On a class of sums of squares related to Hamiltonians with a non periodic magnetic field

Published Version:
Bove, A., Chinni, G., Mughetti, M. (2023). On a class of sums of squares related to Hamiltonians with a non periodic magnetic field. JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS, 526(2), 1-20 [10.1016/j.jmaa.2023.127303].

Availability:
This version is available at: https://hdl.handle.net/11585/928493 since: 2023-06-04

Published:
DOI: http://doi.org/10.1016/j.jmaa.2023.127303

ON A CLASS OF SUMS OF SQUARES RELATED TO HAMILTONIANS WITH A NON PERIODIC MAGNETIC FIELD

ANTONIO BOVE, GREGORIO CHINNI, AND MARCO MUGHETTI

Abstract. We consider the operator in (1.1) and prove that it is analytic hypoelliptic. This operator is linked to a stationary Schrödinger equation with a magnetic field and an anharmonic type potential. It is also a sum of squares of vector fields exhibiting a symplectic characteristic variety. This aspect is discussed in the introduction.

1. Introduction

The aim of this paper is to study the analytic hypoellipticity of a class of sums of squares operators defined in an open subset of $\mathbb{R}^3 = \mathbb{R}^2_x \times \mathbb{R}_y$. More precisely let P be a sum of squares operator of the form

$$P(x, D) = D_1^2 + (D_2 + A(x)D_y)^2 + Q(x)D_y^2, \tag{1.1}$$

where $Q(x) \geq 0$ and

$$A(x) = p_m(x)q(x_2). \tag{1.2}$$

Here $p_m(x)$ is a polynomial in x such that $p_m(0) = 0$ and $p_m(x)$ is a homogeneous function satisfying, for some θ, $0 < \theta \leq 1$, the homogeneity relation

$$p_m(\lambda x_1, \lambda^\theta x_2) = \lambda^m p_m(x). \tag{1.3}$$

Moreover we assume that q is a real analytic function such that $q(0) \neq 0$.

As for Q we suppose that, in a neighborhood of the origin, for some r, $p \in \mathbb{N}$, $r \leq p$, we have

$$C_2|x|^{2r} \geq Q(x) \geq C_1|x|^{2p}, \tag{1.4}$$

Date: January 26, 2023.

1991 *Mathematics Subject Classification.* 35H10, 35H20 (primary), 35B65, 35A20, 35A27 (secondary).

Key words and phrases. Sums of squares of vector fields; Analytic hypoellipticity; Schrödinger equation with magnetic field in dimension 2.

and moreover

$$Q(x) = \sum_{j=1}^{n} f_j(x)^2, \tag{1.5}$$

where $f_j \in C^\omega(U)$, U a neighborhood of the origin in $\mathbb{R}^2$ and are homogeneous of degree $m_j > 0$, $j = 1, \dots, n$.

Inequality (1.4) in particular implies that Q vanishes only for $x = 0$ in the assigned neighborhood of $x = 0$.

Thus the characteristic variety of P is given by

$$\operatorname{Char}(P) = \{(x, y; \xi, \eta) \mid x = 0, \xi = 0, \eta \neq 0\}. \tag{1.6}$$

Hence $\operatorname{Char}(P)$ is a symplectic submanifold of codimension 4 of $T^*U \setminus \{0\}$.

P is a sum of squares of vector fields with real analytic coefficients in $\mathbb{R}^3$ and we may see that it satisfies Hörmander bracket condition, due to the fact that e.g. p_m is a polynomial in x.

The operator P in (1.1) is the stationary Schrödinger equation for a magnetic field with vector potential $(0, A)$ in a degenerate infinite potential well Q.

We observe that if A is a function of the variable x_2 only, then, by a gauge transformation (or a canonical transformation in the microlocal parlance) we may reduce P to the form

$$D_1^2 + D_2^2 + Q(x) D_y^2,$$

which is easier to treat. Hence we are assuming that the vector potential A depends on both variables x_1, x_2.

We also point out that if $|A| \leq C Q^{\frac{1}{2}}$, or, in more correct terms,

$$\max\{|A(x)|, |\partial_{x_2} A(x)|\} \leq C \sum_{j=1}^{n} |f_j(x)|, \tag{1.7}$$

for $x \in U$, U neighborhood of the origin, we could prove the analytic hypoellipticity of (1.1) by using the a priori L^2 half maximal estimate:

$$\|u\|_\varepsilon^2 + \|D_1 u\|_0^2 + \|(D_2 + A(x) D_y) u\|_0^2 + \sum_{j=1}^{n} \|f_j(x) D_y u\|_0^2 \leq C_0 \Big(\langle Pu, u \rangle + \|u\|_0^2 \Big), \tag{1.8}$$

where ε is a suitably small positive number and $\|u\|_\varepsilon$ denotes the ε–Sobolev norm: $\|\langle D \rangle^\varepsilon u\|_0$, with $\langle \xi \rangle^2 = 1 + |\xi|^2$.

The goal of the present paper is to prove the

Theorem 1.1. *The operator P in* (1.1)*, satisfying assumptions* (1.2)–(1.5)*, is analytic hypoelliptic in a neighborhood of the origin in $\mathbb{R}^3$.*

In Section 2 we sketch the proof of this fact when (1.7) is verified. This proof hinges on the a priori estimate (1.8), but unfortunately we are not able to use (1.8) in general.

To treat the general case we use the FBI transform technique we used in [2] to prove analyticity for a sum of squares.

Remark 1.2. We assume throughout the paper that the function A depends only on x. We point out that when A is also a function of y the result may no longer be true. For instance consider the following example: let r, p, q, a be positive integers such that $1 < r < p < q$, $a \geq 1$, and define

$$P_1(x, y, D_x, D_y) = D_1^2 + (D_2 + y^a x_2^{p-1} D_y)^2 + x_1^{2(r-1)} D_y^2 + x_2^{2(q-1)} D_y^2.$$

We have that $\mathrm{Char}(P_1) = \{x_1 = x_2 = 0, \xi_1 = \xi_2 = 0, \eta \neq 0\}$. Hence $\mathrm{Char}(P_1) = \mathrm{Char}(P)$.

Even though there is no definition of stratification associated to P_1, see e.g. [3], we think that $\mathrm{Char}(P_1)$ is stratified according to

$$\Sigma_{1,\pm,\pm} = \{x = \xi = 0, \pm\eta > 0, \pm y > 0\},$$
$$\Sigma_{2,\pm} = \{x = \xi = 0, y = 0, \pm\eta > 0\}.$$

Moreover the strata $\Sigma_{2,\pm}$ are not symplectic. In this particular case we believe, although a proof at this stage is missing, that P_1 is, in a neighborhood of the origin of $\mathbb{R}^3$, Gevrey s hypoelliptic for $s \geq s_0$, where

$$s_0^{-1} = 1 - \frac{1}{a} \cdot \frac{r-1}{r} \cdot \frac{q-p}{q-1}.$$

On the other hand if we are looking at a point in a startum of the form $\Sigma_{1,\pm,\pm}$, the assumptions of the present paper are satisfied and we get real analytic hypoellipticity.

2. The case when (1.7) holds

Definition 2.1. For any N natural number, denote by $\varphi_N = \varphi_N(y)$ a function in $C_0^\infty(\mathbb{R}^m)$. We say that φ_N is an Ehrenpreis sequence of cutoff functions if there is a positive constant R such that for $|\alpha| \leq RN$ we have, for every N

$$|\partial_y^\alpha \varphi_N(y)| \leq C_\varphi^{|\alpha|+1} N^{|\alpha|},$$

where $C_\varphi > 0$ and independent of N.

We also include the definition of Gevrey class:

Definition 2.2. If $U \subset \mathbb{R}^n$ is an open set we say that the function u belongs to the Gevrey class of order $s \geq 1$, $G^s(U)$, if $u \in C^\infty(U)$ and for every compact set $K \subset U$ there exists a positive constant C_K such that

$$\sup_K |\partial_x^\alpha u(x)| \leq C_K^{|\alpha|+1} \alpha!^s.$$

When $s = 1$ we get the class of real analytic functions, also denoted by $C^\omega(U)$.

Let $N \in \mathbb{N}$ be a large integer and denote by $\varphi_N(y) \in C_0^\infty(\mathbb{R})$ an Ehrenpreis cutoff function with support at the origin.

In (1.8) we replace u with $\varphi_N^{(h)} D_y^\alpha u$:

$$\begin{aligned}(2.1)\quad \|\varphi_N^{(h)} D_y^\alpha u\|_\varepsilon^2 + \|D_1 \varphi_N^{(h)} D_y^\alpha u\|_0^2 + \|(D_2 + A(x) D_y)\varphi_N^{(h)} D_y^\alpha u\|_0^2 \\ + \sum_{j=1}^n \|f_j(x) D_y \varphi_N^{(h)} D_y^\alpha u\|_0^2 \\ \leq C_0 \Big(\langle P \varphi_N^{(h)} D_y^\alpha u, \varphi_N^{(h)} D_y^\alpha u \rangle + \|\varphi_N^{(h)} D_y^\alpha u\|_0^2 \Big),\end{aligned}$$

where h is a non negative integer and want to show that, if $Pu \in C^\omega(U)$, the following estimate holds:

$$\|\!|\varphi_N^{(h)} D_y^\alpha u|\!\| \leq C_\#^{4\alpha+h+1} N^{\alpha+h}, \tag{2.2}$$

where

$$\|\!|u|\!\|^2 = \|u\|_\varepsilon^2 + \|D_1 u\|_0^2 + \|(D_2 + A(x) D_y) u\|_0^2 + \sum_{j=1}^n \|f_j(x) D_y u\|_0^2.$$

The analyticity of u ensues choosing $\alpha = N$, $h = 0$.

To prove (2.2) we argue by induction on α: for $\alpha = 0$ the estimate (2.2) is a consequence of the properties of the cutoff functions. Assume that (2.2) holds for $\alpha' \leq \alpha - 1$ and for every h such that $\alpha' + h \leq N$. Then we have to show that (2.2) holds for $\alpha' = \alpha$ and $\alpha + h \leq N$.

Consider the error term $\|\varphi_N^{(h)} D_y^\alpha u\|_0$ first. To this end, denote by χ a smooth cutoff function such that $\chi(t) = 1$ if $|t| \geq 2$ and $\chi(t) = 0$ if $|t| \leq 1$. It turns out that $\chi(N^{-1} D_y) \in OPS^0_{0,0}$. For the Definition of these classes we refer to [18] and to [4] for an application in a similar context.

We have then, via the transposed Leibniz formula,

$$\begin{aligned}&\|\varphi_N^{(h)} D_y^\alpha u\|_0 \\ &\quad \leq \|(1 - \chi(N^{-1} D_y)) \varphi_N^{(h)} D_y^\alpha u\|_0 + \|\chi(N^{-1} D_y) \varphi_N^{(h)} D_y^\alpha u\|_0\end{aligned}$$

$$\leq \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} \|(1-\chi(N^{-1}D_y))D_y^{\alpha-\beta}(\varphi_N^{(h+\beta)}u)\|_0 \\ + \|\chi(N^{-1}D_y)\varphi_N^{(h)}D_y^\alpha u\|_0.$$

By the Calderón–Vaillancourt theorem (see [18]) we have that

$$\sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} \|(1-\chi(N^{-1}D_y))D_y^{\alpha-\beta}(\varphi_N^{(h+\beta)}u)\|_0 \\ \leq \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} C_\chi^{\alpha-\beta+1} N^{\alpha-\beta} \|\varphi_N^{(h+\beta)}u\|_0 \\ \leq \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} C_\chi^{\alpha-\beta+1} N^{\alpha-\beta} C_\varphi^{h+\beta+1} N^{h+\beta} \|u\|_0 \\ \leq C_1^{\alpha+h+1} N^{\alpha+h} \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} \frac{C_\chi C_\varphi}{C_1} \left(\frac{C_\varphi}{C_\chi}\right)^\beta \leq C_2^{\alpha+h+1} N^{\alpha+h}$$

Further

$$\|\chi(N^{-1}D_y)\varphi_N^{(h)}D_y^\alpha u\|_0 \leq N^{-\varepsilon}\|N^\varepsilon \chi(N^{-1}D_y)\langle D\rangle^{-\varepsilon}\langle D\rangle^{\varepsilon}\varphi_N^{(h)}D_y^\alpha u\|_0.$$

Due to the support of the cutoff χ, we see that

$$\sigma\big(N^\varepsilon \chi(N^{-1}D_y)\langle D\rangle^{-\varepsilon}\big) = N^\varepsilon \chi(N^{-1}\eta)(1+|\xi|^2+\eta^2)^{-\frac{\varepsilon}{2}} \in S^0_{0,0}$$

with the $S^0_{0,0}$-semi-norms uniformly bounded with respect to N; thus from the Calderón-Vaillancourt theorem it follows that

$$\|N^\varepsilon \chi(N^{-1}D_y)\langle D\rangle^{-\varepsilon}\|_{\mathcal{L}(L^2,L^2)} \leq C,$$

C being a positive constant independent on N.

Finally, we obtain

$$\|\chi(N^{-1}D_y)\varphi_N^{(h)}D_y^\alpha u\|_0 \leq CN^{-\varepsilon}\|\langle D\rangle^{\varepsilon}\varphi_N^{(h)}D_y^\alpha u\|_0 \\ \leq CN^{-\varepsilon}\|\varphi_N^{(h)}D_y^\alpha u\|_\varepsilon.$$

Hence we get

$$\|\varphi_N^{(h)}D_y^\alpha u\|_0 \leq C_2 N^{-\varepsilon}\|\varphi_N^{(h)}D_y^\alpha u\|_\varepsilon + C_2^{h+\alpha+1}N^{\alpha+h}. \tag{2.3}$$

The first norm above can be absorbed on the left hand side of (2.1), while the second is essentially the desired bound, modulo an adjustment of the constant.

Next we examine the scalar product

$$\langle P\varphi_N^{(h)} D_y^\alpha u, \varphi_N^{(h)} D_y^\alpha u\rangle = \sum_{j=1}^{2} \langle X_j^2 \varphi_N^{(h)} D_y^\alpha u, \varphi_N^{(h)} D_y^\alpha u\rangle + \sum_{j=1}^{n} \langle Z_j^2 \varphi_N^{(h)} D_y^\alpha u, \varphi_N^{(h)} D_y^\alpha u\rangle,$$

with an obvious meaning of the notation. Since

$$[X_j^2, \varphi_N^{(h)} D_y^\alpha] = 2X_j [X_j, \varphi_N^{(h)} D_y^\alpha] - [X_j, [X_j, \varphi_N^{(h)} D_y^\alpha]],$$

we get

$$\begin{aligned}
&\langle P\varphi_N^{(h)} D_y^\alpha u, \varphi_N^{(h)} D_y^\alpha u\rangle \\
&= \sum_{j=1}^{2} \Big(2\langle X_j [X_j, \varphi_N^{(h)} D_y^\alpha] u, \varphi_N^{(h)} D_y^\alpha u\rangle - \langle [X_j, [X_j, \varphi_N^{(h)} D_y^\alpha]] u, \varphi_N^{(h)} D_y^\alpha u\rangle \Big) \\
&+ \sum_{j=1}^{n} \Big(2\langle Z_j [Z_j, \varphi_N^{(h)} D_y^\alpha] u, \varphi_N^{(h)} D_y^\alpha u\rangle - \langle [Z_j, [Z_j, \varphi_N^{(h)} D_y^\alpha]] u, \varphi_N^{(h)} D_y^\alpha u\rangle \Big) \\
&= \sum_{j=1}^{2} \Big(T_{1X}^{(j)} + T_{2X}^{(j)} \Big) + \sum_{j=1}^{n} \Big(T_{1Z}^{(j)} + T_{2Z}^{(j)} \Big).
\end{aligned}$$

Trivially $T_{1X}^{(1)} = T_{2X}^{(1)} = 0$, hence we may start by examining $T_{1X}^{(2)}$. We have $[A(x)D_y, \varphi_N^{(h)} D_y^\alpha] = A(x)[D_y, \varphi_N^{(h)} D_y^\alpha] = A(x)[D_y, \varphi_N^{(h)}] D_y^\alpha = A(x)\varphi_N^{(h+1)} D_y^\alpha$. Hence

$$|T_{1X}^{(2)}| \le \|A(x)\varphi_N^{(h+1)} D_y^\alpha u\|_0 \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0.$$

Applying the Cauchy Schwartz inequality and (1.7) we have

$$|T_{1X}^{(2)}| \le \varepsilon \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0^2 + C_\varepsilon \sum_{j=1}^{n} \|f_j(x)\varphi_N^{(h+1)} D_y^\alpha u\|_0^2.$$

The first term can be absorbed on the left provided ε is small. The summands are estimated by using the formula

$$\varphi_N^{(h+1)} D_y^\alpha = \sum_{\ell=0}^{\alpha-1} (-1)^\ell D_y \varphi_N^{(h+1+\ell)} D_y^{\alpha-\ell-1} + (-1)^\alpha \varphi_N^{(h+\alpha+1)}. \tag{2.4}$$

As a consequence

$$\begin{aligned}
|T_{1X}^{(2)}| &\le \varepsilon \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0^2 \\
&+ C_\varepsilon \sum_{j=1}^{n} \Big(\sum_{\ell=0}^{\alpha-1} \|Z_j \varphi_N^{(h+1+\ell)} D_y^{\alpha-\ell-1} u\|_0 \Big)^2 + C_\varepsilon C_f^2 \|\varphi_N^{(h+\alpha+1)} u\|_0^2
\end{aligned}$$

$$\leq \varepsilon \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0^2 + C_\varepsilon \Big(\sum_{\ell=1}^{\alpha} |||\varphi_N^{(h+\ell)} D_y^{\alpha-\ell} u||| \Big)^2 + (C_\#^{\alpha+h+1} N^{\alpha+h})^2$$

$$\leq \varepsilon \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0^2 + C_\varepsilon \Big(\sum_{\ell=1}^{\alpha} C_\#^{4\alpha+h-3\ell+1} N^{\alpha+h} \Big)^2 + (C_\#^{\alpha+h+1} N^{\alpha+h})^2$$

$$\leq \varepsilon \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0^2 + (C_\#^{4\alpha+h+1} N^{\alpha+h})^2 \Big[\Big(\sum_{\ell=1}^{\alpha} C_\varepsilon^{\frac{1}{2}} C_\#^{-3\ell} \Big)^2 + C_\#^{-6\alpha} \Big]$$

$$\leq \varepsilon \|X_2 \varphi_N^{(h)} D_y^\alpha u\|_0^2 + \delta (C_\#^{4\alpha+h+1} N^{\alpha+h})^2,$$

where δ is a small constant to be chosen later and $C_\#$ is large enough.

Consider now $T_{2X}^{(2)}$. A computation analogous to the above yields $[X_2, [X_2, \varphi_N^{(h)} D_y^\alpha]] = [X_2, A(x) \varphi_N^{(h+1)} D_y^\alpha] = [D_2, A(x)] \varphi_N^{(h+1)} D_y^\alpha + A(x)^2$ $[D_y, \varphi_N^{(h+1)}] D_y^\alpha = (D_2 A) \varphi_N^{(h+1)} D_y^\alpha + A(x)^2 \varphi_N^{(h+2)} D_y^\alpha$, so that, using (1.7), (2.4),

$$|T_{2X}^{(2)}| \leq \|(D_2 A) \varphi_N^{(h+1)} D_y^\alpha u\|_0 \|\varphi_N^{(h)} D_y^\alpha u\|_0$$

$$+ N^{-1} \|A(x) \varphi_N^{(h+2)} D_y^\alpha u\|_0 \cdot N \|A(x) \varphi_N^{(h)} D_y^\alpha u\|_0$$

$$\leq \sum_{j=1}^{n} \|f_j(x) \varphi_N^{(h+1)} D_y^\alpha u\|_0 \|\varphi_N^{(h)} D_y^\alpha u\|_0$$

$$+ \sum_{j_1=1}^{n} N^{-1} \|f_{j_1} \varphi_N^{(h+2)} D_y^\alpha u\|_0 \sum_{j_2=1}^{n} N \|f_{j_2}(x) \varphi_N^{(h)} D_y^\alpha u\|_0$$

$$\leq \sum_{j=1}^{n} \Big(\sum_{\ell=0}^{\alpha-1} \|Z_j \varphi_N^{(h+1+\ell)} D_y^{\alpha-\ell-1} u\|_0 + \|\varphi_N^{(h+1+\alpha)} u\|_0 \Big) \|\varphi_N^{(h)} D_y^\alpha u\|_0$$

$$+ \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \Big(\sum_{\ell_1=0}^{\alpha-1} N^{-1} \|Z_{j_1} \varphi_N^{(h+2+\ell_1)} D_y^{\alpha-\ell_1-1} u\|_0 + N^{-1} \|\varphi_N^{(h+2+\alpha)} u\|_0 \Big)$$

$$\cdot \Big(\sum_{\ell_2=0}^{\alpha-1} N \|Z_{j_2} \varphi_N^{(h+\ell_2)} D_y^{\alpha-\ell_2-1} u\|_0 + N \|\varphi_N^{(h+\alpha)} u\|_0 \Big).$$

Now the first line is treated literally as above, keeping in mind that the factor $\|\varphi_N^{(h)} D_y^\alpha u\|_0$ is treated exactly as we did with the error term of tha a priori estimate. The second and third line are also treated analogously and we remark that the factors N^{-1}, N balance the power of N in the final estimate. As a consequence we obtain that

$$|T_{2X}^{(2)}| \leq (C_2 N^{-\varepsilon} \|\varphi_N^{(h)} D_y^\alpha u\|_\varepsilon)^2 + \delta (C_\#^{4\alpha+h+1} N^{\alpha+h})^2.$$

Consider now $T^{(j)}_{1Z}$. We have that $[Z_j, \varphi^{(h)}_N D^\alpha_y] = f_j(x)\varphi^{(h+1)}_N D^\alpha_y$. Hence

$$
\begin{aligned}
|T^{(j)}_{1Z}| &\leq \|f_j(x)\varphi^{(h+1)}_N D^\alpha_y u\|_0 \|Z_j \varphi^{(h)}_N D^\alpha_y u\|_0 \\
&\leq \Big(\sum_{\ell=0}^{\alpha-1} \|Z_j \varphi^{(h+1+\ell)}_N D^{\alpha-1-\ell}_y u\|_0 + \|\varphi^{(h+1+\alpha)}_N u\|_0\Big) \|Z_j \varphi^{(h)}_N D^\alpha_y u\|_0,
\end{aligned}
$$

which is handled exactly as for $T^{(2)}_{1X}$.

The last term to estimate is $T^{(j)}_{2Z}$. We have $[Z_j, [Z_j, \varphi^{(h)}_N D^\alpha_y]] = [Z_j, f_j(x)\varphi^{(h+1)}_N D^\alpha_y] = f_j^2(x)\varphi^{(h+2)}_N D^\alpha_y$. This is treated exactly as the term with A^2 in the expression of the double commutator in $T^{(2)}_{2X}$.

If δ is small enough the argument proves (2.2) and hence the real analyticity of u.

This completes the proof of Theorem 1.1 with the hypothesis (1.7).

3. Background on the FBI transform and related pseudodifferential operators

We are going to use a pseudodifferential calculus introduced by Grigis and Sjöstrand in the paper [13]. We recall below the main definitions and properties to make this paper self-consistent and readable. For further details we refer to the paper [13] and to the lecture notes [22].

3.1. **The FBI Transform.** We define the FBI transform of a temperate distribution u as

$$
Tu(x, \lambda) = \int_{\mathbb{R}^n} e^{i\lambda\varphi(x,y)} u(y) dy, \qquad x \in \mathbb{C}^n,
$$

where $\lambda \geq 1$ is a large parameter, φ is a holomorphic function such that $\det \partial_x \partial_y \varphi \neq 0$, $\operatorname{Im} \partial_y^2 \varphi > 0$.

Here ∂_x denotes the complex derivative with respect to the complex variable x.

Example 3.1.1. The classical phase function is $\varphi(x, y) = \frac{i}{2}(x - y)^2$ corresponding to the FBI transform

$$
Tu(x, \lambda) = \int_{\mathbb{R}^n} e^{-\frac{\lambda}{2}(x-y)^2} u(y) dy, \tag{3.1.1}
$$

where $(x - y)^2 = \sum_{i=1}^n (x_i - y_i)^2$, $x \in \mathbb{C}^n$, $y \in \mathbb{R}^n$.

To the phase φ there corresponds a weight function $\Phi(x)$, defined as

$$
\Phi(x) = \sup_{y \in \mathbb{R}^n} - \operatorname{Im} \varphi(x, y), \qquad x \in \mathbb{C}^n.
$$

We may take a slightly different perspective. Let us consider $(x_0, \xi_0) \in \mathbb{C}^{2n}$ and a real valued real analytic function $\Phi(x)$ defined near x_0, such that Φ is strictly plurisubharmonic and

$$\frac{2}{i}\,\partial_x \Phi(x_0) = \xi_0.$$

Denote by $\psi(x, y)$ the holomorphic function defined near $(x_0, \bar{x}_0)$ by

$$\psi(x, \bar{x}) = \Phi(x). \tag{3.1.2}$$

Because of the plurisubharmonicity of Φ, we have

$$\det \partial_x \partial_y \psi \neq 0. \tag{3.1.3}$$

Another consequence often used in the sequel is an estimate of $\operatorname{Re}\psi(x, \bar{y})$: in fact since $\overline{\psi(x, \bar{y})} = \psi(y, \bar{x})$, we have, with a slight abuse of notation,

$$\begin{aligned}
\operatorname{Re}\psi(x, \bar{y}) &= \frac{1}{2}\left(\psi(x, \bar{y}) + \psi(y, \bar{x})\right)\\
&= \frac{1}{2}\left(\psi(x, \bar{x}) + \psi(y, \bar{y})\right) + \frac{1}{2}\left[\partial_y \psi(x, \bar{x})(\bar{y} - \bar{x}) + \partial_y \psi(y, \bar{y})(\bar{x} - \bar{y})\right]\\
&\quad + \frac{1}{4}\left[\partial_y^2 \psi(x, \bar{x})((\bar{y} - \bar{x}))^2 + \partial_y^2 \psi(y, \bar{y})(\bar{x} - \bar{y})^2\right] + \cdots\\
&= \frac{1}{2}\left(\Phi(x) + \Phi(y)\right) - \frac{1}{2}\partial_x \partial_y \psi(x, \bar{x})(y - x)\overline{(y - x)} + \cdots
\end{aligned}$$

which is expressed as

$$\operatorname{Re}\psi(x, \bar{y}) - \frac{1}{2}\left[\Phi(x) + \Phi(y)\right] \sim -|x - y|^2. \tag{3.1.4}$$

Example 3.1.2. If $\varphi(x, y) = \varphi_0(x, y) = \frac{i}{2}(x - y)^2$ we have that $\Phi_0(x) = \sup_{y \in \mathbb{R}^n} -\operatorname{Im}\varphi_0(x, y) = \frac{1}{2}\sup_{y \in \mathbb{R}^n} -\operatorname{Re}(x' - y + ix'')^2 = \frac{|\operatorname{Im} x|^2}{2}$.

3.2. **Pseudodifferential Operators.** Let $\lambda \geq 1$ be a large positive parameter. We write

$$\tilde{D} = \frac{1}{\lambda} D, \qquad D = \frac{1}{i}\partial.$$

Denote by $q(x, \xi, \lambda)$ an analytic classical symbol and by $Q(x, \tilde{D}, \lambda)$ the formal classical pseudodifferential operator associated to q:

$$Q(x, \tilde{D}, \lambda) u(x, \lambda) = \left(\frac{\lambda}{2\pi}\right)^n \iint e^{i\lambda(x - y)\xi} q(x, \xi, \lambda) u(y, \lambda)\, dy\, d\xi.$$

The above expression is formal and we realize it by choosing an integration path in the (complex) variable ξ of the form

$$\xi = \frac{2}{i}\frac{\partial \Phi}{\partial x}(x) + iR(\overline{x - y}),$$

where R is a sufficiently large positive constant and $|x-y| \leq r$, with r a sufficiently small positive constant.

Using "Kuranishi's trick" one may represent formally $Q(x, \tilde{D}, \lambda)$ as (3.2.1)

$$Q(x, \tilde{D}, \lambda) u(x, \lambda) = \left(\frac{\lambda}{2i\pi}\right)^n \iint e^{2\lambda(\psi(x,\theta)-\psi(y,\theta))} \tilde{q}(x, \theta, \lambda) u(y) dy d\theta.$$

We also can see, because of (3.1.4), that $\theta = \bar{y}$ is a good integration path for (3.2.1):

$$(3.2.2) \qquad Q^{\Omega} u(x, \lambda) = \left(\frac{\lambda}{\pi}\right)^n \int_{\Omega} e^{2\lambda\psi(x,\bar{y})} \tilde{q}(x, \bar{y}, \lambda) u(y) e^{-2\lambda\Phi(y)} L(dy),$$

where $L(dy) = (2i)^{-n} dy \wedge d\bar{y}$, Ω is a small neighborhood of $(x_0, \bar{x}_0) \in \mathbb{C}^{2n}$. We remark that $Q^{\Omega} u(x)$ is an holomorphic function of x.

We refer to Appendix A for the details on how to obtain the good contour $\theta = \bar{y}$ from a classical one.

The advantages of such a definition are:

1- if the principal symbol is real, Q^{Ω} is formally self adjoint in $L^2(\Omega, e^{-2\lambda\Phi})$ where $L^2(\Omega, e^{-2\lambda\Phi}) = \{u \mid \|u\|_{\Phi} < +\infty\}$ and
$$\|u\|_{\Phi}^2 = \int_{\Omega} e^{-2\lambda\Phi(x)} |u(x)|^2 L(dx).$$

2- If $\tilde{q}$ is a classical symbol of order zero, Q^{Ω} is uniformly bounded as $\lambda \to +\infty$, from $H_{\Phi}(\Omega)$ into itself.

Here $H_{\Phi}(\Omega)$ is the space of all holomorphic functions $u(x, \lambda)$ such that for every $\varepsilon > 0$ we have

$$|u(x, \lambda)| \leq C e^{\lambda(\Phi(x)+\varepsilon)}, \qquad C > 0, \qquad x \in \Omega$$

with C independent of x and λ.

For future reference we also recall that the identity operator can be realized as

$$(3.2.3) \quad I^{\Omega} u(x, \lambda) = \left(\frac{\lambda}{\pi}\right)^n \int_{\Omega} e^{2\lambda\psi(x,\bar{y})} i(x, \bar{y}, \lambda) e^{-2\lambda\Phi(y)} u(y, \lambda) L(dy),$$

for a suitable analytic classical symbol $i(x, \xi, \lambda)$ (see also the Appendix.) Moreover we have the following estimate (see [13] and [21], Section 12)

$$(3.2.4) \qquad \|I^{\Omega} u - u\|_{\Phi - d^2/C} \leq C' \|u\|_{\Phi + d^2/C},$$

for suitable positive constants C and C'. Here we denoted by

$$(3.2.5) \qquad d(x) = \operatorname{dist}(x, \complement\Omega),$$

the distance of x to the boundary of Ω.

As a consequence of the above definitions one may see that if q_0 is the principal symbol of Q and $\tilde{q}_0$, i_0 are the principal parts of $\tilde{q}$ and i respectively, we have that

$$\tilde{q}_0(x, \bar{x}) = q_0\left(x, \frac{2}{i}\frac{\partial \Phi}{\partial x}(x)\right) i_0(x, \bar{x}).$$

3.3. **Some Pseudodifferential Calculus.** We start with a proposition on the composition of two pseudodifferential operators.

Proposition 3.3.1 ([13], Proposition 1.3)**.** *Let Q_1 and Q_2 be pseudodifferential operators of order zero. Then they can be composed and*

$$Q_1^\Omega \circ Q_2^\Omega = (Q_1 \circ Q_2)^\Omega + R^\Omega,$$

where R^Ω is an error term whose norm is $\mathscr{O}(1)$ as an operator from $H_{\Phi+(1/C)d^2}$ to $H_{\Phi-(1/C)d^2}$.

We shall need also a lower bound for an elliptic operator of order zero.

We use the following notation:

$$\Lambda_\Phi = \left\{ \left(x, \frac{2}{i}\frac{\partial \Phi}{\partial x}(x, \bar{x})\right), \text{ for } (x, \bar{x}) \in \Omega \times \overline{\Omega} \right\}. \tag{3.3.1}$$

Proposition 3.3.2. *Let Q a zero order pseudodifferential operator defined on Ω as above. Assume further that its principal symbol $q_0(x, \xi, \lambda)$ satisfies*

$$|q_{0|_{\Lambda_\Phi \cap \pi^{-1}(\Omega)}}| \geq c_0 > 0.$$

Here π denotes the projection onto the first factor in $\mathbb{C}^n_x \times \mathbb{C}^n_\xi$. Then

$$\|u\|_{\tilde{\Phi}} + \|Q^\Omega u\|_\Phi \geq C\|u\|_\Phi, \tag{3.3.2}$$

where

$$\tilde{\Phi}(x) = \Phi(x) + \frac{1}{C}d^2(x), \tag{3.3.3}$$

and d has been defined in (3.2.5).

Proof. We have

$$Q^\Omega u(x, \lambda) - q_{0|_{\Lambda_\Phi}}(x, \lambda) I^\Omega u(x, \lambda)$$
$$= \left(\frac{\lambda}{\pi}\right)^n \int_\Omega e^{2\lambda\psi(x,\bar{y})} \left[\tilde{q}(x, \bar{y}, \lambda) - q_{0|_{\Lambda_\Phi}}(x, \lambda) i(x, \bar{y}, \lambda)\right] e^{-2\lambda\Phi(y)} u(y) L(dy).$$

The absolute value of the term in square brackets may be estimated by $C(|x - y| + \lambda^{-1})$ (see formula (1.12) in [13].) Then

$$\|Q^\Omega u - q_{0|_{\Lambda_\Phi}} I^\Omega u\|^2_\Phi \leq C\lambda^{-2}\|u\|^2_\Phi$$

$$
\begin{aligned}
+ C \int_{\Omega} \left| \left(\frac{\lambda}{\pi} \right)^n \int_{\Omega} e^{-\lambda\Phi(x)+2\lambda\psi(x,\bar{y})-\lambda\Phi(y)} |x-y| e^{-\lambda\Phi(y)} u(y) L(dy) \right|^2 L(dx) \\
\leq C \left(\frac{\lambda}{\pi} \right)^{2n} \int_{\Omega} \left(\int_{\Omega} e^{-\lambda/C|x-y|^2} |x-y| L(dy) \right) \\
\cdot \left(\int_{\Omega} e^{-\lambda/C|x-y|^2} |x-y| e^{-2\lambda\Phi(y)} |u(y)|^2 L(dy) \right) L(dx) \\
+ C\lambda^{-2} \|u\|_{\Phi}^2 \leq C\lambda^{-1} \|u\|_{\Phi}^2.
\end{aligned}
$$

Using (3.2.4) we may conclude that

$$
\begin{aligned}
\|Q^{\Omega} u\|_{\Phi} &\geq \|q_{0|_{\Lambda_\Phi}} I^{\Omega} u\|_{\Phi} - C\lambda^{-1/2} \|u\|_{\Phi} \\
&\geq \|q_{0|_{\Lambda_\Phi}} u\|_{\Phi} - \|q_{0|_{\Lambda_\Phi}} (I^{\Omega} - 1) u\|_{\Phi} - C\lambda^{-1/2} \|u\|_{\Phi} \\
&\geq c_0 \|u\|_{\Phi} - C\|u\|_{\tilde{\Phi}} - C\lambda^{-1/2} \|u\|_{\Phi}.
\end{aligned}
$$

This proves the assertion. □

3.4. **A subelliptic estimate for sums of squares on the FBI side.** Let $X_1(x,\xi), \dots, X_\nu(x,\xi)$ be classical analytic symbols of the first order defined in an open neighborhood Ω of $(x_0,\xi_0) \in \Lambda_\Phi$. We assume also that the $X_{j|_{\Lambda_\Phi}}$ are real valued, so that we may think of the corresponding pseudodifferential operators as formally self-adjoint in H_Φ. Let

$$
L(x, \tilde{D}) = \sum_{j=1}^{\nu} X_j^2(x, \tilde{D}). \tag{3.4.1}
$$

Arguing as in [13] we see that the Ω-realization of L can be written as

$$
L^{\Omega} = \sum_{j=1}^{\nu} (X_j^{\Omega})^2 + \mathscr{O}(\lambda^2), \tag{3.4.2}
$$

where $\mathscr{O}(\lambda^2)$ is continuous from $H_{\tilde{\Phi}}$ to $H_{\Phi-(1/C)d^2}$ with norm bounded by $C'\lambda^2$ (see Proposition (3.3.1).)

We assume also that there is a commutator of length $\nu(x_0,\xi_0)$ which is elliptic at $(x_0,\xi_0) \in \Lambda_\Phi$ and that it involves the minimal number of factors.

Next we state the a priori estimate on the FBI side. For the proof we refer to the paper [1].

Theorem 3.4.1. *Let L^{Ω} be as in (3.4.2). We write $r = \nu(x_0, \xi_0)$. Let Ω_1 be a neighborhood of (x_0, ξ_0) such that $\Omega_1 \Subset \Omega$, then*

$$\lambda^{\frac{2}{r}} \|u\|_{\Phi}^2 + \sum_{j=1}^{\nu} \|X_j^{\Omega} u\|_{\Phi}^2 \leq C \left(\langle L^{\Omega} u, u \rangle_{\Phi} + \lambda^{\alpha} \|u\|_{\Phi, \Omega \setminus \Omega_1}^2 \right), \tag{3.4.3}$$

for a suitable positive constant C.

From Theorem 3.4.1 we deduce

Corollary 3.4.2. *With the same notation of Theorem 3.4.1 we have*

$$\lambda^{\frac{2}{r}} \|u\|_{\Phi} \leq C \left(\|L^{\Omega} u\|_{\Phi} + \lambda^{\alpha} \|u\|_{\Phi, \Omega \setminus \Omega_1} \right). \tag{3.4.4}$$

Our purpose is to use Corollary 3.4.2 in order to proof the analytic hypoellipticity of a sums of squares operator. The technique was introduced in [2] and we include the proof of the main theorem in order to make this paper self contained.

3.5. **A criterion of analytic hypoellipticity.** We start by considering an operator of the form

$$P(x, D) = \sum_{i,j=1}^{N} X_i(x, D) a_{ij}(x, D) X_j(x, D) + \sum_{j=1}^{N} b_j(x, D) X_j(x, D) + c(x, D), \tag{3.5.1}$$

where $D_j = D_{x_j} = i^{-1} \partial_{x_j}$ and the $a_{ij}(x, \xi)$, $b_j(x, \xi)$, $c(x, \xi)$ are analytic symbols of order zero such that

$$[a_{ij}]_{i,j=1,\dots,N} + [\bar{a}_{ji}]_{i,j=1,\dots,N} \geq c, \tag{3.5.2}$$

where $c > 0$ is a positive constant.

Let $(x_0, \xi_0) \in \operatorname{Char}(P)$. We assume that $\operatorname{Char}(P)$ is a symplectic real analytic manifold.

Denote by U a neighborhood of the point (x_0, ξ_0) in $\mathbb{R}^{2n}$. We assume that there exists a function $r \colon U \to [0, +\infty[$, real analytic in U, such that

1. $r(x_0, \xi_0) = 0$ and $r(x, \xi) > 0$ in $U \setminus \{(x_0, \xi_0)\}$.
2. There exist real analytic functions, $\alpha_{j,k}(x, \xi)$, defined in U, such that

$$\{r(x, \xi), X_j(x, \xi)\} = \sum_{\ell=1}^{N} \alpha_{j,\ell}(x, \xi) X_\ell(x, \xi), \tag{3.5.3}$$

where $j = 1, \dots, N$, $(x, \xi) \in U$.

We recall the definition of s–Gevrey wave front set of a distribution. In particular, for $s = 1$, we obtain the definition of analytic wave front set.

Definition 3.5.1. Let $(x_0, \xi_0) \in U \subset T^*\mathbb{R}^n \setminus 0$. We say that $(x_0, \xi_0) \notin WF_s(u)$ if there exist a neighborhood Ω of $x_0 - i\xi_0 \in \mathbb{C}^n$ and positive constants C_1, C_2 such that

$$|e^{-\lambda\Phi_0(x)} Tu(x, \lambda)| \leq C_1 e^{-\lambda^{1/s}/C_2},$$

for every $x \in \Omega$. Here T denotes the classical FBI transform, i.e. that in (3.1.1).

Then we have the theorem

Theorem 3.5.2 ([2]). *Assume, as above, that there exists a function $r \in C^\omega(U)$, satisfying conditions (1), (2). Then we have that if $(x_0, \xi_0) \notin WF_a(Pu)$, then $(x_0, \xi_0) \notin WF_a(u)$.*

Proof. Let us write $\tilde{D} = \lambda^{-1} D$, where λ denotes a large positive parameter. The operator P in (3.5.1) then becomes

$$(3.5.4) \quad P(x, \tilde{D}; \lambda) = \sum_{i,j=1}^{N} X_i(x, \tilde{D}) a_{ij}(x, \tilde{D}; \lambda) X_j(x, \tilde{D}) + \lambda^{-1} \sum_{j=1}^{N} b_j(x, \tilde{D}; \lambda) X_j(x, \tilde{D}) + \lambda^{-2} c(x, \tilde{D}; \lambda),$$

with condition (3.5.2) still holding.

We now perform an FBI tranformation on P and we still denote by P the resulting pseudodifferential operator. The cotangent bundle $T^*\mathbb{R}^n$ is thus locally transformed into Λ_{Φ_0}, where Φ_0 denotes the weight function corresponding to the FBI transform with the classical phase function φ_0. Note that Λ_{Φ_0} is contained in $\mathbb{C}^{2n}$ and has real dimension $2n$.

The next step consists in moving away from Λ_{Φ_0} and, following Sjöstrand, [20], we use a canonical deformation of Φ_0 for this purpose.

Let $r(x, \xi)$ be the real analytic function whose existence is guaranteed by our assumptions, or rather its FBI transform. Define the deformed weight function Φ_t, where t denotes a small non negative parameter, as the solution to the following Hamilton–Jacobi equation:

$$(3.5.5) \quad \begin{cases} 2\dfrac{\partial \Phi_t(x)}{\partial t} - r\left(x, \dfrac{2}{i}\dfrac{\partial \Phi_t(x)}{\partial x}\right) = 0, \\ \Phi_t(x)\Big|_{t=0} = \Phi_0(x). \end{cases}$$

We have that $\Lambda_{\Phi_t} = \exp itH_r(\Lambda_{\Phi_0})$.

Next we want to deduce a priori estimates for P with the weight function Φ_0 replaced by Φ_t.

First we write (3.5.3) as

$$\{r(x,\xi), X(x,\xi)\} = \alpha(x,\xi)X(x,\xi), \tag{3.5.6}$$

where X denotes a vector whose components are the symbols of the vector fields $X_1(x,\xi),\dots,X_N(x,\xi)$ and α denotes a $N\times N$ matrix with entries being real analytic symbols.

Denote by Y_j^t the symbol $X_j \circ \exp(itH_r)$, or the restriction to Λ_{Φ_t} of the holomorphic extension of X_j, $j = 1,\dots,N$. We have

$$\partial_t Y^t(x,\xi) = i\{r, X\} \circ \exp(itH_r)(x,\xi),$$

for t small enough. We deduce then that

$$\begin{cases} \partial_t Y^t(x,\xi) = i\alpha \circ \exp(itH_r) Y^t(x,\xi), \\ Y^t(x,\xi)\big|_{t=0} = X(x,\xi). \end{cases}$$

From the above equation we obtain that there is a $N\times N$ matrix whose entries are real analytic symbols with a real analytic dependence on the real parameter t, $b_t(x,\xi)$, such that

$$Y^t(x,\xi) = b_t(x,\xi)X(x,\xi), \tag{3.5.7}$$

and that $b_{t=0}(x,\xi) = \mathrm{Id}_N$. In particular b_t is non singular, provided t is small.

Denote by X^t the holomorphic extension of $\operatorname{Re} Y^t$; since X is real on Λ_{Φ_0}, using (3.5.7), we have that

$$X^t(x,\xi) = \beta_t(x,\xi)X(x,\xi), \tag{3.5.8}$$

where $\beta_{t=0}(x,\xi) = \mathrm{Id}_N$. In particular β_t is non singular, provided t is small.

Then we may write

$$\begin{aligned} P(x,\tilde{D}) = \sum_{i,j=1}^{N} X_i^t(x,\tilde{D}) a_{ij}^t(x,\tilde{D};\lambda) X_j^t(x,\tilde{D}) \\ + \lambda^{-1}\sum_{j=1}^{N} b_j^t(x,\tilde{D};\lambda)X_j^t(x,\tilde{D}) + \lambda^{-2}c^t(x,\tilde{D};\lambda), \end{aligned} \tag{3.5.9}$$

where a_{ij}^t, b_j^t, c^t are symbols of order zero with real analytic dependence on t.

It is also obvious from what has been said before that the fields X_j^t, $j = 1, \dots, N$, also satisfy Hörmander condition with the same bracket length, r, as the X_j.

We may thus apply Theorem 3.4.1 and conclude that the following *a priori* estimate holds ([1]):

$$\lambda^{\frac{2}{r}} \|u\|_{\Phi_t,\Omega_1} \leq C \left(\|Pu\|_{\Phi_t,\Omega} + \lambda^\alpha \|u\|_{\Phi_t,\Omega\setminus\Omega_1} \right), \tag{3.5.10}$$

where $\Omega_1 \subset\subset \Omega$, α is a fixed positive integer and P denotes the realization on Ω of the given operator P.

Let us now assume that $(x_0, \xi_0) \notin WF_a(Pu)$. We may choose Ω in such a way that

$$\|Pu\|_{\Phi_0,\Omega} \leq Ce^{-\lambda/C}, \tag{3.5.11}$$

for a suitable positive constant C. From

$$\Phi_t(x) = \Phi_0(x) + \frac{1}{2} \int_0^t r\left(x, \frac{2}{i} \partial_x \Phi_s(x) \right) ds, \tag{3.5.12}$$

using the fact that $r\big|_{\Lambda_{\Phi_0}} \geq 0$, and recalling that $\Lambda_{\Phi_t} = \exp(itH_r)\Lambda_{\Phi_0}$, we deduce that $r\big|_{\Lambda_{\Phi_t}} \geq 0$ so that

$$\Phi_t(x) \geq \Phi_0(x), \qquad x \in \Omega. \tag{3.5.13}$$

Hence, by (3.5.13) and (3.5.11),

$$\|Pu\|_{\Phi_t,\Omega} \leq Ce^{-\lambda/C}, \tag{3.5.14}$$

for a suitable positive constant C.

Let us now estimate the second term in the right hand side of (3.5.10). We point out that

$$r\big|_{\Lambda_{\Phi_0}\cap\Omega\setminus\Omega_1} \geq a > 0.$$

It follows, because of (3.5.12), that

$$\Phi_t(x) \geq \Phi_0(x) + c't, \qquad x \in \Omega \setminus \Omega_1. \tag{3.5.15}$$

Then

$$\begin{aligned} \|u\|^2_{\Phi_t,\Omega\setminus\Omega_1} &= \int_{\Omega\setminus\Omega_1} e^{-2\lambda\Phi_t(x)} |u(x)|^2 L(dx) \\ &\leq \int_{\Omega\setminus\Omega_1} e^{-2\lambda\Phi_0(x)-2\lambda c't} |u(x)|^2 L(dx) \\ &\leq Ce^{-2\lambda c't}\lambda^N \\ &\leq Ce^{-\lambda c''t}, \qquad t > 0. \end{aligned}$$

By (3.5.10) we deduce that $\|u\|_{\Phi_t,\Omega_1} \leq C\exp(-\lambda t/C)$, for a suitable positive constant C. Let now $\Omega_2 \subset\subset \Omega_1$ be a neighborhood of x_0 such that $\Phi_t \leq \Phi_0 + t/(2C)$ in Ω_2. We conclude that

$$\|u\|^2_{\Phi_0,\Omega_2} \leq Ce^{-\lambda t/C}, \qquad t > 0.$$

This proves the theorem. □

4. Proof of the Theorem

In this section we conclude the proof of Theorem 1.1. Let P be as in (1.1) satisfying the assumptions (1.2) through (1.5). In view of Theorem (3.5.2) we have to construct a real analytic function r microlocally defined in a conic neighborhood of the point $(x_0, \xi_0) \in \mathrm{Char}(P)$, satisfying conditions (1), (2).

We start by performing a real analytic change of variables of the form

$$\begin{aligned} x_1' &= x_1 & \xi_1' &= \xi_1 \\ x_2' &= x_2 g(x_2) & \xi_2' &= \frac{1}{q(x_2)}\xi_2 \\ y' &= y & \eta' &= \eta, \end{aligned} \tag{4.1}$$

where g is to be found so that the above change is canonical. Thus

$$1 \equiv \{\xi_2', x_2'\} = \{\frac{1}{q(x_2)}\xi_2, x_2 g(x_2)\} = \frac{1}{q(x_2)}\partial_{x_2}(x_2 g(x_2)).$$

The latter gives the relation

$$g + x_2\frac{\partial g}{\partial x_2} = q.$$

If, in a neighborhood of the origin, $q(x_2) = \sum_{j=0}^{\infty} q_j x_2^j$ from the above equation we get

$$g(x_2) = \sum_{j=0}^{\infty} q_j \frac{x_2^j}{j+1},$$

or

$$x_2 g(x_2) = \int_0^{x_2} q(t)dt.$$

Then the operator P becomes

$$P'(x', D_{x'}, D_{y'}) = D^2_{x_1'} + q^2(x_2(x_2'))(D_{x_2'} + p_m(x')D_{y'})^2 + Q(x')D^2_{y'} \tag{4.2}$$

Without any loss of generality we may suppose that $(x_0, \xi_0) = (0, e_3)$ and define

$$r(x, y; \xi, \eta) = C_0 P(x, \xi, \eta) + \rho(x, y, \xi, \eta), \tag{4.3}$$

where
$$\rho = (\eta - 1)^2 + \left(y\eta + \frac{1}{m+\theta} \langle x, \xi \rangle_\theta \right)^2 ,$$
where $\langle x, \xi \rangle_\theta = x_1\xi_1 + \theta x_2 \xi_2$.

In the new variables (4.1) r becomes
$$r(x', y'; \xi', \eta') = C_0 P(x_1', x_2(x_2'), \xi_1', q(x_2(x_2'))\xi_2', \eta') \\ + \rho(x_1', x_2(x_2'), y', \xi_1', q(x_2(x_2'))\xi_2', \eta'),$$
where
$$\rho(x_1', x_2(x_2'), y', \xi_1', q(x_2(x_2'))\xi_2', \eta') \\ = (\eta' - 1)^2 + \left(y'\eta' + \frac{1}{m+\theta} \left(x_1'\xi_1' + \theta q(x_2(x_2')) x_2(x_2') \xi_2' \right) \right)^2 \\ = (\eta' - 1)^2 + \left(g(x', y', \xi', \eta') \right)^2 .$$
Since P obviously satisfies the reproduction condition, it is enough to verify it for ρ. Furthermore, since $D_{y'}$ commutes with the operator, it is enough to verify the reproduction property for g.

As a consequence of the change of varaibles in (4.1), we have to reproduce the vector fields $\partial_{x_1'}$, $\partial_{x_2'} + p_m(x_1', x_2(x_2'))\partial_{y'}$, $f_j(x_1', x_2(x_2'))\partial_{y'}$, $j = 1, \dots, n$. It is useful to have the notation $p_m'(x') = p_m(x_1', x_2(x_2'))$ and analogously $f_j'(x') = f_j(x_1', x_2(x_2'))$.

We preliminarily observe that
$$\left(x_1' \partial_{x_1'} + \theta q(x_2(x_2')) x_2(x_2') \partial_{x_2'} \right) p_m'(x') = m p_m'(x'),$$
and analogously for the f_j'. Hence
$$\{g, \xi_2' + p_m'\eta'\} = -p_m'\eta' - \frac{\theta}{m+\theta}\xi_2' + \frac{m}{m+\theta} p_m'\eta' \\ = -\frac{\theta}{m+\theta} \left(\xi_2' + p_m'\eta' \right) .$$
$$\{g, \xi_1'\} = -\frac{1}{m+\theta}\xi_1'.$$
$$\left\{ g, f_j'\eta' \right\} = -f_j'\eta' + \frac{m_j}{m+\theta} f_j'\eta' = \left(\frac{m_j}{m+\theta} - 1 \right) f_j'\eta'.$$
We are then allowed to use Theorem 3.5.2 to prove the analyticity of P'. Since the change of variables in (4.1) is real analytic, this implies the real analytic hypoellipticity of P in (1.1). This completes the proof of the theorem.

A. Appendix: Pseudodifferential Operators on the FBI Side

In this appendix we include some known facts about pseudodifferential operators and FBI transform in order to make the text of the paper self contained. Our main references are [21] and [13].

Let us start with an operator in the form (3.2.1)

(A.1)
$$Q(x, \tilde{D}, \lambda)u(x, \lambda) = \left(\frac{\lambda}{2i\pi}\right)^n \iint e^{2\lambda(\psi(x,\theta)-\psi(y,\theta))}\tilde{q}(x, \theta, \lambda)u(y)dyd\theta,$$

where we use Kuranishi's trick.

We recall what we mean by "Kuranishi's trick". Let $\omega(x, y, \theta)$ denote a phase function which we assume to be holomorphic in x, y and θ. In (3.2.1) we have $\omega(x, y, \theta) = \psi(x, \theta) - \psi(y, \theta)$. We further assume that $\omega(x, x, \theta) = 0$ and
$$\frac{\partial^2 \omega}{\partial x \partial \theta}(x, x, \theta) \neq 0,$$
for $(x, y, \theta) \in U \times U \times V \subset \mathbb{C}^{3n}$.

Because of the assumptions we may write
$$\omega(x, y, \theta) = i(x - y)\xi(x, y, \theta),$$
where $\xi(x, y, \theta)$ depends holomorphically on its arguments and the map $\theta \mapsto \xi(x, y, \theta)$ is a holomorphic diffeomorphism possibly in a smaller open subset. Denote by $\theta(x, y, \xi)$ its inverse. Then the integral in (3.2.1) can be written as
$$\left(\frac{\lambda}{2i\pi}\right)^n \iint e^{i\lambda(x-y)\xi}\tilde{q}(x, \theta(x, y, \xi), \lambda)u(y, \lambda) \det\left(\frac{\partial \theta}{\partial \xi}\right) dyd\xi,$$
where $\tilde{q}(x, \theta(x, y, \xi), \lambda) \det\left(\frac{\partial \theta}{\partial \xi}\right)$ is the "three variable realization" of q.

We remark that a good integration path for the standard pseudodifferential operator becomes a good integration path for the non standard one.

Next we want to show that $\theta = \bar{y}$ is a good integration path for the operator with a non standard phase.

To this end we consider the image of a classic good integration path through the map $\xi \mapsto \theta(x, y, \xi)$. From formula
$$2(\psi(x, \theta) - \psi(y, \theta)) = i(x - y)\xi,$$
we deduce that, with obvious abuse of notation,
$$2\frac{\partial \psi}{\partial x}(x, \theta)(x - y) - \frac{\partial^2 \psi}{\partial x^2}(x, \theta)(x - y)^2 + \mathscr{O}((x - y)^3) = i(x - y)\xi.$$

Hence, modulo the scalar product with a factor $x-y$, writing $\theta = \bar{x} + (\theta - \bar{x})$, we have

$$\frac{2}{i}\frac{\partial\psi}{\partial x}(x,\bar{x}) + \frac{2}{i}\frac{\partial^2\psi}{\partial x\partial y}(x,\bar{x})(\theta-\bar{x}) + \mathscr{O}((\theta-\bar{x})^2) - \frac{\partial^2\psi}{\partial x^2}(x,\bar{x})(x-y) \\ + \mathscr{O}((x-y)(\theta-\bar{x})) + \mathscr{O}((x-y)^2) = \xi,$$

or

$$\frac{2}{i}\frac{\partial^2\psi}{\partial x\partial y}(x,\bar{x})(\theta-\bar{x}) + \mathscr{O}((\theta-\bar{x})^2) - \frac{\partial^2\psi}{\partial x^2}(x,\bar{x})(x-y) \\ + \mathscr{O}((x-y)(\theta-\bar{x})) + \mathscr{O}((x-y)^2) = \xi - \frac{2}{i}\frac{\partial\Phi}{\partial x}(x).$$

Now if ξ is on a good integration path we have

$$\left(\frac{2}{i}\frac{\partial^2\Phi}{\partial x\partial\bar{x}}(x) + \mathscr{O}((\theta-\bar{x})) + \mathscr{O}((x-y))\right)(\theta-\bar{x}) \\ = iR(\overline{x-y}) + \frac{\partial^2\Phi}{\partial x^2}(x)(x-y) + \mathscr{O}((x-y)^2).$$

Hence, since the matrix $\frac{2}{i}\frac{\partial^2\Phi}{\partial x\partial\bar{x}}(x)$ is strictly positive, applying the inverse function theorem, we obtain that

$$\theta - \bar{x} = iRM(\overline{x-y}) + M_1(x-y),$$

where $M = M(x,y,\bar{x},\bar{y})$ is a real analytic $n\times n$ matrix in its arguments and the same holds for M_1.

As a consequence we have a representation formula for the image of the classic good integration path

$$\theta - \bar{y} = (iRM + \mathrm{Id})(\overline{x-y}) + M_1(x-y). \tag{A.2}$$

To show that $\theta = \bar{y}$ is a good integration path we are going to prove that the homotopy

$$\theta - \bar{y} = t\big\{(iRM + \mathrm{Id})(\overline{x-y}) + M_1(x-y)\big\}, \quad t\in[0,1], \tag{A.3}$$

has the following properties:

(i) For every t (A.3) is a good integration path.
(ii) Using Stokes theorem (A.3) can be replaced by $\theta = \bar{y}$ modulo exponentially decreasing errors in λ.

(i). First we observe that when $t = 1$ (A.3) is a good path by definition. As a consequence if $\delta > 0$ and $\delta \le t \le 1$ we have that (A.3) is a good path provided $R \ge R(\delta) > 0$ and $|x-y|$ is small enough. On the other hand if $0 \le t \le \delta$, (A.3) is just a small perturbation of $\theta = \bar{y}$, which is a good path because of (3.1.4). This proves the first assertion.

(ii). To start with let us denote the the path in (A.3) by $\Gamma(t)$:

$$(y,\theta) = (y, \bar{y} + tg(y,\bar{y})) = z_t(y,\bar{y}), \quad t \in [0,1]. \tag{A.4}$$

Consider the integral in (A.1)

$$Q(x,\tilde{D},\lambda)u(x,\lambda) = \left(\frac{\lambda}{2i\pi}\right)^n \iint_{\Gamma(t)} e^{2\lambda(\psi(x,\theta)-\psi(y,\theta))}\tilde{q}(x,\theta,\lambda)u(y)dyd\theta.$$

Denote by

$$f_\lambda(x,y,\theta) = e^{2\lambda(\psi(x,\theta)-\psi(y,\theta))}\tilde{q}(x,\theta,\lambda)u(y),$$

so that we have to study the integral

$$\begin{aligned}
&\left(\frac{\lambda}{2i\pi}\right)^n \iint_{\Gamma(t)} f_\lambda(x,y,\theta)dyd\theta \\
&\quad = \left(\frac{\lambda}{2i\pi}\right)^n \iint_{\Omega\times\overline{\Omega}} f_\lambda(x,z_t(y,\bar{y}))\det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) dy\wedge d\bar{y} \\
&\qquad = \left(\frac{\lambda}{\pi}\right)^n \iint_{\Omega\times\overline{\Omega}} f_\lambda(x,z_t(y,\bar{y}))\det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) L(dy)
\end{aligned}$$

where we chose the positive orientation in $\Omega\times\overline{\Omega}$ and $dy\wedge d\bar{y} = \pm(2i)^n L(dy)$ and $L(dy)$ denotes the Lebesgue measure in $\mathbb{R}^{2n}$.

Let us compute

$$I(t) = \partial_t \iint_{\Omega\times\overline{\Omega}} f_\lambda(x,z_t(y,\bar{y}))\det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) dy\wedge d\bar{y}.$$

We have

$$\begin{aligned}
I(t) &= \iint_{\Omega\times\overline{\Omega}} \langle \partial_{(y,\bar{y})} f_\lambda(x,z_t(y,\bar{y})), (0,g(y,\bar{y}))\rangle \det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) dy\wedge d\bar{y} \\
&\quad + \iint_{\Omega\times\overline{\Omega}} f_\lambda(x,z_t(y,\bar{y}))\partial_t\left(\det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right)\right) dy\wedge d\bar{y} \\
&= I_1(t) + I_2(t).
\end{aligned}$$

Let us consider I_2. We have

$$\begin{aligned}
\partial_t \det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) &= \sum_{\ell=1}^n \det\left[e_1 + t\frac{\partial g}{\partial \bar{y}_1}\cdots\frac{\partial g}{\partial \bar{y}_\ell}\cdots e_n + t\frac{\partial g}{\partial \bar{y}_n}\right] \\
&= \sum_{\ell=1}^n\sum_{k=1}^n \frac{\partial g_k}{\partial \bar{y}_\ell}A_{k\ell},
\end{aligned}$$

where $A_{k\ell}$ denotes the cofactor of the (k, ℓ) entry in the matrix and the first line has the expression of the given matrix according to its columns. We now recall the Jacobi identity for the cofactors:

$$\sum_{\ell=1}^{n} \frac{\partial A_{k\ell}}{\partial \bar{y}_\ell} = 0. \tag{A.5}$$

As a consequence we get

$$\partial_t \det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) = \sum_{\ell=1}^{n}\sum_{k=1}^{n} \partial_{\bar{y}_\ell}\left(A_{k\ell} g_k\right).$$

Plugging this into the expression for I_2 we get

$$\begin{aligned} I_2(t) = \iint_{\Omega\times\overline{\Omega}} f_\lambda(x, z_t(y,\bar{y})) \sum_{\ell=1}^{n}\sum_{k=1}^{n} \partial_{\bar{y}_\ell}\left(A_{k\ell} g_k\right) dy \wedge d\bar{y} \\ = \iint_{\Omega\times\overline{\Omega}} f_\lambda(x, z_t(y,\bar{y})) dy \wedge d\bar{\omega}, \end{aligned}$$

where

$$\bar{\omega} = \sum_{\ell=1}^{n} (-1)^{n+\ell-1} \left(\sum_{k=1}^{n} A_{k\ell} g_k\right) d\bar{y}_1 \wedge \cdots \wedge \widehat{d\bar{y}_\ell} \wedge \cdots \wedge d\bar{y}_n.$$

On the other hand

$$\begin{aligned} & d_{\bar{y}}(f_\lambda(x, z_t(y,\bar{y}))dy \wedge \bar{\omega}) \\ &\qquad = f_\lambda(x, z_t(y,\bar{y}))dy \wedge d_{\bar{y}}\bar{\omega} + d_{\bar{y}} f_\lambda(x, z_t(y,\bar{y}))dy \wedge \bar{\omega} \\ &= f_\lambda(x, z_t(y,\bar{y}))dy \wedge d_{\bar{y}}\bar{\omega} + \sum_{j,\ell=1}^{n} (\partial_{\theta_j} f_\lambda)(\delta_{j\ell} + t\partial_{\bar{y}_\ell} g_j) d\bar{y}_\ell \wedge dy \wedge \bar{\omega} \\ &= f_\lambda(x, z_t(y,\bar{y}))dy \wedge d_{\bar{y}}\bar{\omega} + \sum_{\ell=1}^{n} \langle \partial_\theta f_\lambda, e_\ell + t\partial_{\bar{y}_\ell} g\rangle d\bar{y}_\ell \wedge dy \wedge \bar{\omega} \\ &= f_\lambda(x, z_t(y,\bar{y}))dy \wedge d_{\bar{y}}\bar{\omega} + \sum_{\ell=1}^{n}\sum_{k=1}^{n} \langle \partial_\theta f_\lambda, e_\ell + t\partial_{\bar{y}_\ell} g\rangle A_{k\ell} g_k dy \wedge d\bar{y}. \end{aligned}$$

Since, analogously to what we did above,

$$\sum_{\ell=1}^{n} \partial_{\bar{y}_\ell}(\bar{y} + tg)_j A_{k\ell} = \det\big(\partial_{\bar{y}}(\bar{y} + tg)\big)\delta_{kj},$$

we may conclude that

$$d_{\bar{y}}(f_\lambda(x, z_t(y,\bar{y}))dy \wedge \bar{\omega})$$

$$= f_\lambda(x, z_t(y, \bar{y}))dy \wedge d_{\bar{y}}\bar{\omega} + \sum_{k=1}^{n} \partial_{\theta_k} f_\lambda g_k \det\left(\mathrm{Id}_n + t\frac{\partial g}{\partial \bar{y}}\right) dy \wedge d\bar{y}.$$

The last line above gives the integrands of I_1 and I_2. Thus

$$I(t) = \iint_{\Omega\times\overline{\Omega}} d(f_\lambda dy \wedge \bar{\omega}) = \iint_{\Omega\times\partial\overline{\Omega}} f_\lambda dy \wedge \bar{\omega},$$

and tha latter term is exponentially decreasing w.r.t. λ. Thus

$$\int_0^1 I(t)dt = \iint_{\Gamma(1)} f_\lambda dy d\theta - \iint_{\Gamma(0)} f_\lambda dy d\theta = \mathscr{O}(e^{-c\lambda}),$$

for a suitable constant $c > 0$. This proves (ii).

References

[1] P. Albano, A. Bove and G. Chinni, *Minimal Microlocal Gevrey Regularity for "Sums of Squares"*, International Mathematics Research Notices, Vol. 2009, No. **12**, 2275-2302.

[2] P. Albano and A. Bove, *Wave front set of solutions to sums of squares of vector fields*, Mem. Amer. Math. Soc., **221** (2013), no. 1039.

[3] P. Albano, A. Bove and M. Mughetti, *Analytic Hypoellipticity for Sums of Squares and the Treves Conjecture*, J. Funct. Anal. **274** (2018), no. 10, 2725–2753.

[4] A. Bove and M. Mughetti, *Analytic hypoellipticity for sums of squares in the presence of symplectic non Treves strata*, J. Inst. Math. Jussieu **19**(6) (2020), 1877–1888.

[5] A. Bove and M. Mughetti, *Analytic Hypoellipticity for Sums of Squares and the Treves Conjecture, II*, Analysis and PDE **10** (7) (2017), 1613–1635.

[6] A. Bove and M. Mughetti, *Gevrey Regularity for a Class of Sums of Squares of Monomial Vector Fields*, Advances in Math. **373**(2020) 107323, 35pp.

[7] A. Bove and M. Mughetti, *Analytic regularity for solutions to sums of squares: an assessment*, Complex Analysis and its synergies **6**(2020), no. 2. https://doi.org/10.1007/s40627-020-00055-8.

[8] 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.

[9] G. Chinni, *Gevrey regularity for a generalization of the Oleĭnik–Radkeviç operator*, J. Math. Anal. Appl. **415** (2014), 948–962.

[10] G. Chinni and P. D. Cordaro, *On global analytic and Gevrey hypoellipticity on the torus and the Métivier inequality*, Communications in P.D.E. **42**(1) (2017), 121–141.

[11] G. Chinni, *On the sharp Gevrey regularity for a generalization of the Métivier operator*, preprint, 2021.

[12] G. Chinni, *(Semi)-global analytic hypoellipticity for a class of sums of squares which fail to be locally analytic hypoelliptic*, Proc. A.M.S., to appear, DOI: https://doi.org/10.1090/proc/14464.

[13] A. Grigis and J. Sjöstrand, *Front d'onde analytique et sommes de carrés de champs de vecteurs*, Duke Math. J. **52** (1985), no. 1, 35–51.

[14] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171.
[15] L. Hörmander, *Uniqueness Theorems and Wave Front Sets for Solutions of Linear Differential Equations with Analytic Coefficients*, Communications Pure Appl. Math. **24** (1971), 671–704.
[16] L. Hörmander, *The Analysis of Partial Differential Operators, I*, Springer Verlag, 1985.
[17] L. Hörmander, *The Analysis of Partial Differential Operators, III*, Springer Verlag, 1985.
[18] H. Kumano-Go, *Pseudodifferential Operators* , MIT Press, Cambridge, Mass.–London, 1981.
[19] L. Rothschild and E. M. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math. **137** (3 - 4) (1976), 247–320.
[20] J. Sjöstrand, *Analytic wavefront set and operators with multiple characteristics*, Hokkaido Math. J. **12**(1983), 392–433.
[21] J. Sjöstrand, *Singularités analytiques microlocales*, Astérisque **95** (1982).
[22] J. Sjöstrand, *Lectures on resonances*, unpublished lecture notes, 2002, `http://www.math. polytechnique.fr/~sjoestrand/CoursgbgWeb.pdf`.
[23] 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.
[24] F. Treves, *On the analyticity of solutions of sums of squares of vector fields*, Phase space analysis of partial differential equations, Bove, Colombini, Del Santo ed.'s, 315-329, Progr. Nonlinear Differential Equations Appl., **69**, Birkhäuser Boston, Boston, MA, 2006.
[25] F. Treves, *Analytic Partial Differential Equations*, Grundlehren der Mathematischen Wissenschaften, vol. 359, Springer Verlag, 2022.

Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, Bologna Italy
E-mail address: `bove@bo.infn.it`

Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, Bologna Italy
E-mail address: `gregorio.chinni3@unibo.it`

Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, Bologna Italy
E-mail address: `marco.mughetti@unibo.it`