# THE PRESENCE OF SYMPLECTIC STRATA IMPROVES THE GEVREY REGULARITY FOR SUMS OF SQUARES

*By*

PAOLO ALBANO AND ANTONIO BOVE

**Abstract.** We consider a class of operators of the type sum of squares of real analytic vector fields satisfying the Hörmander bracket condition. The Poisson-Treves stratification is associated to the vector fields. We show that if the deepest stratum in the stratification, i.e., the stratum associated to the longest commutators, is symplectic, then the Gevrey regularity of the solution is better than the minimal Gevrey regularity given by the Derridj-Zuily theorem.

## 1 Introduction and statement of the results

We are concerned with the regularity of the solutions of PDEs of the type "sums of squares" of vector fields. While in the $C^\infty$ category this problem was settled by L. Hörmander in [13], since the mid-seventies it has been well known that the problem of the analytic regularity of the solutions is much more involved.

There are examples in three variables showing that, even though the coefficients of the vector fields are real analytic, the solutions of such equations have only Gevrey regularity, i.e., something half way between $C^\infty$ and real analytic.

In 1999, F. Treves [19] conjectured that the analytic regularity of the solutions of sums of squares equations depends on the geometry of the characteristic set of the operator, i.e., the set on which the symbol of the operator vanishes. This conjecture remains open.

It is clear at this point that the study of the Gevrey regularity of solutions of sums of squares is an important step toward a possible proof of the Treves conjecture. The characteristic variety of a differential operator of type "sum of squares" can be decomposed as the disjoint union of real analytic manifolds, which are called strata and form a local stratification. After giving a prescription on how to construct this stratification, Treves conjectured that the operator is analytic hypoelliptic if and only if the strata are symplectic manifolds. The strata are ordered



JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. 134 (2018)
DOI 10.1007/s11854-018-0005-3

using the length of the Poisson bracket of vector fields that define them. In this paper, we examine the case when the last strata, i.e., the strata defined by the longest brackets, are symplectic and show that the regularity can be improved with respect to the minimal obtained in [9] and [10].

Let us start formulating our problem in a more quantitative way. Let $\mathcal{O}$ be an open subset of $\mathbb{R}^n$, and consider $N$ real analytic vector fields

$$X_j(x, D),\ j = 1, \dots, N,$$

satisfying the Hörmander bracket condition; see 1 below for a statement of this condition.

In this paper, we study the Gevrey hypoellipticity of the operator

$$P(x, D) = \sum_{j=1}^{N} X_j(x, D)^2. \tag{1.1}$$

We recall that a function $u$ is of **Gevrey class** $s(\geq 1)$ **in** $\mathcal{O}$ and write $u \in G^s(\mathcal{O})$ if for every compact set $K \subset \mathcal{O}$, there exists a positive constant $C$ such that

$$|\partial^\alpha u(x)| \leq C^{|\alpha|+1}(\alpha!)^s \quad \text{for all } x \in K$$

for every multi-index $\alpha \in \mathbb{N}^n$. In particular, $G^1(\mathcal{O})$ is the set $C^\omega(\mathcal{O})$ of all real analytic functions defined on $\mathcal{O}$.

Furthermore, the operator $P$ is said to be **Gevrey** $s$ **hypoelliptic in** $\mathcal{O}$ if for every $U \subset\subset \mathcal{O}$ and every distribution $u \in \mathcal{D}'(\mathcal{O})$, the fact that $Pu_{|_U} \in G^s(U)$ implies that $u_{|_U} \in G^s(U)$.

A result of Derridj and Zuily (see [9], [10] and [2] for a microlocal version) relates the Gevrey order of hypoellipticity to the Hörmander bracket condition. More precisely, assume that, at any point of $\mathcal{O}$,

1. the Lie algebra generated by $X_1, \dots, X_N$ and their commutators is of dimension $n$;
2. only commutators of length[1] at most $r$ are needed to generate the algebra.

Derridj and Zuily proved that $P$ is Gevrey $r$ hypoelliptic in $\mathcal{O}$. The index $r$ is the order of minimal Gevrey hypoellipticity, i.e., without other assumptions on the vector fields, this is the best index one can hope for.

In order to study the analytic hypoellipticity of a sum of squares operator with real analytic coefficients, Treves (see [19] and [6]) introduced a stratification with analytic submanifolds, which we call the Poisson–Treves stratification, associated to the (real analytic) vector fields $X_1, \dots, X_N$. More precisely,

[1]The **length** of an iterated commutator is the number of vector fields involved in the Lie bracket; for instance, the commutator $[X_i, [X_i, X_k]]$ is of length 3.

- the characteristic set

$$\mathrm{Char}(X_1, \ldots, X_N) = \{(x, \xi) \mid X_1(x, \xi) = \cdots = X_N(x, \xi) = 0\}$$

is decomposed as a disjoint union of connected real analytic submanifolds, the strata;
- the symplectic form has constant rank on each stratum;
- if $\Sigma_{\nu_j,k}$, $k = 1, \ldots, m_j$, is one of the strata (here, the index $k$ counts the number of disjoint strata of this type), all the Poisson brackets of symbols of the vector fields of length $< \nu_j$ vanish on $\Sigma_{\nu_j,k}$, but there is at least one bracket of length $\nu_j$ which is nonzero[2].

We call $\nu_j$ the **depth** of the strata $\Sigma_{\nu_j,k}$. To be definite, we assume that there exist positive integers $1 < \nu_1 < \nu_2 < \ldots < \nu_{p+1}$ and $m_1, \ldots, m_{p+1}$ such that

$$\mathrm{Char}(X_1, \ldots, X_N) = \bigcup_{j=1}^{p+1} \bigcup_{k=1}^{m_j} \Sigma_{\nu_j,k}.$$

Then the $\Sigma_{\nu_{p+1},k}$ are the **deepest strata**.

In the present paper, we show that under a suitable technical assumption, the following (microlocal) rule of thumb holds: if the deepest stratum of the Poisson-Treves stratification is a symplectic submanifold, then the Gevrey regularity given by the Derridj-Zuily theorem can be improved. Analogous phenomena have also been studied in [4], [6] in particular situations.

The motivation of this work is to make a step forward in the general philosophy underlying Treves' conjecture: the more the symplectic is the strata, the closer the Gevrey regularity is to 1. Actually, the presence of symplectic strata increases the size of the region on which the operator behaves "elliptically".

In order to give a precise statement, it is necessary to microlocalize the problem. Denoting by $WF_s(u)$ the Gevrey $s$ ($\geq 1$) wave front set[3] of the distribution $u$ (see also Subsection 2.1 below), we may define the **Gevrey $s$ (micro)-hypoellipticity** of $P$ as follows:

$$\text{if } (x_0, \xi_0) \in \mathrm{Char}(P) \setminus WF_s(Pu), \text{ then } (x_0, \xi_0) \notin WF_s(u).$$

In order to state a general theorem, we need a geometrical assumption. This is formulated in terms of the existence of a certain real analytic function (the Hamiltonian function) on the cotangent bundle.

**(H)** Let $(x_0, \xi_0) \in \Sigma_{\nu_{p+1},k}$ and $U$ be a neighborhood of the point $(x_0, \xi_0)$ in $\mathbb{R}^{2n}$. There exists a real analytic function $r : U \longrightarrow [0, +\infty[$ such that

[2]The **length** of a Poisson bracket is the number of fields forming it.
[3]The analytic wave front set $WF_1(u)$ is also denoted by $WF_a(u)$.

(1) $r(x_0, \xi_0) = 0$;
(2) $r(x, \xi) > 0$ in
$$A_+ = U \setminus \left[ \{(x_0, \xi_0)\} \cup \left( \bigcup_{j=1}^{p} \cup_{\ell=1}^{m_j} \Sigma_{\nu_j, \ell} \right) \cup \left( \bigcup_{\substack{\ell=1 \\ \ell \neq k}}^{m_{p+1}} \Sigma_{\nu_{p+1}, \ell} \right) \right];$$
(3) there exist real analytic functions, $\alpha_{j,k}(x, \xi)$ defined in $U$ such that
$$\{r(x, \xi), X_j(x, \xi)\} = \sum_{h=1}^{N} \alpha_{jh}(x, \xi) X_h(x, \xi), \tag{1.2}$$
where $j = 1, \ldots, N$ and
$$\{r(x, \xi), X_j(x, \xi)\} = \partial_\xi r(x, \xi) \partial_x X_j(x, \xi) - \partial_x r(x, \xi) \partial_\xi X_j(x, \xi)$$
denotes the Poisson bracket of $r$ and $X_j$.

**Remark 1.1.** We point out that the nontrivial part of condition (**H**) is relation (1.2). We stress the fact that even though the above assumption may seem a technical and restrictive assumption, it has been verified in large classes of problems; see [1] for a detailed discussion of its role. As far as the present paper is concerned, (1.2) allows us to deduce a priori microlocal weighted estimates on complex domains for the operator $P$ using the FBI transform and the FBI characterization of the analytic and Gevrey wave front sets; again see [1].

**Remark 1.2.** On the one hand, the fact that the deepest stratum is symplectic is not explicitly required in the above assumption. On the other hand, we expect that the existence of a function $r$ that vanishes at $(x_0, \xi_0)$ and that is positive on $A_+$ implies that $\Sigma_{\nu_{p+1},k}$ is a symplectic stratum; see [1, Appendix B].

**Theorem 1.1.** *Let $P$ be as in* (1.1)*, where the vector fields have real analytic coefficients and satisfy Hörmander's bracket condition, and assume* (**H**)*. If* $(x_0, \xi_0) \in \Sigma_{\nu_{p+1},k} \setminus WF_{\nu_p}(Pu)$*, for some* $k = 1, \ldots, m_{p+1}$*, then* $(x_0, \xi_0) \notin WF_{\nu_p}(u)$*.*

In other words, if we take a point in a deepest stratum (of depth $\nu_{p+1}$) and assume that such a stratum is symplectic, then we have a Gevrey index of (micro-)hypoellipticity equal to the depth of the "above" stratum (of depth $\nu_p$.) Hence, the index of Gevrey hypoellipticity is improved by $\nu_{p+1} - \nu_p$ units with respect to the minimal Gevrey regularity; see [2].

**Remark 1.3.** Our method applies also to the seemingly more general operator
$$P(x, D) = \sum_{i,j=1}^{N} X_i(x, D) a_{ij}(x, D) X_j(x, D) + \sum_{j=1}^{N} b_j(x, D) X_j(x, D) + c(x, D),$$

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 0 such that

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

for some positive constant $c$.

We now discuss some classes of operators satisfying the above assumptions. Let $\mathcal{O}$ be an open neighborhood of the origin in $\mathbb{R}^{n+1}$ with the variables $(x,t)$, $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$, and consider the operator

$$P = \sum_{j=1}^{n} D_j^2 + \sum_{\ell=1}^{m} \left(b_\ell x^{\alpha^{(\ell)}-1} D_t\right)^2, \quad (x,t) \in \mathcal{O}, \tag{1.3}$$

where $m$ is a positive integer and

- $b_\ell$, $\ell = 1, \dots, m$, are nonzero real numbers;
- $x^{\alpha^{(\ell)}-1} := x_1^{\alpha_1^{(\ell)}-1} \cdots x_n^{\alpha_n^{(\ell)}-1}$, $\ell = 1, \dots, m$, and $\alpha_j^{(\ell)} \geq 2$ for every $\ell = 1, \dots, m$, $j = 1, \dots, n$.

In particular, an operator of the form $P = D_x^2 + D_y^2 + x^{2(p-1)} y^{2(q-1)} D_t^2$, with $p, q \geq 2$ integers, is of the form (1.3).

The 2-dimensional Schrödinger operator associated to $P$, $D_x^2 + D_y^2 + x^{2(p-1)} y^{2(q-1)}$, has been studied by Simon in [15] and [16]. The peculiarity of such an operator is that it has discrete spectrum even if the potential $x^{2(p-1)} y^{2(q-1)}$ vanishes at $xy = 0$, i.e., the operator is not globally elliptic.

**Theorem 1.2.** *Let $P$ be as in* (1.3).

(i) *Let $k \in \{1, \dots, n-1\}$, $i_1, \dots, i_k \in \{1, \dots, n\}$ and let $j_1, \dots, j_{n-k}$ be such that $\{i_1, \dots, i_k, j_1, \dots, j_{n-k}\} = \{1, \dots, n\}$. Then a stratum is given by the formula*

$$\Sigma_{i_1,\dots,i_k,\varepsilon} = \{(x,\xi) \mid \xi_1 = \cdots = \xi_n = 0,\ x_{i_1} = \cdots = x_{i_k} = 0, x_{j_1} x_{j_2} \cdots x_{j_{n-k}} \neq 0\},$$

*where $\varepsilon$ is an ordered string of $n-k+1$ "+" or "−" signs* [4]. *Furthermore, if $k < n$ and $(x,\xi) \in \Sigma_{i_1,\dots,i_k,\varepsilon} \setminus WF_{s_k}(Pu)$, then $(x,\xi) \notin WF_{s_k}(u)$, with*

$$s_k = \min_{\ell=1,\dots,m} \Big( \sum_{r=1}^{k} \alpha_{i_r}^{(\ell)} - k + 1 \Big).$$

*The number $s_k$ is the depth of the stratum $\Sigma_{i_1,\dots,i_k,\varepsilon}$.*

[4] If $k = n-3$ and $\varepsilon = \{+,-,-\}$, then $\Sigma_{i_1,\dots,i_{n-2},\varepsilon} = \{(x,\xi) \mid \xi_1 = \cdots = \xi_n = x_{i_1} = \cdots = x_{i_{n-2}} = 0, +x_{i_{n-1}} > 0,\ -x_{i_n} > 0,\ -\tau > 0\}$

(ii) *If $k = n$, the stratum is*

$$\Sigma_{1,\dots,n,\pm} = \{(x, \xi) \,|x_1 = \cdots = x_n = \xi_1 = \cdots = \xi_n = 0, \ \pm\tau > 0\}.$$

*Furthermore, if $(x, \xi) \in \Sigma_{1,\dots,n,\pm} \setminus WF_{\tilde{s}_n}(Pu)$, then $(x, \xi) \notin WF_{\tilde{s}_n}(u)$, with*

$$\tilde{s}_n = \max_{\{i_1,\dots,i_{n-1}\}\subset\{1,\dots,n\}} \min_{\ell\in\{1,\dots,m\}} \sum_{r=1}^{n-1} \alpha_{i_r}^{(\ell)} - n.$$

**Remarks 1.4.** (i) The number $\tilde{s}_n$ is better (lower) than the depth of the stratum $\Sigma_{1,\dots,n,\pm}$, which is

$$\min_{\ell\in\{1,\dots,m\}} \sum_{r=1}^{n} \alpha_r^{(\ell)} - n + 1.$$

(ii) The formally more general case of the operator

$$P(x, D) = \sum_{j,k=1}^{n} a_{jk}(x, t) D_j D_k + \sum_{\ell=1}^{m} \left(b_\ell(x, t) x^{\alpha^{(\ell)}-1} D_t\right)^2$$

where $m$ is a positive integer,

- $[a_{jk}(x)]_{j,k=1,\dots,n-1}$ is a family of positive definite matrices with real analytic entries;
- $b_\ell(x, t)$, $\ell = 1, \dots, m$ are positive real analytic functions,

can be treated using the same method.

We believe the Gevrey thresholds in Theorem 1.2 to be optimal, based on the following model.

Consider the special case

$$P = D_x^2 + D_y^2 + x^{2(p-1)} y^{2(q-1)} D_t^2,$$

where $p, q \in \mathbb{N}$, $q \geq p > 1$. The operator $P$ is Gevrey $q$ hypoelliptic, by Theorem 1.2. The Gevrey index obtained from the Derridj-Zuily theorem is $p + q - 1 > q$.

**Theorem 1.3.** *For the operator $P$ above, the index $q$ is optimal.*

**Remark 1.5.** The case $p = 1$ in Theorem 1.3 boils down to the generalized Baouendi–Goulaouic operator; in this case, the operator with $p = 1$ is $G^q$ hypoelliptic and not better. Thus it is irrelevant for our purpose.

## 2 Preliminaries

**2.1 The FBI transform.** Define the **Fourier-Bros-Iagolnitzer (FBI) transform** of a tempered distribution $u$ as

$$Tu(x,\lambda) = \int_{\mathbb{R}^n} e^{i\lambda\varphi(x,y)}u(y)dy, \tag{2.1}$$

where $\lambda \geq 1$ is a large parameter and $\varphi$ is a holomorphic function such that $\det \partial_x\partial_y\varphi \neq 0$, $\operatorname{Im} \partial_y^2\varphi > 0$. Here $\partial_x$ denotes the complex derivative with respect to the complex variable $x$.

Unless stated otherwise, we always take $\varphi$ to be the **classical phase function** $\varphi(x,y) = \frac{i}{2}(x-y)^2$.

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

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

We may take a slightly different perspective. Consider $(x_0,\xi_0) \in \mathbb{C}^{2n}$ and a real-valued analytic function $\Phi(x)$ defined near $x_0$ such that $\Phi$ is strictly plurisubharmonic and $\frac{2}{i}\,\partial_x\Phi(x_0) = \xi_0$. Denote by $\psi(x,y)$ the holomorphic function defined near $(x_0,\overline{x_0})$ by

$$\psi(x,\overline{x}) = \Phi(x). \tag{2.2}$$

Because of the plurisubharmonicity of $\Phi$,

$$\det \partial_x\partial_y\psi \neq 0 \tag{2.3}$$

and

$$\operatorname{Re} \psi(x,\overline{y}) - \frac{1}{2}\left[\Phi(x) + \Phi(y)\right] \sim -|x-y|^2. \tag{2.4}$$

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

**Definition 2.1.** 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 $\Omega$ 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 using the classical phase function.

**Theorem 2.1** ([2, Theorem 1.4]). *Let P be as in (1.1) and* $(x_0, \xi_0) \notin WF_r(Pu)$ *be a point in a stratum of depth r. Then* $(x_0, \xi_0) \notin WF_r(u)$.

This is a sort of microlocal Derridj-Zuily theorem. Next we need to define pseudodifferential operators on the FBI side.

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

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

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

Using "Kuranishi's trick" (see, e.g., [14, Proposition .1.3], one may represent $Q(x, \tilde{D}, \lambda)$ as

$$Qu(x, \lambda) = \left(\frac{\lambda}{2i\pi}\right)^n \int e^{2\lambda(\psi(x,\theta)-\psi(y,\theta))}\tilde{q}(x, \theta, \lambda)u(y)dyd\theta. \tag{2.5}$$

Here $\tilde{q}$ denotes a formal classic analytic symbol defined in a neighborhood of $(x_0, \overline{x_0})$, which we may write as $\Omega \times \overline{\Omega}$.

To realize the above operator, we need a prescription for the integration path. This is accomplished by transforming the classical integration path via the Kuranishi change of variables and an application of Stokes' theorem. We obtain

$$Q^\Omega u(x, \lambda) = \left(\frac{\lambda}{\pi}\right)^n \int_\Omega e^{2\lambda\psi(x,\bar{y})}\tilde{q}(x, \bar{y}, \lambda)u(y)e^{-2\lambda\Phi(y)}L(dy), \tag{2.6}$$

where $L(dy) = (2i)^{-n}dy \wedge d\bar{y}$, the integration path is $\theta = \bar{y}$, and $\Omega$ is a small neighborhood of $x_0$. We remark that $Q^\Omega u(x)$ is a holomorphic function of $x$.

The above definition has some useful consequences.

1- If the principal symbol is real, $Q^\Omega$ is formally self-adjoint in $L^2(\Omega, e^{-2\lambda\Phi})$.

2- If $\tilde{q}$ is a classical symbol of order 0, $Q^\Omega$ 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 $\varepsilon > 0$ and for every compact $K \subset\subset \Omega$, there exists a constant $C > 0$ such that $|u(x, \lambda)| \leq Ce^{\lambda(\Phi(x)+\varepsilon)}$, for $x \in K$ and $\lambda \geq 1$.

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

$$I^\Omega u(x, \lambda) = \left(\frac{\lambda}{\pi}\right)^n \int_\Omega e^{2\lambda\psi(x,\bar{y})}i(x, \bar{y}, \lambda)e^{-2\lambda\Phi(y)}u(y, \lambda)L(dy) \tag{2.7}$$

for a suitable analytic classical symbol $i(x, \xi, \lambda)$. Moreover, we have the estimate (see [12] and [17])

$$\|I^\Omega u - u\|_{\Phi-d^2/C} \leq C'\|u\|_{\Phi+d^2/C} \tag{2.8}$$

for suitable positive constants $C$ and $C'$. Here

$$d(x) = \operatorname{dist}(x, \complement\Omega) \tag{2.9}$$

is the distance of $x$ from the boundary of $\Omega$, and

$$\|u\|_\Phi^2 = \int_\Omega e^{-2\lambda\Phi(x)}|u(x)|^2 L(dx). \tag{2.10}$$

Now we state an a priori estimate for an operator of the type "sum of squares" on the FBI side. The estimate is optimal as far as the subellipticity index (or the Gevrey regularity) is concerned. We refer to [2] and [1] for the details.

Let $X_1(x,\xi), \dots, X_N(x,\xi)$ be classical analytic symbols of the first order defined in an open neighborhood $\Omega$ 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 $H_\Phi$. Let

$$\begin{aligned} P(x,\tilde{D}) = \sum_{i,j=1}^{N} X_i(x,\tilde{D}) a_{ij}(x,\tilde{D},\lambda) X_j(x,\tilde{D}) \\ + \lambda^{-1} \sum_{j=1}^{N} b_j(x,\tilde{D},\lambda) X_j(x,\tilde{D}) + \lambda^{-2} c(x,\tilde{D},\lambda), \end{aligned} \tag{2.11}$$

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

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

where $c > 0$ is a positive constant. Let $P^\Omega$ be the $\Omega$-realization of $P$ (in the sense of [12]).

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

**Theorem 2.2.** *Let* $\Omega_1 \subset\subset \Omega$. *Then*

$$\lambda^{\frac{2}{\nu}} \|u\|_\Phi \leq C \left( \|P^\Omega u\|_\Phi + \lambda^\alpha \|u\|_{\Phi,\Omega\setminus\Omega_1} \right), \tag{2.12}$$

*where* $\alpha$ *is a positive integer.*

## 3 Proof of Theorem 1.1

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

$$P(x,\tilde{D}) = \sum_{j=1}^{N} X_j(x,\tilde{D})^2. \tag{3.1}$$

We now perform an FBI tranformation on $P$ and still denote by $P$ the resulting pseudodifferential operator. The cotangent bundle $T^*\mathbb{R}^n$ is thus locally transformed into $\Lambda_{\Phi_0}$, where $\Phi_0$ denotes the weight function corresponding to the FBI transform phase function $\varphi_0$. Note that $\Lambda_{\Phi_0}$ is contained in $\mathbb{C}^{2n}$ and has real dimension $2n$.

The next step consists of canonically moving away from $\Lambda_{\Phi_0}$. Following Sjöstrand [18], we use a canonical deformation of $\Phi_0$ 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 $\Phi_t$, where $t$ denotes a small nonnegative parameter, as the solution to the following Hamilton–Jacobi equation:

$$\begin{cases} 2\dfrac{\partial \Phi_t(x)}{\partial t} - r\left(x, \dfrac{2}{i}\dfrac{\partial \Phi_t(x)}{\partial x}\right) = 0, \\ \Phi_t(x)\big|_{t=0} = \Phi_0(x). \end{cases} \tag{3.2}$$

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

Next we want to deduce a priori estimates for $P$ with the weight function $\Phi_0$ replaced by $\Phi_t$. First we write (1.2) as

$$\{r(x,\xi), X(x,\xi)\} = \alpha(x,\xi)X(x,\xi), \tag{3.3}$$

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

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

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

for $t$ small enough. We now deduce 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

$$Y^t(x,\xi) = b_t(x,\xi)X(x,\xi) \tag{3.4}$$

and $b_0(x,\xi) = \mathrm{Id}_N$. In particular, $b_t$ is nonsingular, provided $t$ is small.

Denote by $X^t$ the holomorphic extension of Re $Y^t$; since $X$ is real on $\Lambda_{\Phi_0}$, using (3.4), we have

$$X^t(x, \xi) = \beta_t(x, \xi) X(x, \xi), \tag{3.5}$$

where $\beta_0(x, \xi) = \mathrm{Id}_N$. In particular, $\beta_t$ is nonsingular, provided $t$ is small.

Then we may write

$$\begin{aligned} P(x, \tilde{D}) = &\sum_{i,j=1}^{N} X_i^t(x, \tilde{D}) a_{ij}^t(x, \tilde{D}; \lambda) X_j^t(x, \tilde{D}) \\ &+ \lambda^{-1} \sum_{j=1}^{N} b_j^t(x, \tilde{D}; \lambda) X_j^t(x, \tilde{D}) + \lambda^{-2} c^t(x, \tilde{D}; \lambda), \end{aligned} \tag{3.6}$$

where $a_{ij}^t$, $b_j^t$, $c^t$ are symbols of order 0 with real analytic dependence on $t$.

It is also obvious from what has been said before that the fields $X_j^t$, $j = 1, \ldots, N$, also satisfy Hörmander condition with the same bracket length $\nu_{p+1}$ as that associated to the fields $X_j$. We may thus apply Theorem 2.2 and obtain the a priori estimate ([2])

$$\lambda^{\frac{2}{\nu_{p+1}}} \|u\|_{\Phi_t, \Omega_1} \leq C \left( \|Pu\|_{\Phi_t, \Omega} + \lambda^{\alpha} \|u\|_{\Phi_t, \Omega \setminus \Omega_1} \right), \tag{3.7}$$

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

Let us now assume that $(x_0, \xi_0) \notin WF_{\nu_p}(Pu)$. We may choose $\Omega$ in such a way that

$$\|Pu\|_{\Phi_0, \Omega} \leq C e^{-\lambda^{1/\nu_p}/C} \tag{3.8}$$

for a suitable positive constant $C$.

Furthermore, taking $\Omega$ small enough and still denoting by $\Sigma_{\nu_j, i}$ the image of $\Sigma_{\nu_j, i}$ under the complex canonical transformation associated to the FBI transform, we may assume that for every $\rho \in \partial\Omega \cap \Sigma_{\nu_j, i}$, $j < p+1$, there exists a neighborhood $W$ of $\rho$ in $\mathbb{C}^{2n}$ such that

$$\|u\|_{\Phi_0, W} \leq C e^{-\lambda^{\frac{1}{\nu_j}}/C}. \tag{3.9}$$

Here, we have used Theorem 2.1. There is no loss of generality in assuming that $\partial\Omega \cap \Sigma_{\nu_{p+1}, i} = \varnothing$, $i \neq k$.

By compactness, there are finitely many open sets $W$, say $W_1, \ldots, W_\ell$, where an estimate of the kind (3.9) is satisfied. From

$$\Phi_t(x) = \Phi_0(x) + \frac{1}{2} \int_0^t r\left(x, \frac{2}{i} \partial_x \Phi_s(x)\right) ds, \tag{3.10}$$

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

$$\Phi_t(x) \geq \Phi_0(x), \quad x \in \Omega. \tag{3.11}$$

Hence, by (3.11), and (3.8),

$$\|Pu\|_{\Phi_t,\Omega} \leq Ce^{-\lambda^{1/\nu_p}/C} \tag{3.12}$$

for a suitable positive constant $C$.

Let us now estimate the second term in the right hand side of (3.7). Due to assumption (**H**), for a suitable positive number $a$,

$$r_{|\Lambda_{\Phi_0}\cap\Omega\setminus(\Omega_1\cup W_1\cup\cdots\cup W_\ell)} \geq a > 0.$$

It follows, because of (3.10), that

$$\Phi_t(x) \geq \Phi_0(x) + c't, \quad x \in \Omega \setminus (\Omega_1 \cup W_1 \cdots \cup W_\ell). \tag{3.13}$$

Then

$$\begin{aligned}
\|u\|^2_{\Phi_t,\Omega\setminus\Omega_1} &= \int_{\Omega\setminus(\Omega_1\cup W_1\cup\cdots\cup W_\ell)} e^{-2\lambda\Phi_t(x)}|u(x)|^2 L(dx) \\
&\quad + \int_{W_1\cup\cdots\cup W_\ell} e^{-2\lambda\Phi_t(x)}|u(x)|^2 L(dx) \\
&\leq \int_{\Omega\setminus(\Omega_1\cup W_1\cup\cdots\cup W_\ell)} e^{-2\lambda\Phi_0(x)-2\lambda c't}|u(x)|^2 L(dx) \\
&\quad + C \max_{j\in\{1,\dots,p\}} e^{-\lambda^{\frac{1}{\nu_j}}/C} \\
&\leq C\left(e^{-2\lambda c't}\lambda^N + e^{-\lambda^{\frac{1}{\nu_p}}/C}\right) \leq Ce^{-\lambda^{\frac{1}{\nu_p}}c''t}, \quad t > 0.
\end{aligned}$$

By (3.7), we deduce that $\|u\|_{\Phi_t,\Omega_1} \leq C\exp(-\lambda^{\frac{1}{\nu_p}}t/C)$ 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 $\Omega_2$. We conclude that

$$\|u\|^2_{\Phi_0,\Omega_2} \leq Ce^{-\lambda^{\frac{1}{\nu_p}}t/C}, \quad t > 0.$$

In other words, $(x_0, \xi_0) \notin WF_{\nu_p}(u)$. This completes our proof.

## 4 Proof of Theorem 1.2

Part (i) of the theorem is the minimal Gevrey regularity proved in [2]; see Theorem 2.1.

To prove part (ii), it suffices to show that there exists a function $r$ satisfying conditions (1), (2) and (1.2) of (**H**). We may assume, without loss of generality, that

$$(x_0, t_0, \xi_0, \tau_0) = (0, 0, 0, 1) \in \mathbb{R}^{2(n+1)}.$$

Let $U$ be a neighborhood of $(0, 0, 0, 1)$ in $\mathbb{R}^{2(n+1)}$. We claim that the function

$$r(x, t, \xi, \tau) = \sum_{j=1}^{n} \xi_j^2 + \sum_{\ell=1}^{m} \left(b_\ell x^{\alpha^{(\ell)}-1}\tau\right)^2 + t^2\tau^2 + (\tau-1)^2$$

satisfies all the requirements.

Establishing condition (1) and (1.2) is straightforward. Condition (2) is verified if $r$ is positive in the intersection of the deepest stratum, $\{(0, t, 0, \tau) | \pm\tau > 0\}$, with $\partial U$; and this follows by the inequality $r(x, t, \xi, \tau) \geq t^2\tau^2 + (\tau-1)^2$. The conclusion follows by applying Theorem 1.1 to the present situation. This completes our proof.

## 5 Proof of Theorem 1.3

We need the following preliminary lemma.

**Lemma 5.1.** *Let $a(x)$ be a nowhere vanishing real analytic function defined in a neighborhood of* 0*. Then the operator*

$$L_a(x, y, D_x, D_y, D_t) = D_x^2 + D_y^2 + a(x)^2 y^{2(q-1)} D_t^2,$$

*defined in a small neighborhood of the origin in $\mathbb{R}^3$, is $G^q$ Gevrey hypoelliptic and not better.*

An indirect proof has been given in [8] for the special case $q = 2$.

**Proof.** To be definite, we consider an open ball $B_R \subset \mathbb{R}^2$ with center at the origin and radius $R$, and let $A$ be a positive constant such that $\max_{(x,y)\in B_R} a(x) \leq A$. Let $\varphi_0 = \varphi_0(y)$, $\lambda_0$ be the first eigenfunction, respectively eigenvalue, of the operator $-\partial_y^2 + y^{2(q-1)}$, i.e.,

$$-\partial_y^2\varphi_0(y) + y^{2(q-1)}\varphi_0(y) = \lambda_0\varphi_0(y).$$

It is known that $\lambda_0 > 0$ and that $\varphi_0$ is a positive rapidly decreasing function. In particular, $0 \leq \varphi_0(y) \leq C_0$, $y \in \mathbb{R}$, for a suitable positive constant $C_0$.

For a positive parameter $\rho$, we define

$$s_\rho(x, y) = \varphi_0\left(A^{\frac{1}{2q}}\rho y\right) e^{xA^{\frac{1}{2q}}\sqrt{\lambda_0}\rho} e^{-C_1\rho}.$$

Here, $C_1$ is a positive constant such that $RA^{\frac{1}{2q}}\sqrt{\lambda_0} - C_1 < 0$. In particular, there exists $C_2 > 0$ such that

$$e^{xA^{\frac{1}{2q}}\sqrt{\lambda_0}\rho}e^{-C_1\rho} \le e^{-C_2\rho}$$

for $x \in [-R, R]$ and $\rho \ge 1$.

Now let $u_\rho$ be the solution of the Dirichlet problem

$$\begin{cases} (D_x^2 + D_y^2 + a(x)y^{2(q-1)}\rho^{2q})u_\rho(x, y) = 0, & \text{on } B_R, \\ u_\rho(x, y) = s_\rho(x, y), & \text{on } \partial B_R. \end{cases}$$

The existence of the (classical) solution of this Dirichlet problem is a classical result of the theory of elliptic equations; see, e.g., [11, Corollary 6.9]. Using the maximum principle, we deduce that

$$s_\rho(x, y) \le u_\rho(x, y) \le C_0 e^{-C_2\rho}, \quad (x, y) \in B_R,\ \rho \ge 1. \tag{5.1}$$

Now set

$$v(x, y, t) = \int_1^{+\infty} e^{it\rho^q} u_\rho(x, y)\, d\rho.$$

By (5.1), this integral is convergent. A direct computation shows that

$$L_a(x, y, D_x, D_y, D_t)v(x, y, t) = 0.$$

Furthermore, for every positive integer $k$,

$$\varphi_0(0) \int_1^{+\infty} \rho^{kq} e^{-C_1\rho}\, d\rho \le |D_t^k v(0, 0, 0)| \le C_0 \int_1^{+\infty} \rho^{kq} e^{-C_2\rho}\, d\rho,$$

i.e., $|D_t^k v(0, 0, 0)| \sim k!^q$. This shows that the function $v$ has Gevrey regularity not better than Gevrey $q$. □

Next we prove Theorem 1.3. We argue by contradiction, assuming that the operator $P$ is Gevrey $s$ hypoelliptic, with $s \in [1, q[$ in a neighborhood $\mathcal{O}$ of the origin in $\mathbb{R}^3$.

Let $\varepsilon$ denote a parameter so small that $(2\varepsilon, 0, 0) \in \mathcal{O}$, and let $Q \subset\subset \mathcal{O}$ be a open cube of side $2\varepsilon$ centered at $(2\varepsilon, 0, 0)$. viz,

$$Q = ]\varepsilon, 3\varepsilon[ \times ]-\varepsilon, \varepsilon[ \times ]-\varepsilon, \varepsilon[.$$

We now use the above argument on the optimality of the Gevrey $q$ regularity for the operator $L_a$. In particular, we showed that there is a Gevrey $q$ (and not better) solution of

$$Pv = 0 \quad \text{in } Q. \tag{5.2}$$

Let $\chi = \chi(y)$ be a cut-off function of class $G^s$ such that

$$\chi(y) = \begin{cases} 1, & y \le \varepsilon/4, \\ 0, & y \ge \varepsilon/2. \end{cases}$$

Then $P(\chi v) = [P, \chi]v$. Furthermore, defining

$$\tilde{Q} = \mathcal{O} \cap \Big( ]\varepsilon, 3\varepsilon[ \times ]-\varepsilon, \infty[ \times ]-\varepsilon, \varepsilon[ \Big),$$

we have that $\chi v \in G^q(\tilde{Q}) \setminus G^s(\tilde{Q})$ (by the definition of $v$). We claim that $[P, \chi]v \in G^s(\tilde{Q})$. Indeed, the commutator $[P, \chi]v$ is supported in

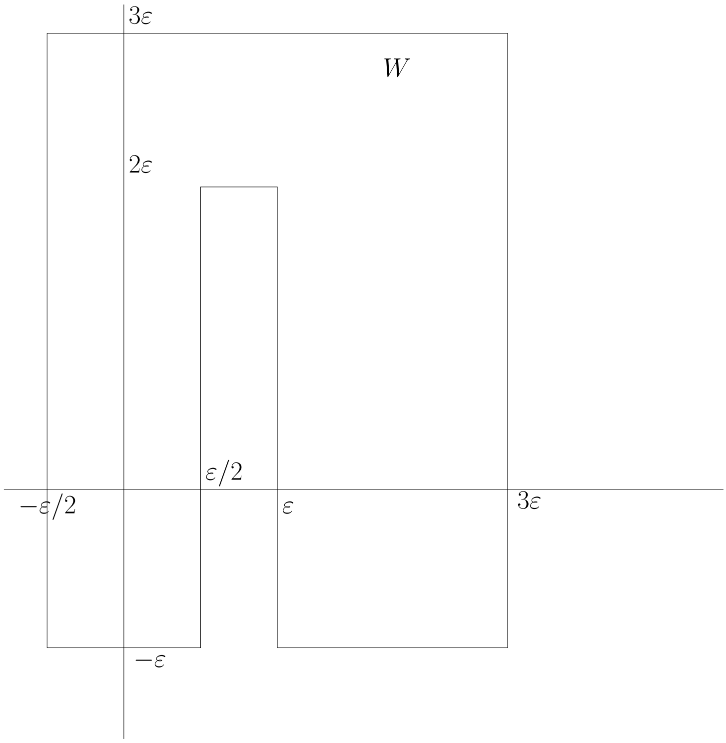
$$Q_0 = \Big( ]\varepsilon, 3\varepsilon[ \times ]\varepsilon/4, \varepsilon/2[ \times ]-\varepsilon, \varepsilon[ \Big) \subset Q.$$

Now the operator $P$ is elliptic in $Q_0$ with real analytic coefficients; hence, in particular, it is $G^s$ hypoelliptic in $Q_0$. Thus $v \in G^s(Q_0)$, and we deduce that $[P, \chi]v \in G^s(Q_0) \cap G^s(\tilde{Q})$. We have found a subset $\tilde{Q} \subset\subset \mathcal{O}$ and two functions $w = \chi v \in G^q(\tilde{Q}) \setminus G^s(\tilde{Q})$, $g = [P, \chi]v \in G^s(\tilde{Q})$ such that $Pw = g$ in $\tilde{Q}$. Define

$$u(x, y, t) = \begin{cases} w(x, y, t), & (x, y, t) \in \tilde{Q}, \\ 0, & (x, y, t) \in \mathcal{O} \setminus \tilde{Q}, \end{cases}$$

and let

$$W = \Big\{ \Big( ]\varepsilon, 3\varepsilon[ \times ]-\varepsilon, 3\varepsilon[ \Big) \cup \Big( ]\frac{\varepsilon}{2}, \varepsilon[ \times ]2\varepsilon, 3\varepsilon[ \Big) \cup \Big( ]-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}[ \times ]-\varepsilon, 3\varepsilon[ \Big) \Big\},$$




$$U = W \times ]-\varepsilon, \varepsilon[.$$

We may assume that $\varepsilon > 0$ is such that $U \subset \mathcal{O}$.

Then we have $u \in \mathcal{D}'(\mathcal{O})$ such that for a suitable neighborhood of the origin, $U \subset\subset \mathcal{O}$, $Pu \in G^s(U)$. On the other hand, $u \in G^q(U) \setminus G^s(U)$ near the origin, which contradicts the assumption that $P$ is Gevrey $s$ hypoelliptic in $\mathcal{O}$.

This completes our proof of the optimality.

**Note added in proofs.** Since the acceptance of this paper, the sufficient part of Treves conjecture has been disproved by Albano, Bove and Mughetti in the papers [3] and [5]. The necessary part is still an open problem. Furthermore, there are no conjectures as to when a real analytic sum of squares is analytic hypoelliptic, not to mention Gevrey hypoellipticity.

## References


[1] 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.

[2] P. Albano, A. Bove, and G. Chinni, *Minimal microlocal Gevrey regularity for "sums of squares"*, Int. Math. Res. Not. IMRN, **2009**, 2275–2302.

[3] P. Albano, A. Bove, and M. Mughetti, *Analytic hypoellipticity for sums of squares and the Treves conjecture*, preprint, 2016, `http://arxiv.org/abs/1605.03801`.

[4] A. Bove, *Gevrey hypo-ellipticity for sums of squares of vector fields: some examples*, in *Geometric Analysis of PDE and Several Complex Variables*, Contemp. Math. **368** (2005), 41–68.

[5] A. Bove and M. Mughetti, *Analytic hypoellipticity for sums of squares and the Treves Conjecture, II*, Analysis and PDE **10** (2017), 1613–1635.

[6] A. Bove and D. S. Tartakoff, *A class of sums of squares with a given Poisson-Treves stratification*, J. Geom. Anal. **13** (2003), 391–420.

[7] 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.

[8] P. Cordaro and N. Hanges, *Hypoellipticity in spaces of ultradistributions–study of a model case*, Israel J. Math. **191** (2012), 771–789.

[9] M. Derridj and C. Zuily, *Régularité analytique et Gevrey d'opérateurs elliptiques dégénérés*, J. Math. Pures Appl. (9) **52** (1973), 65–80.

[10] M. Derridj and C. Zuily, *Sur la régularité Gevrey des opérateurs de Hörmander*, J. Math. Pures Appl. (9) **52** (1973), 309–336.

[11] D. Gilbarg and N. S. Trudinger *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 2001.

[12] A. Grigis and J. Sjöstrand, *Front d'onde analytique et somme de carrés de champs de vecteurs*, Duke Math. J. **52** (1985), 35–51.

[13] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171.

[14] L. Hórmander, *Fourier integral operators. I*, Acta Math. **127** (1971), 79–183.

[15] B. Simon, *Some quantum operators with discrete spectrum but classically continuous spectrum*, Ann. Physics **146** (1983), 209–220.

[16] B. Simon, *Nonclassical eigenvalue asymptotics*, J. Funct. Anal. **53** (1983), 84–98.

[17] J. Sjöstrand, *Singularités analytiques microlocales*, Astérisque **95** (1982).

[18] J. Sjöstrand, *Analytic wavefront set and operators with multiple characteristics*, Hokkaido Math. J. **12** (1983), 392–433.

[19] F. Treves, *Symplectic geometry and analytic hypo-ellipticity*, in *Differential Equations: La Pietra 1996*, Amer. Math. Soc., Providence, RI, 1999, pp. 201–219.



*Paolo Albano and Antonio Bove*
DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI BOLOGNA
PIAZZA DI PORTA SAN DONATO 5
40127 BOLOGNA, ITALY
email: paolo.albano@unibo.it,bove@bo.infn.it


(Received June 9, 2014)