}_1 = 0$, then $\gamma_{(\bar{x}, \bar{\xi})}(t) = \gamma_{(\bar{x}', \bar{\xi}')} (t) = (0, \bar{x}', 0, \bar{\xi}') + t(e_1, 0)$ is actually a null bicharacteristic curve of X_1 issued from a point $(0, \bar{x}', 0, \bar{\xi}')$ of Σ_1 . Assume $t \neq 0$ and compute X_2 and X_3 on such a curve. We obtain

$$(8.1) \quad \begin{aligned} X_2(\gamma_{(\bar{x}', \bar{\xi}')} (t)) &= t^{p-1} [(\alpha(\bar{x}') + t \tilde{a}_{22}(t, \bar{x}')) \bar{\xi}_2 \\ &\quad + t^{q-p} \tilde{a}_{23}(t, \bar{x}') \bar{\xi}_3] \\ X_3(\gamma_{(\bar{x}', \bar{\xi}')} (t)) &= t^{p-1} [(\lambda(\bar{x}') \alpha(\bar{x}') + t \tilde{a}_{32}(t, \bar{x}')) \bar{\xi}_2 \\ &\quad + t^{q-p} \tilde{a}_{33}(t, \bar{x}') \bar{\xi}_3]. \end{aligned}$$

Assume that the point $(0, \bar{x}', 0, \bar{\xi}')$ is in a neighborhood of $(0, e_3)$. Then $\xi_3 \neq 0$ and also $\alpha(\bar{x}') \neq 0$ by Theorem 6.3.1. On the other hand nothing

is known a priori about the function λ . We point out explicitly that we chose X_2 as the field having a non-zero $\partial/\partial x_2$ coefficient near $(0, e_3)$, thus breaking the X_2 - X_3 symmetry. This is evidently no restriction of generality, provided we bear in mind that analogous statements hold if we interchange the roles of X_2 and X_3 .

When $t \neq 0$ we may consider the characteristic set of $X_2(\gamma_{(\bar{x}', \bar{\xi}')} (t))$; we obtain that $X_2(\gamma_{(\bar{x}', \bar{\xi}')} (t)) = 0$ if and only if

$$\bar{\xi}_2 = -t^{q-p} \frac{\tilde{a}_{23}(t, \bar{x}')}{\alpha(\bar{x}') + t\tilde{a}_{22}(t, \bar{x}')} \bar{\xi}_3.$$

Let us now compute $X_3(\gamma_{(\bar{x}', \bar{\xi}')} (t)) \Big|_{X_2(\gamma_{(\bar{x}', \bar{\xi}')} (t))=0}$; we get

$$(8.2) \quad \frac{\alpha(\bar{x}')}{\alpha(\bar{x}') + t\tilde{a}_{22}(t, \bar{x}')} t^{q-1} \left[-\lambda(\bar{x}') \tilde{a}_{23}(t, \bar{x}') + \tilde{a}_{33}(t, \bar{x}') + \frac{t}{\alpha(\bar{x}')} \det \tilde{A}(t, \bar{x}') \right] \bar{\xi}_3$$

where the quantity in square brackets is that playing a role in Equation (6.3.6) and is non-zero provided $t \neq 0$. We also point out that the coefficient $\alpha(\bar{x}')(\alpha(\bar{x}') + t\tilde{a}_{22}(t, \bar{x}'))^{-1}$ is also non-zero at $t = 0$.

The above discussion motivates the following

Definition 8.1. *We say that the fields X_1, X_2, X_3 of (6.3.4) are in Case I_0 or of type I_0 if*

$$(8.3) \quad -\lambda(0)\tilde{a}_{23}(0, 0) + \tilde{a}_{33}(0, 0) \neq 0.$$

This means that, as $t \rightarrow 0$

$$X_3(\gamma_{(\bar{x}', \bar{\xi}')} (t)) \Big|_{X_2(\gamma_{(\bar{x}', \bar{\xi}')} (t))=0} \sim t^{q-1},$$

uniformly with respect to $\bar{x}', \bar{\xi}_3 \neq 0$.

Assume now that (8.3) no longer holds and let

$$(8.4) \quad -\lambda(0)\tilde{a}_{23}(t, 0) + \tilde{a}_{33}(t, 0) + \frac{t}{\alpha(0)} \det \tilde{A}(t, 0) \sim t^r,$$

as $t \rightarrow 0$. Then we say that the fields X_1, X_2, X_3 of (6.3.4) are in case I_r or of type I_r , $r > 0$. This implies that

$$X_3(\gamma_{(\bar{x}', \bar{\xi}')} (t)) \Big|_{X_2(\gamma_{(\bar{x}', \bar{\xi}')} (t))=0} \sim t^{q-1+r},$$

for $t \rightarrow 0$ and \bar{x}' in a small neighborhood of the origin.

The first and second examples in Section 7.1 for Case I operators are of type I_0 , while the third example is of type I_{q-p} .

We will find this property to be relevant for the Gevrey hypoellipticity threshold of the corresponding sums of squares operators.

9. GEVREY REGULARITY FOR SUMS OF SQUARES OF VECTOR FIELDS OF TYPE I_0

In this section our purpose is to deduce microlocal Gevrey estimates for operators of type I_0 . For the sake of simplicity we slightly modify our notation in (6.3.4). Thus let us consider three vector fields of the form

$$(9.1) \quad \begin{aligned} X_1(x, D) &= D_1 \\ X_2(x, D) &= a_{21}(x)D_1 + x_1^{p-1}f_2(x)D_2 + x_1^{q-1}g_2(x)D_3 \\ X_3(x, D) &= a_{31}(x)D_1 + x_1^{p-1}f_3(x)D_2 + x_1^{q-1}g_3(x)D_3, \end{aligned}$$

where f_j and g_j , $j = 2, 3$, are real analytic functions defined in a neighborhood of the origin and such that (6.3.5) becomes

$$(9.2) \quad \begin{aligned} f_2(0, x') &\neq 0 \\ f_3(0, x') &= \lambda(x')f_2(0, x') \end{aligned}$$

and

$$(9.3) \quad \begin{bmatrix} g_2(0) \\ g_3(0) \end{bmatrix} \neq 0.$$

Moreover (6.3.6) becomes

$$(9.4) \quad -\lambda(x')g_2(x) + g_3(x) + \frac{x_1}{f_2(0, x')} \det \begin{bmatrix} \frac{f_2 - f_2(0, x')}{x_1} & g_2(x) \\ \frac{f_3 - f_3(0, x')}{x_1} & g_3(x) \end{bmatrix} \neq 0$$

if $x_1 \neq 0$. Now the assumption that our operator $\sum_{j=1}^3 X_j^2$ is of type I_0 means that

$$(9.5) \quad -\lambda(0)g_2(0) + g_3(0) \neq 0.$$

The latter implies (9.4), while (9.4) makes sense due to (9.3).

Lemma 9.1. *Let α , β and γ be real analytic functions defined in a neighborhood of the origin in \mathbb{R}^3 . Then we can find real analytic functions a , b and c such that*

$$\alpha(x)D_1 + \beta(x)x_1^{p-1}D_2 + \gamma(x)x_1^{q-1}D_3 = a(x)X_1 + b(x)X_2 + c(x)X_3.$$

Proof. This very useful lemma is a simple consequence of the assumptions, and says that the span of the vector fields $\{D_1, x^{p-1}D_2, x_{q-1}D_3\}$ is that same as that of the vector fields $\{X_j\}$. Using elementary row and column operations on the matrix on the right hand side of (9.1) the Lemma states the invertibility of the matrix

$$(9.6) \quad \begin{bmatrix} f_2(x) & g_2(x) \\ f_3(x) & g_3(x) \end{bmatrix}$$

which, in view of (9.2) is equivalent to the invertibility of the matrix

$$(9.7) \quad \begin{bmatrix} 1 & g_2(x) \\ 0 & g_3(x) - \lambda(x)g_2(x) \end{bmatrix}.$$

But this is just (9.5) (all locally). \square

Lemma 9.2. *For $j = 1, 2, 3$, and m an integer,*

$$(9.8) \quad [X_j, D_3^m] = \sum_{\ell=1}^m \binom{m}{\ell} \sum_{h=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h D_3^{m-\ell},$$

where

$$(9.9) \quad |\partial^\alpha \tilde{\gamma}_{jh}^{(\ell)}| \lesssim C_{jh}^{\ell+|\alpha|} (\ell + |\alpha|)!$$

Equivalently,

$$(9.10) \quad D_3^m X_j = - \sum_{\ell=0}^m \binom{m}{\ell} \sum_{h=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h D_3^{m-\ell},$$

where $\tilde{\gamma}_{jh}^{(0)} = -\delta_{jh}$.

Proof. This is just an iteration of the previous Lemma. \square

Let denote by φ a cut off function identically equal to one in a neighborhood of the origin in \mathbb{R}^3 . Due to the special form of our coordinates and the fact that the characteristic manifold is symplectic, we may assume that φ is independent of the variable x_1 : in fact we may always take φ as a product of three such cut off functions each depending on a single coordinate, x_j , and every x_1 -derivative landing on $\varphi(x_1)$ would leave a cut off supported in a region where x_1 is bounded away from zero, hence in a region where the operator is (uniformly, microlocally) elliptic. Thus we take $\varphi(x) = \varphi(x')$. Here φ is assumed to be a function of Ehrenpreis-Hörmander type (see e.g. [15], [20]), i.e., denoting by U our neighborhood of the origin, then φ_j has the following property: for

any \tilde{U} compactly contained in U , and for any fixed $r \in \mathbb{N}$, we choose $\varphi_j = \varphi_{j,r} \in C_0^\infty(U)$, $\varphi \equiv 1$ on \tilde{U} and such that, with a *universal* constant (i.e., depending only on the dimension of the Euclidean space in which we work) C_0 such that

$$|\varphi^{(k)}(x)| \leq \left(\frac{C_0}{\text{dist}(\tilde{U}, U^c)} \right)^{k+1} r^k \text{ for } k \leq 3r.$$

It is a well known fact that the operator

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

is C^∞ -hypoelliptic and satisfies an a priori estimate of the form

$$(9.11) \quad \sum_{j=1}^3 \|X_j u\|^2 + \|u\|_{1/q}^2 \leq C (|\langle Pu, u \rangle| + \|u\|^2),$$

where u is a rapidly decreasing smooth function, $\|\cdot\|_s$ denotes the usual Sobolev norm of order s and $\|\cdot\| = \|\cdot\|_0$ is the L^2 norm.

We want to obtain a bound for an expression of the form

$$(9.12) \quad \|X_j \varphi(x') D_3^r u\|,$$

where, since we are in a microlocal neighborhood of the point $(0, e_3)$, D_3 is an elliptic operator. It is well known that obtaining a bound for (9.12) of the type $\|X_j \varphi(x') D_3^r u\| \leq C^{r+1} r!^s$ allows us to deduce that P is Gevrey (micro-)hypoelliptic of order s .

Remark 1. *We would like to mention here that in the case of the second example of Section 7.1, i.e. the Oleinik-Radkevič operator, the authors in [3] proved that one has $G^{q/p}$ hypoellipticity and that this bound is optimal.*

Instead of bounding the quantity in (9.12), for technical reasons we want to bound the more general quantity:

$$(9.13) \quad \|X_j x_1^a \varphi^{(b)} D_3^{r-c} u\| + \|x_1^a \varphi^{(b)} D_3^{r-c} u\|_{1/q}$$

where a, b and c are positive integers with $a \leq q$ but b and c bounded only by r . Using (9.11), we see that (9.13) is bounded by

$$(9.14) \quad |\langle P x_1^a \varphi^{(b)} D_3^{r-c} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle| + \|x_1^a \varphi^{(b)} D_3^{r-c} u\|^2,$$

modulo a positive constant in front of everything appearing in the above formula. We need to move P in (9.14) to the right (onto u); the term with the L^2 norm will be easier to handle. Writing $P = \Sigma X^2$ and then $[X^2, V] = X[X, V] + [X, V]X$ with $V = x_1^a \varphi^{(b)} D_3^{r-c}$, we find

$$(9.15) \quad \begin{aligned} \langle P x_1^a \varphi^{(b)} D_3^{r-c} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle &= \langle x_1^a \varphi^{(b)} D_3^{r-c} P u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle \\ &+ \sum_{j=1}^3 \langle X_j [X_j, x_1^a \varphi^{(b)} D_3^{r-c}] u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle \\ &+ \sum_{j=1}^3 \langle [X_j, x_1^a \varphi^{(b)} D_3^{r-c}] X_j u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle. \end{aligned}$$

The first of the above right hand side terms is good, since we assume Pu to be analytic, even 0.

The second and third terms on the right hand side in (9.15) have many common features, which we may treat with the help of Lemma 9.2.

For $j=1$, we have

$$(9.16) \quad [X_1, x_1^a \varphi^{(b)} D_3^{r-c}] = a x_1^{a-1} \varphi^{(b)} D_3^{r-c},$$

and so

$$(9.17) \quad [X_1, x_1^a \varphi^{(b)} D_3^{r-c}] X_1 = X_1 a x_1^{a-1} \varphi^{(b)} D_3^{r-c} - a(a-1) x_1^{a-2} \varphi^{(b)} D_3^{r-c}.$$

For $j = 2, 3$,

$$(9.18) \quad [X_j, x_1^a \varphi^{(b)} D_3^{r-c}] = f(x) x_1^{p-1} x_1^a \varphi^{(b+1)} D_3^{r-c} + x_1^a \varphi^{(b)} \sum_{\ell=1}^{r-c} \binom{r-c}{\ell} \sum_{h=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h D_3^{r-c-\ell},$$

with f analytic, and thus using Lemma 9.2 again, ($j = 2, 3$)

$$\begin{aligned}
(9.19) \quad & [X_j, x_1^a \varphi^{(b)} D_3^{r-c}] X_j = f(x) x_1^{p-1} x_1^a \varphi^{(b+1)} D_3^{r-c} X_j + \\
& + x_1^a \varphi^{(b)} \sum_{\ell=1}^{r-c} \binom{r-c}{\ell} \sum_{h=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h D_3^{r-c-\ell} X_j = \\
& = f(x) x_1^{p-1} x_1^a \varphi^{(b+1)} \sum_{\ell=0}^{r-c} \binom{r-c}{\ell} \sum_{h=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h D_3^{r-c-\ell} + \\
& + x_1^a \varphi^{(b)} \sum_{\ell=1}^{r-c} \binom{r-c}{\ell} \sum_{h=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h \sum_{\ell_1=0}^{r-c-\ell} \binom{r-c-\ell}{\ell_1} \sum_{k=1}^3 \tilde{\gamma}_{jk}^{(\ell_1)} X_k D_3^{r-c-\ell-\ell_1}.
\end{aligned}$$

Going back to (9.15), the first term we have seen is harmless as it contains Pu . In the second, we integrate by parts and use a weighted Schwarz inequality. Since X_j^* is equal to $-X_j$ modulo a zero order term, the second term on the right in (9.15), using (9.16) and (9.18), becomes

$$\begin{aligned}
(9.20) \quad & \left| \sum_{j=1}^3 \langle X_j [X_j, x_1^a \varphi^{(b)} D_3^{r-c}] u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle \right| \leq \\
& \leq \varepsilon \sum_{k=1}^3 \|X_k x_1^a \varphi^{(b)} D_3^{r-c} u\|^2 + C_\varepsilon \|a x_1^{a-1} \varphi^{(b)} D_3^{r-c} u\|^2 + \\
& + C_\varepsilon \|f(x) x_1^{p-1} x_1^a \varphi^{(b+1)} D_3^{r-c} u\|^2 + \\
& + C_\varepsilon \sum_{\substack{1 \leq h \leq 3 \\ 2 \leq j \leq 3}} \|x_1^a \varphi^{(b)} \sum_{\ell=1}^{r-c} \binom{r-c}{\ell} \tilde{\gamma}_{jh}^{(\ell)} X_h D_3^{r-c-\ell} u\|^2.
\end{aligned}$$

This expression we leave for the moment and treat the issues which arise in the double commutator needed for the last term in (9.15), those which have already been expanded in (9.17) and (9.19).

We may continue with (9.17) in (9.15):

$$\begin{aligned}
(9.21) \quad & \langle [X_1, x_1^a \varphi^{(b)} D_3^{r-c}] X_1 u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle = \\
& = \langle X_1 a x_1^{a-1} \varphi^{(b)} D_3^{r-c} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle \\
& \quad - \langle a(a-1) x_1^{a-2} \varphi^{(b)} D_3^{r-c} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle.
\end{aligned}$$

We shall also continue with (9.19) in (9.15): for $j = 2, 3$

$$\begin{aligned}
(9.22) \quad & \langle [X_j, x_1^a \varphi^{(b)} D_3^{r-c}] X_j u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle = \\
& = \sum_{\substack{\ell=0 \\ h=1,2,3}}^{r-c} \binom{r-c}{\ell} \langle f(x) \tilde{\gamma}_{jh}^{(\ell)} x_1^{p-1} x_1^a \varphi^{(b+1)} X_h D_3^{r-c-\ell} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle \\
& + \sum_{\ell+\ell_1=1}^{r-c} \binom{r-c}{\ell, \ell_1} \langle x_1^a \varphi^{(b)} \sum_{h,k=1}^3 \tilde{\gamma}_{jh}^{(\ell)} X_h \tilde{\gamma}_{jk}^{(\ell_1)} X_k D_3^{r-c-\ell-\ell_1} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle.
\end{aligned}$$

Here we have used the ‘multinomial’ notation for brevity:

$$\binom{\alpha}{\beta, \gamma} = \frac{\alpha!}{\beta! \gamma! (\alpha - \beta - \gamma)!} = \binom{\alpha}{\beta} \binom{\alpha - \beta}{\gamma}.$$

Before collecting our individual terms we throw in a kind of ‘symmetrization’ of the first term on the left, for errors will often appear in this form. In so doing, we will encounter one more commutator, which is covered under the fourth and fifth terms on the right, hence contributing nothing new. We also drop the subscripts on the vector fields now. From (9.13), (9.14), and (9.15), (9.20), (9.21), (9.22) we have, for any positive ε ,

$$\begin{aligned}
(9.23) \quad & \|X x_1^a \varphi^{(b)} D_3^{r-c} u\| + \|x_1^a \varphi^{(b)} D_3^{r-c} u\|_{1/q} + \|x_1^a \varphi^{(b)} X D_3^{r-c} u\| \lesssim \\
& \lesssim \|x_1^a \varphi^{(b)} D_3^{r-c} P u\| + \|x_1^a \varphi^{(b)} D_3^{r-c} u\| \\
& + \varepsilon \|X x_1^a \varphi^{(b)} D_3^{r-c} u\| + C_\varepsilon \|a x_1^{a-1} \varphi^{(b)} D_3^{r-c} u\| + \\
& + C_\varepsilon \|f(x) x_1^{p-1} x_1^a \varphi^{(b+1)} D_3^{r-c} u\| + \\
& + C_\varepsilon \sum_{2 \leq j \leq 3} \|x_1^a \varphi^{(b)} \sum_{\ell=1}^{r-c} \binom{r-c}{\ell} \tilde{\gamma}_{jh}^{(\ell)} X D_3^{r-c-\ell} u\| + \\
& + |\langle X a x_1^{a-1} \varphi^{(b)} D_3^{r-c} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle|^{1/2} \\
& + |\langle a(a-1) x_1^{a-2} \varphi^{(b)} D_3^{r-c} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle|^{1/2} + \\
& + \sum_{\substack{\ell=0 \\ j=2,3}}^{r-c} \binom{r-c}{\ell} |\langle f(x) \tilde{\gamma}_{jh}^{(\ell)} x_1^{p-1} x_1^a \varphi^{(b+1)} X D_3^{r-c-\ell} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle|^{1/2} \\
& + \sum_{\substack{\ell+\ell_1=1 \\ j=2,3}}^{r-c} \binom{r-c}{\ell, \ell_1} |\langle x_1^a \varphi^{(b)} \tilde{\gamma}_j^{(\ell)} X \tilde{\gamma}_j^{(\ell_1)} X D_3^{r-c-\ell-\ell_1} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle|^{1/2} = \\
& = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9.
\end{aligned}$$

9.1. **The term I_1 .** This term is harmless since Pu is real analytic, even zero, in the support of all φ .

9.2. **The term I_2 .** This term will be bounded by a small multiple of (9.13) if we take the support of all the localizing functions small, and hence may be absorbed.

9.3. **The term I_3 .** This term is already a small multiple of (9.13), hence absorbable for ε small.

9.4. **The term I_4 .** This term, $C_\varepsilon \|ax_1^{a-1}\varphi^{(b)}D_3^{r-c}u\|$, exhibits an overall gain (in the norm) of $1/q$, but pays for it with a decrease in the power of x_1 . We will consider this term further below.

9.5. **The term I_5 .** This term, bounded at once by

$$C_f C_\varepsilon \|x_1^{p-1}x_1^a\varphi^{(b+1)}D_3^{r-c}u\|,$$

suffers a new derivative on φ but gains the factor x_1^{p-1} ,

9.6. **The term I_6 .** This term, easily bounded by

$$\tilde{C}_\varepsilon \sup_{1 \leq \ell \leq r-c} (C_\gamma r)^\ell \|x_1^a\varphi^{(b)}XD_3^{r-c-\ell}u\|$$

in view of the estimates (9.9), where C_γ depends only on the coefficients of the X_j and their first few derivatives. This term will be further treated under I_8 below, where also the term with $\ell = 0$ appears, though with a small constant in front.

9.7. **The term I_7 .** This term,

$$|\langle a(a-1)x_1^{a-2}\varphi^{(b)}D_3^{r-c}u, x_1^a\varphi^{(b)}D_3^{r-c}u \rangle|^{1/2},$$

is bounded exactly as is I_4 above once one power of a is moved to the left and the Schwarz inequality applied.

9.8. **The term I_8 .** This term,

$$\sum_{\substack{\ell=0 \\ j=2,3}}^{r-c} \binom{r-c}{\ell} |\langle f(x)\tilde{\gamma}_j^{(\ell)}x_1^{p-1}x_1^a\varphi^{(b+1)}XD_3^{r-c-\ell}u, x_1^a\varphi^{(b)}D_3^{r-c}u \rangle|^{1/2},$$

permits us to move $x_1^{p-1}\varphi^{(b+1)}$ to the right and $\varphi^{(b)}$ to the left, apply the Schwarz inequality and bring both f and $\tilde{\gamma}_j^{(\ell)}$ out of the norm. The result is

$$(9.24) \quad \varepsilon \|x^a \varphi^{(b)} X D_3^{r-c} u\| + \sup_{\ell \geq 1} (C_\gamma r)^\ell \|x^a \varphi^{(b)} X D_3^{r-c-\ell} u\| \\ + C_\varepsilon \|x_1^{a+p-1} \varphi^{(b+1)} D_3^{r-c} u\|.$$

The last of these is exactly like I_5 above, while the supremum, has been met in I_6 above. The first term, which we note carries the small constant ε , will be absorbed on the left hand side of (9.13) once the X is commuted to the left.

9.9. The term I_9 . This term,

$$\sum_{\substack{\ell+\ell_1=1 \\ j=2,3}}^{r-c} \binom{r-c}{\ell, \ell_1} |\langle x_1^a \varphi^{(b)} \tilde{\gamma}_j^{(\ell)} X \tilde{\gamma}_j^{(\ell_1)} X D_3^{r-c-\ell-\ell_1} u, x_1^a \varphi^{(b)} D_3^{r-c} u \rangle|^{1/2},$$

carries with it some of the features of all of the above terms. We want to move X to the right, use the weighted Schwarz inequality, and estimate the derivatives $\tilde{\gamma}_j^{(\ell)}$ just as we have done before. But two things may happen: in first commuting X to the left another derivative may fall on $\tilde{\gamma}_j^{(\ell)}$ (doing no harm - the estimates on these derivatives are flexible enough to handle one or two more derivatives by changing the constant a bit, uniformly in r). But the coefficient $x_1^a \varphi^{(b)}$ may also be differentiated by X . No matter - this has happened often before, and either φ receives one more derivative gains a coefficient of x_1^{p-1} , as in I_5 , or x_1^a becomes x_1^{a-1} as in I_4 .

Putting these results together, the error terms, apart from those which may be absorbed on the left, we have arrived at

Lemma 9.3. *For any a, b, c , and r we have the estimate*

$$\|X x_1^a \varphi^{(b)} D_3^{r-c} u\| + \|x_1^a \varphi^{(b)} D_3^{r-c} u\|_{1/q} + \|x_1^a \varphi^{(b)} X D_3^{r-c} u\| \lesssim \\ (9.25) \quad \|x_1^a \varphi^{(b)} D_3^{r-c} P u\| + \sup_{1 \leq \ell \leq r-c} (C_\gamma r)^\ell \|x_1^a \varphi^{(b)} X D_3^{r-c-\ell} u\| \\ + \|x_1^{a-1} \varphi^{(b)} D_3^{r-c} u\| + \|x_1^{a+p-1} \varphi^{(b+1)} D_3^{r-c} u\| = J_1 + J_2 + J_3 + J_4.$$

And these terms are of four distinct types: the first involves Pu and is harmless; the second exhibits a gain of ℓ powers of D_3 at the expense

of ℓ powers of r ; iteration will lead to $(Cr)^r \sim \tilde{C}^r r!$, which by itself would lead to analytic growth.

For the final two terms, J_3 and J_4 , we argue as follows:

- 1) In treating terms where a power of x has been differentiated, we invoke subellipticity, writing

$$\|x_1^{a-1} \varphi^{(b)} D_3^{r-c} u\| = \|x_1^{a-1} \varphi^{(b)} D_3^{r-c-1/q} u\|_{1/q} + E$$

and estimate E using the standard calculus of pseudo-differential operators - giving rise to a sum of terms, in which a typical term has k more derivatives on φ and k fewer powers of D_3 , modulo an error with no derivatives on u . This trade-off, D_3 's being transferred from u to φ , is the sort that would lead to analyticity. At any rate, the principal contribution is similar to the second term on the left of Lemma 9.3 with a decreased by one and c increased by $1/q$.

- 2) We observe that when $a = 0$ (at the outset, for instance) this kind of term does not arise; thus J_4 will be the first term to arise, starting from $\|\varphi D_3^r u\| : \|\varphi D_3^r u\| \rightarrow \|x_1^{p-1} \varphi^{(1)} D_3^{r-1/q} u\|_{1/q}$.
- 3) Alternatively, when, as in 2) just above, one does add $p-1$ powers of x_1 , add a full derivative to φ , one *may* reach a total of $q-1$ x 's, in which case one invokes Lemma 9.1, and writes $x^{q-1} D_3 = \sum b_j X_j$ and does *not* employ the $1/q$ - 'shunt' in the first item just above. If the powers of x do not permit this, we use the subellipticity again.
- 4) All together, then, we observe that after s steps of type 1) and k steps of type 3), starting from $a = b = c = 0$, we will have, as 'worst' errors,

$$C^{k+s} r^k \|x_1^{k(p-1)-s} \varphi^{(k)} D_3^{r-\ell-\frac{k+s-1}{q}} u\|,$$

where after the last step we have *not* taken $D_3^{1/q}$ and moved it to be part of the norm; for this time, assuming that we have approximately $q-1$ powers of x_1 , we will use Lemma 9.1 to 'create' an X .

Whenever possible (when the powers of x_1 grow to $q-1$, we do not take advantage of the subelliptic $1/q$ gain but combine x^{q-1} with D_3 to produce an X instead. This may happen t times. The result is that

after $s + k + t$ iterations we have an expression

$$(9.26) \quad r^\ell \|\varphi^{(k)} x_1^{k(p-1)-s-t(q-1)} D_3^{r-\ell-\frac{k+s-t}{q}-t} u\|.$$

Now $|\varphi^{(k)}| \leq C^{k+1} r^k$, so that (9.26) is bounded by

$$(9.27) \quad C^{k+1} r^{\ell+k} \|x_1^{k(p-1)-s-t(q-1)} D_3^{r-\ell-\frac{k+s-t}{q}-t} u\|_{L^2(\text{supp}\varphi)}.$$

Since we are looking for powers of x_1 as close to zero as possible (where we started) to gauge the effect of returning to the starting point, it is natural to take

$$t = \frac{k(p-1) - s}{q-1},$$

or its integer part. This choice of t reduces the quantity in (9.27) to the following

$$(9.28) \quad C^{k+1} r^{\ell+k} \|D_3^{r-(\ell+k\frac{p}{q})} u\|.$$

Upon iteration we get

$$(9.29) \quad C^{\sum_j k_j} C^r r^{\sum_j \ell_j + \sum_j k_j} \|D_3^{r-\sum_j \ell_j - \frac{p}{q} \sum_j k_j} u\|,$$

where

$$r - \sum_j \ell_j - \frac{p}{q} \sum_j k_j \sim 0.$$

Let us write $K = \sum_j k_j$ and $L = \sum_j \ell_j$. Then

$$K + L = r \frac{K + L}{r} \sim r \frac{K + L}{L + \frac{p}{q} K} \leq r \frac{q}{p},$$

since $q \geq p$. This ends the proof of the following

Theorem 9.1. *The operator*

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

where the X_j are given in (9.1), is Gevrey hypoelliptic of order $\frac{q}{p}$, i.e.

$$Pu = f \in C^\omega \quad \text{implies that microlocally } u \in G^s, \quad s \geq \frac{q}{p}.$$

Remarks The results given are microlocal. To provide a proof in all detail would entail introducing cut-off functions which are local in space x and also in the frequency variables ξ . This can be done, and has been carried out in all detail in [24] and [25] in the analytic case and in [4] in

the Gevrey category. One introduces localizing functions with the local behavior used here and conic localization in the frequency variables, all cut-off near the origin in the dual variables (with analytic error) in the manner detailed in [25]. We omit details here, as they would largely repeat [25] and risk rendering the exposition unreadable.

In the case of the Oleinik-Radkevich model we know from [3] that these results are optimal and that in particular the result is analytic hypoelliptic if and only if $p = q$. We also strongly believe, but have not yet been able to prove, that every ‘threshold’ obtained in this paper is also sharp.

REFERENCES

- [1] M.S. BAOUENDI AND C. GOULAOUIC, *Analyticity for Degenerate Elliptic Equations and Applications*, Proc. Symp. in Pure Math. **23** (1971), pp. 79-84.
- [2] M.S. BAOUENDI AND C. GOULAOUIC, *Non analytic-hypoellipticity for some degenerate operators*, Bull. A.M.S. **78** (1972), pp. 483-486.
- [3] A. BOVE AND D.S. TARTAKOFF, *Optimal Non-Isotropic Gevrey Exponents for Sums of Squares of Vector Fields*, Comm. in P. D. E. **22**(1997), 1263-1282.
- [4] A. BOVE AND D.S. TARTAKOFF, *Propagation of Gevrey regularity for a class of hypoelliptic equations*, Trans. Amer. Math. Soc. **348** (1996) (7), 2533-2575.
- [5] A. BOVE AND D.S. TARTAKOFF, in preparation.
- [6] M. CHRIST, *On Local and Global Analytic and Gevrey Hypoellipticity*, Journées E.D.P. St.-Jean-de-Monts, 1995, exposé IX, 7 pp.
- [7] M. CHRIST, *Examples Pertaining to a Conjecture on Gevrey Class Hypoellipticity*, preprint, 1/97.
- [8] M. CHRIST, *Intermediate Gevrey Exponents Occur*, Comm. in P. D. E., 1997.
- [9] M. CHRIST, *Hypoellipticity: Geometrization and Speculation*, 1998.
- [10] M. CHRIST, *The Szegő Projection Need Not Preserve Global Analyticity*, Annals of Math. **143** (1996), pp. 301-330.
- [11] M. CHRIST AND D. GELLER, *Counterexamples to Analytic Hypoellipticity for Domains of Finite Type*, Annals of Math. **135** (1992), pp. 551-566.
- [12] M. DERRIDJ, *Sur une classe d'opérateurs différentiels hypoelliptiques à coefficients analytiques*, Sem. Goulaouic-Schwartz, 1970-1971, Équations aux dérivées partielles et analyse fonctionnelle, Exp. no. 12, 6 pp., Centre de Math., École Polytechnique, Paris, (1971).
- [13] M. DERRIDJ AND D.S. TARTAKOFF, *Local Analyticity for \square_b and the $\bar{\partial}$ -Neumann Problem at Certain Weakly Pseudo-Convex Points*, Comm. P. D. E. **13**(12) (1988), pp. 1521-1600.
- [14] M. DERRIDJ AND C. ZUILY, *Sur la régularité Gevrey des opérateurs de Hörmander*, J. de Math. Pures et Appl. **52** (1973), pp. 309-336.
- [15] L. EHRENPREIS, *Solutions of some Problems of Division IV*, Amer. J. Math. **82** (1960), pp. 522-588.
- [16] V.V.GRUŠIN, *On a class of elliptic pseudodifferential operators degenerate on a submanifold*, Math. USSR Sbornik **13** (1971), pp. 155-185.

- [17] N. HANGES AND A. HIMONAS, *Singular Solutions for Sums of Squares of Vector Fields*, Comm. in P.D.E. **16**(8,9) (1991), pp. 1503-1511.
- [18] N. HANGES AND A. HIMONAS, *Non-analytic hypoellipticity in the presence of symplecticity*, preprint.
- [19] L. HÖRMANDER, *Hypoelliptic Second Order Differential Equations*, Acta Math. **119** (1967), pp. 147-171.
- [20] L. HÖRMANDER, *Uniqueness Theorems and Wave Front Sets for Solutions of Linear Differential Equations with Analytic Coefficients*, Comm. Pure Appl. Math., **24** (1971), pp. 671-704.
- [21] O. OLEINIK, *On the analyticity of solutions to partial differential equations and systems*, Soc. Math. de France, Astérisque **2-3** (1973), pp. 272-285.
- [22] 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), pp. 221-222.
- [23] D.S. TARTAKOFF, *Local Analytic Hypoellipticity for \square_b on Non-Degenerate Cauchy Riemann Manifolds*, Proc. Nat. Acad. Sci. U.S.A. **75** (1978), pp. 3027-3028.
- [24] D.S. TARTAKOFF, *On the Local Real Analyticity of Solutions to \square_b and the $\bar{\partial}$ -Neumann Problem*, Acta Math. **145** (1980), pp. 117-204.
- [25] 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.
- [26] F. TREVES, *Analytic Hypo-ellipticity of a Class of Pseudo-Differential Operators with Double Characteristics and Application to the $\bar{\partial}$ -Neumann Problem*, Comm. in P.D.E. **3** (6-7) (1978), pp. 475-642.
- [27] F. TREVES, *Hypo-Analytic structures, Local Theory*, Princeton University Press, Princeton, N.J., 1992.
- [28] F. TREVES, *Symplectic Geometry and Analytic Hypoellipticity, Differential equations: La Pietra 1996*, Proc. Sympos. Pure Math., 65, 201-219, Amer. Math. Soc., Providence, RI, 1999.
- [29] F. TREVES, *Introduction to Pseudodifferential and Fourier Integral Operators*, voll. I, II, Plenum Press, New York and London, 1980.

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI BOLOGNA, PIAZZA DI PORTA S. DONATO 5, 40127 BOLOGNA, ITALIA, AND ISTITUTO NAZIONALE DI FISICA NUCLEARE, SEZIONE DI BOLOGNA

E-mail address: bove@dm.unibo.it

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT CHICAGO, M/C 247, 851 S. MORGAN ST., CHICAGO IL 60607

E-mail address: dst@uic.edu