ANALYSIS & PDE

Volume 16 No. 3 2023

CYRIL LETROUIT

SUBELLIPTIC WAVE EQUATIONS ARE NEVER OBSERVABLE

msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 3, pp. 643–678
DOI: 10.2140/apde.2023.16.643

msp

SUBELLIPTIC WAVE EQUATIONS ARE NEVER OBSERVABLE

CYRIL LETROUIT

It is well known that observability (and, by duality, controllability) of the elliptic wave equation, i.e., with a Riemannian Laplacian, in time T_0 is almost equivalent to the geometric control condition (GCC), which stipulates that any geodesic ray meets the control set within time T_0. We show that in the subelliptic setting, the GCC is never satisfied, and that subelliptic wave equations are never observable in finite time. More precisely, given any subelliptic Laplacian $\Delta = -\sum_{i=1}^{m} X_i^* X_i$ on a manifold M, and any measurable subset $\omega \subset M$ such that $M \backslash \omega$ contains in its interior a point q with $[X_i, X_j](q) \notin \mathrm{Span}(X_1, \ldots, X_m)$ for some $1 \leqslant i, j \leqslant m$, we show that, for any $T_0 > 0$, the wave equation with subelliptic Laplacian Δ is not observable on ω in time T_0.

The proof is based on the construction of sequences of solutions of the wave equation concentrating on geodesics (for the associated sub-Riemannian distance) spending a long time in $M \backslash \omega$. As a counterpart, we prove a positive result of observability for the wave equation in the Heisenberg group, where the observation set is a well-chosen part of the phase space.

1. Introduction

1.1. ***Setting.*** Let $n \in \mathbb{N}^*$ and let M be a smooth connected compact manifold of dimension n with a nonempty boundary ∂M. Let μ be a smooth volume on M. We consider $m \geqslant 1$ smooth vector fields $X_1, \ldots, X_m$ on M which are not necessarily independent, and we assume that the following Hörmander condition [1967] holds:

The vector fields $X_1, \ldots, X_m$ and their iterated brackets $[X_i, X_j]$, $[X_i, [X_j, X_k]]$, etc. span the tangent space $T_q M$ at every point $q \in M$.

MSC2010: primary 22E25, 35H20, 35L05, 93B05, 93B07; secondary 35H10, 35S05, 78A05.
Keywords: subelliptic, wave equation, observability, sub-Riemannian.

We consider the sub-Laplacian Δ defined by

$$\Delta = -\sum_{i=1}^{m} X_i^* X_i = \sum_{i=1}^{m} X_i^2 + \mathrm{div}_\mu(X_i) X_i,$$

where the star designates the transpose in $L^2(M, \mu)$ and the divergence with respect to μ is defined by $L_X \mu = (\mathrm{div}_\mu X)\mu$, where L_X stands for the Lie derivative. Then Δ is hypoelliptic; see [Hörmander 1967, Theorem 1.1].

We consider Δ with Dirichlet boundary conditions and the domain $D(\Delta)$ which is the completion in $L^2(M, \mu)$ of the set of all $u \in C_c^\infty(M)$ for the norm $\|(\mathrm{Id}-\Delta)u\|_{L^2}$. We also consider the operator $(-\Delta)^{1/2}$ with domain $D((-\Delta)^{1/2})$ which is the completion in $L^2(M, \mu)$ of the set of all $u \in C_c^\infty(M)$ for the norm $\|(\mathrm{Id} - \Delta)^{1/2} u\|_{L^2}$.

Consider the wave equation

$$\begin{cases} \partial_{tt}^2 u - \Delta u = 0 \quad \text{in } (0, T) \times M, \\ u = 0 \quad \text{on } (0, T) \times \partial M, \\ (u_{|t=0}, \partial_t u_{|t=0}) = (u_0, u_1), \end{cases} \tag{1}$$

where $T > 0$. It is well known (see for example [Garetto and Ruzhansky 2015, Theorem 2.1; Engel and Nagel 2000, Chapter II, Section 6]) that for any $(u_0, u_1) \in D((-\Delta)^{1/2}) \times L^2(M)$, there exists a unique solution

$$u \in C^0(0, T; D((-\Delta)^{\frac{1}{2}})) \cap C^1(0, T; L^2(M)) \tag{2}$$

to (1) (in a mild sense).

We set

$$\|v\|_{\mathcal{H}} = \left(\int_M \sum_{j=1}^{m} (X_j v(x))^2 \, d\mu(x) \right)^{\frac{1}{2}}. \tag{3}$$

Note that $\|v\|_{\mathcal{H}} = \|(-\Delta)^{1/2} v\|_{L^2(M,\mu)}$.

The natural energy of a solution is

$$E(u(t, \cdot)) = \tfrac{1}{2}(\|\partial_t u(t, \cdot)\|^2_{L^2(M,\mu)} + \|u(t, \cdot)\|^2_{\mathcal{H}}).$$

If u is a solution of (1), then

$$\frac{d}{dt} E(u(t, \cdot)) = 0,$$

and therefore the energy of u at any time is equal to

$$\|(u_0, u_1)\|^2_{\mathcal{H}\times L^2} = \|u_0\|^2_{\mathcal{H}} + \|u_1\|^2_{L^2(M,\mu)}.$$

In this paper, we investigate exact observability for the wave equation (1).

Definition 1. Let $T_0 > 0$ and $\omega \subset M$ be a μ-measurable subset. The subelliptic wave equation (1) is exactly observable on ω in time T_0 if there exists a constant $C_{T_0}(\omega) > 0$ such that, for any $(u_0, u_1) \in D((-\Delta)^{1/2}) \times L^2(M)$, the solution u of (1) satisfies

$$\int_0^{T_0} \int_\omega |\partial_t u(t, x)|^2 \, d\mu(x) \, dt \geqslant C_{T_0}(\omega) \|(u_0, u_1)\|^2_{\mathcal{H}\times L^2}. \tag{4}$$

1.2. ***Main result.*** Our main result is the following.

Theorem 2. *Let $T_0 > 0$ and let $\omega \subset M$ be a measurable subset. We assume that there exist $1 \leqslant i, j \leqslant m$ and q in the interior of $M \backslash \omega$ such that $[X_i, X_j](q) \notin \mathrm{Span}(X_1(q), \ldots, X_m(q))$. Then the subelliptic wave equation (1) is not exactly observable on ω in time T_0.*

Consequently, using a duality argument (see Section 4.2), we obtain that exact controllability also does not hold in any finite time.

Definition 3. Let $T_0 > 0$ and $\omega \subset M$ be a measurable subset. The subelliptic wave equation (1) is exactly controllable on ω in time T_0 if for any $(u_0, u_1) \in D((-\Delta)^{1/2}) \times L^2(M)$ there exists $g \in L^2((0, T_0) \times M)$ such that the solution u of

$$\begin{cases} \partial_{tt}^2 u - \Delta u = \mathbf{1}_\omega g & \text{in } (0, T_0) \times M, \\ u = 0 & \text{on } (0, T_0) \times \partial M, \\ (u_{|t=0}, \partial_t u_{|t=0}) = (u_0, u_1), \end{cases} \tag{5}$$

satisfies $u(T_0, \cdot\,) = 0$.

Corollary 4. *Let $T_0 > 0$ and let $\omega \subset M$ be a measurable subset. We assume that there exist $1 \leqslant i, j \leqslant m$ and q in the interior of $M \backslash \omega$ such that $[X_i, X_j](q) \notin \mathrm{Span}(X_1(q), \ldots, X_m(q))$. Then the subelliptic wave equation (1) is not exactly controllable on ω in time T_0.*

In what follows, we denote by $\mathcal{D}$ the set of all vector fields that can be decomposed as linear combinations with smooth coefficients of the X_i:

$$\mathcal{D} = \mathrm{Span}(X_1, \ldots, X_m) \subset TM.$$

$\mathcal{D}$ is called the *distribution* associated to the vector fields $X_1, \ldots, X_m$. For $q \in M$, we denote by $\mathcal{D}_q \subset T_q M$ the distribution $\mathcal{D}$ taken at point q.

The assumptions of Theorem 2 are satisfied as soon as the interior U of $M \setminus \omega$ is nonempty and $\mathcal{D}$ has constant rank $< n$ in U. Indeed, under these conditions, we can argue by contradiction: assume that for any $q \in U$ and any $1 \leqslant i, j \leqslant m$, it holds $[X_i, X_j](q) \in \mathrm{Span}(X_1(q), \ldots, X_m(q)) = \mathcal{D}_q$. Then we have $[\mathcal{D}, \mathcal{D}] \subset \mathcal{D}$ in U, i.e., $\mathcal{D}$ is involutive. By Frobenius's theorem, $\mathcal{D}$ is then completely integrable, which contradicts Hörmander's condition.

The following examples show that the assumptions of Theorem 2 are also satisfied in some nonconstant-rank cases:

Example 5. In the Baouendi–Grushin case, for which $X_1 = \partial_{x_1}$ and $X_2 = x_1 \partial_{x_2}$ are vector fields on $(-1, 1)_{x_1} \times \mathbb{T}_{x_2}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the corresponding sub-Laplacian $\Delta = X_1^2 + X_2^2$ (here, $\mu = dx_1\, dx_2$ for simplicity) is elliptic outside of the singular submanifold $S = \{x_1 = 0\}$. Therefore, the corresponding subelliptic wave equation is observable on any open subset containing S (with some finite minimal time of observability, see [Bardos et al. 1992]), but according to Theorem 2, it is not observable in any finite time on any subset ω such that the interior of $M \setminus \omega$ has a nonempty intersection with S.

Example 6. In the Martinet case, the vector fields are $X_1 = \partial_{x_1}$ and $X_2 = \partial_{x_2} + x_1^2 \partial_{x_3}$ on $(-1, 1)_{x_1} \times \mathbb{T}_{x_2} \times \mathbb{T}_{x_3}$, and the corresponding sub-Laplacian is $\Delta = X_1^2 + X_2^2$ (again, $\mu = dx_1\, dx_2\, dx_3$ for simplicity).

Then, we have $[X_1, X_2] = 2x_1\partial_{x_3}$. The only points at which this bracket belongs to the distribution $\mathrm{Span}(X_1, X_2)$ are the points for which $x_1 = 0$. Since this set of points has empty interior, the assumptions of Theorem 2 are satisfied as soon as $M \setminus \omega$ has nonempty interior.

Remark 7. The assumption of compactness on M is not necessary; we may remove it and just require that the subelliptic wave equation (1) in M is well-posed. It is for example the case if M is complete for the sub-Riemannian distance induced by $X_1, \dots, X_m$ since Δ is then essentially self-adjoint [Strichartz 1986].

Remark 8. Theorem 2 remains true if M has no boundary. In this case, (1) is well-posed in a space slightly smaller than (2): a condition of null average has to be added since nonzero constant functions on M are solutions of (1); see Section 1.5. The observability inequality of Theorem 2 remains true in this space of solutions; anticipating the proof, we notice that the spiraling normal geodesics of Proposition 17 still exist (since their construction is purely local), and we subtract from the initial datum u_0^k of the localized solutions constructed in Proposition 16 their spatial average $\int_M u_0^k \, d\mu$.

Remark 9. Thanks to abstract results (see for example [Miller 2012]), Theorem 2 remains true when the subelliptic wave equation (1) is replaced by the subelliptic half-wave equation $\partial_t u + i\sqrt{-\Delta}u = 0$ with Dirichlet boundary conditions.

1.3. ***Ideas of the proof.*** In the sequel, we define a normal geodesic[1] to be the projection on M of a bicharacteristic (parametrized by time) for the principal symbol of the wave equation (1). We will give a more detailed definition in Section 1.4.

The proof of Theorem 2 mainly requires two ingredients:

(1) There exist solutions of the free subelliptic wave equation (1) whose energy concentrates along any given normal geodesic.

(2) There exist normal geodesics which "spiral" around curves transverse to $\mathcal{D}$, and which therefore remain arbitrarily close to their starting point on arbitrarily large time intervals.

Combining the two above facts, the proof of Theorem 2 is straightforward (see Section 4.1). Note that the first point follows from the general theory of propagation of complex Lagrangian spaces, while the second point is the main novelty of this paper.

Since our construction is purely local (meaning that it does not "feel" the boundary and only relies on the local structure of the vector fields), we can focus on the case where there is a (small) open neighborhood V of the origin O such that $V \subset M \backslash \omega$, and $[X_i, X_j](O) \notin \mathcal{D}_O$ for some $1 \leqslant i, j \leqslant m$. In the sequel, we assume it is the case.

Let us give an example of vector fields where the spiraling normal geodesics used in the proof of Theorem 2 are particularly simple. We consider the three-dimensional manifold with boundary $M_1 = (-1, 1)_{x_1} \times \mathbb{T}_{x_2} \times \mathbb{T}_{x_3}$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z} \approx (-1, 1)$ is the one-dimensional torus. We endow M_1 with

[1] This terminology is common in sub-Riemannian geometry, and it is justified by the fact that we can naturally associate to the vector fields $X_1, \dots, X_m$ a metric structure on M for which these projected paths are geodesics; see [Montgomery 2002].

the vector fields $X_1 = \partial_{x_1}$ and $X_2 = \partial_{x_2} - x_1\partial_{x_3}$. This is the Heisenberg manifold with boundary. We endow M_1 with an arbitrary smooth volume μ. The normal geodesics we consider are given by

$$\begin{aligned} x_1(t) &= \varepsilon \sin(t/\varepsilon), \\ x_2(t) &= \varepsilon \cos(t/\varepsilon) - \varepsilon, \\ x_3(t) &= \varepsilon(t/2 - \varepsilon \sin(2t/\varepsilon)/4). \end{aligned} \tag{6}$$

They spiral around the x_3-axis $x_1 = x_2 = 0$.

Here, one should think of ε as a small parameter. In the sequel, we denote by x_ε the normal geodesic with parameter ε.

Clearly, given any $T_0 > 0$, for ε sufficiently small, we have $x_\varepsilon(t) \in V$ for every $t \in (0, T_0)$. Our objective is to construct solutions u^k of the subelliptic wave equation (1) such that $\|(u_0^k, u_1^k)\|_{\mathcal{H}\times L^2} = 1$ and the energy of $u^k(t, \cdot)$ concentrates outside of an open set V_t containing $x_\varepsilon(t)$, i.e.,

$$\int_{M_1\setminus V_t} \big(|\partial_t u^k(t, x)|^2 + (X_1 u_k(t, x))^2 + (X_2 u_k(t, x))^2\big)\, d\mu(x)$$

tends to 0 as $k \to +\infty$ uniformly with respect to $t \in (0, T_0)$. As a consequence, the observability inequality (4) fails.

The construction of solutions of the free wave equation whose energy concentrates on geodesics is classical in the elliptic (or Riemannian) case; these are the so-called Gaussian beams, for which a construction can be found for example in [Ralston 1982]. Here, we adapt this construction to our subelliptic (sub-Riemannian) setting, which does not raise any problem since the normal geodesics we consider stay in the elliptic part of the operator Δ. It may also be directly justified with the theory of propagation of complex Lagrangian spaces (see Section 2).

In the case of general vector fields $X_1, \dots, X_m$, the existence of spiraling normal geodesics also has to be justified. For that purpose, we first approximate $X_1, \dots, X_m$ by their *nilpotent approximations*, and we then prove that, for these approximations, such a family of spiraling normal geodesics exists, as in the Heisenberg case.

1.4. ***Normal geodesics.*** In this section, we explain in more details what *normal geodesics* are. As said before, they are natural extensions of Riemannian geodesics since they are projections of bicharacteristics.

We denote by $S^m_{\mathrm{phg}}(T^*((0, T) \times M))$ the set of polyhomogeneous symbols of order m with compact support and by $\Psi^m_{\mathrm{phg}}((0, T) \times M)$ the set of associated polyhomogeneous pseudodifferential operators of order m whose distribution kernel has compact support in $(0, T) \times M$ (see Appendix A).

We set $P = \partial^2_{tt} - \Delta \in \Psi^2_{\mathrm{phg}}((0, T) \times M)$, whose principal symbol is

$$p_2(t, \tau, x, \xi) = -\tau^2 + g^*(x, \xi),$$

with τ the dual variable of t and g^* the principal symbol of $-\Delta$. For $\xi \in T^*M$, we have (see Appendix A)

$$g^* = \sum_{i=1}^m h^2_{X_i}.$$

Here, given any smooth vector field X on M, we denote by h_X the Hamiltonian function (momentum map) on T^*M associated with X defined in local (x,ξ)-coordinates by $h_X(x,\xi)=\xi(X(x))$.

In $T^*(\mathbb{R}\times M)$, the Hamiltonian vector field $\vec{p}_2$ associated with p_2 is given by $\vec{p}_2 f=\{p_2, f\}$, where $\{\cdot,\cdot\}$ denotes the Poisson bracket (see Appendix A). Since $\vec{p}_2 p_2=0$, we get that p_2 is constant along the integral curves of $\vec{p}_2$. Thus, the characteristic set $\mathcal{C}(p_2)=\{p_2=0\}$ is preserved by the flow of $\vec{p}_2$. Null-bicharacteristics are then defined as the maximal integral curves of $\vec{p}_2$ which live in $\mathcal{C}(p_2)$. In other words, the null-bicharacteristics are the maximal solutions of

$$\begin{cases} \dot{t}(s)=-2\tau(s), \\ \dot{x}(s)=\nabla_\xi g^*(x(s),\xi(s)), \\ \dot{\tau}(s)=0, \\ \dot{\xi}(s)=-\nabla_x g^*(x(s),\xi(s)), \\ \tau^2(0)=g^*(x(0),\xi(0)). \end{cases} \tag{7}$$

This definition needs to be adapted when the null-bicharacteristic meets the boundary ∂M, but in the sequel, we only consider solutions of (7) on time intervals where $x(t)$ does not reach ∂M.

In the sequel, we take $\tau=-\frac{1}{2}$, which gives $g^*(x(s),\xi(s))=\frac{1}{4}$. This also implies that $t(s)=s+t_0$ and, taking t as a time parameter, we are led to solve

$$\begin{cases} \dot{x}(t)=\nabla_\xi g^*(x(t),\xi(t)), \\ \dot{\xi}(t)=-\nabla_x g^*(x(t),\xi(t)), \\ g^*(x(0),\xi(0))=\frac{1}{4}. \end{cases} \tag{8}$$

In other words, the t-variable parametrizes null-bicharacteristics in a way that they are traveled at speed 1.

Remark 10. In the subelliptic setting, the cosphere bundle S^*M can be decomposed as $S^*M=U^*M\cup S\Sigma$, where $U^*M=\left\{g^*=\frac{1}{4}\right\}$ is a cylinder bundle, $\Sigma=\{g^*=0\}$ is the characteristic cone and $S\Sigma$ is the sphere bundle of Σ; see [Colin de Verdière et al. 2018, Section 1].

We denote by $\phi_t : S^*M\to S^*M$ the (normal) geodesic flow defined by $\phi_t(x_0,\xi_0)=(x(t),\xi(t))$, where $(x(t),\xi(t))$ is a solution of the system given by the first two lines of (8) and initial conditions (x_0,ξ_0). Note that any point in $S\Sigma$ is a fixed point of ϕ_t and that the other normal geodesics are traveled at speed 1 since we took $g^*=\frac{1}{4}$ in U^*M (see Remark 10).

The curves $x(t)$ which solve (8) are geodesics (i.e., local minimizers) for a sub-Riemannian metric g; see [Montgomery 2002, Theorem 1.14].

1.5. ***Observability in some regions of phase-space.*** We have explained in Section 1.3 that the existence of solutions of the subelliptic wave equation (1) concentrated on spiraling normal geodesics is an obstruction to observability in Theorem 2. Our goal in this section is to state a result ensuring observability if one “removes” in some sense these normal geodesics.

For this result, we focus on a version of the Heisenberg manifold described in Section 1.3 which has *no boundary*. This technical assumption avoids us using boundary microlocal defect measures in the proof, which, in this sub-Riemannian setting, are difficult to handle. As a counterpart, we need to consider solutions of the wave equation with null initial average, in order to get well-posedness.

We consider the Heisenberg group G, that is, $\mathbb{R}^3$ with the composition law

$$(x_1, x_2, x_3) \star (x_1', x_2', x_3') = (x_1 + x_1', x_2 + x_2', x_3 + x_3' - x_1 x_2').$$

Then $X_1 = \partial_{x_1}$ and $X_2 = \partial_{x_2} - x_1 \partial_{x_3}$ are left-invariant vector fields on G. Since $\Gamma = \sqrt{2\pi}\mathbb{Z} \times \sqrt{2\pi}\mathbb{Z} \times 2\pi\mathbb{Z}$ is a co-compact subgroup of G, the left quotient $M_H = \Gamma \backslash G$ is a compact three-dimensional manifold and, moreover, X_1 and X_2 are well-defined as vector fields on the quotient. We call M_H endowed with the vector fields X_1 and X_2 the "Heisenberg manifold without boundary". Finally, we define the Heisenberg Laplacian $\Delta_H = X_1^2 + X_2^2$ on M_H. Since $[X_1, X_2] = -\partial_{x_3}$, it is a hypoelliptic operator. We endow M_H with an arbitrary smooth volume μ.

We introduce the space

$$L_0^2 = \left\{ u_0 \in L^2(M_H),\ \int_{M_H} u_0 \, d\mu = 0 \right\}$$

and we consider the operator Δ_H whose domain $D(\Delta_H)$ is the completion in L_0^2 of the set of all $u \in C_c^\infty(M_H)$ with null-average for the norm $\|(\mathrm{Id} - \Delta_H)u\|_{L^2}$. Then, $-\Delta_H$ is positive definite and we consider $(-\Delta_H)^{1/2}$ with domain $D((-\Delta_H)^{1/2}) = \mathcal{H}_0 := L_0^2 \cap \mathcal{H}(M_H)$. The wave equation

$$\begin{cases} \partial_{tt}^2 u - \Delta_H u = 0 & \text{in } \mathbb{R} \times M_H, \\ (u_{|t=0}, \partial_t u_{|t=0}) = (u_0, u_1) \in D((-\Delta_H)^{\frac{1}{2}}) \times L_0^2, \end{cases} \tag{9}$$

admits a unique solution $u \in C^0(\mathbb{R}; D((-\Delta_H)^{1/2})) \cap C^1(\mathbb{R}; L_0^2)$.

We note that $-\Delta_H$ is invertible in L_0^2. The space $\mathcal{H}_0$ is endowed with the norm $\|u\|_{\mathcal{H}}$ (defined in (3) and also equal to $\|(-\Delta_H)^{1/2} u\|_{L^2}$), and its topological dual $\mathcal{H}_0'$ is endowed with the norm $\|u\|_{\mathcal{H}_0'} := \|(-\Delta_H)^{-1/2} u\|_{L^2}$.

We note that $g^*(x, \xi) = \xi_1^2 + (\xi_2 - x_1 \xi_3)^2$ and hence the null-bicharacteristics are solutions of

$$\begin{aligned} \dot{x}_1(t) &= 2\xi_1, & \dot{\xi}_1(t) &= 2\xi_3(\xi_2 - x_1\xi_3), \\ \dot{x}_2(t) &= 2(\xi_2 - x_1\xi_3), & \dot{\xi}_2(t) &= 0, \\ \dot{x}_3(t) &= -2x_1(\xi_2 - x_1\xi_3), & \dot{\xi}_3(t) &= 0. \end{aligned} \tag{10}$$

The spiraling normal geodesics described in Section 1.3 correspond to $\xi_1 = \cos(t/\varepsilon)/2$, $\xi_2 = 0$ and $\xi_3 = 1/(2\varepsilon)$. In particular, the constant ξ_3 is a kind of rounding number reflecting the fact that the normal geodesic spirals at a certain speed around the x_3-axis. Moreover, ξ_3 is preserved under the flow (somehow, the Heisenberg flow is completely integrable), and this property plays a key role in the proof of Theorem 11 below and justifies that we state it only for the Heisenberg manifold (without boundary).

As said above, normal geodesics corresponding to a large momentum ξ_3 are precisely the ones used to contradict observability in Theorem 2. We expect to be able to establish observability if we consider only solutions of (1) whose ξ_3 (in a certain sense) is not too large. This is the purpose of our second main result.

Set

$$V_\varepsilon = \left\{ (x, \xi) \in T^* M_H : |\xi_3| > \frac{1}{\varepsilon} (g_x^*(\xi))^{\frac{1}{2}} \right\}.$$

Note that since ξ_3 is constant along null-bicharacteristics, V_ε and its complement V_ε^c are invariant under the bicharacteristic equations (10).

In the next statement, we define a horizontal strip to be the periodization under the action of the co-compact subgroup Γ of a set of the form

$$\{(x_1, x_2, x_3) : (x_1, x_2) \in [0, \sqrt{2\pi})^2,\ x_3 \in I\},$$

where I is a strict open subinterval of $[0, 2\pi)$.

Theorem 11. *Let $B \subset M_H$ be an open subset and suppose that B is sufficiently small, so that $\omega = M_H \backslash B$ contains a horizontal strip. Let $a \in S^0_{\mathrm{phg}}(T^*M_H)$, $a \geqslant 0$, such that, denoting by $j : T^*\omega \to T^*M_H$ the canonical injection,*

$$j(T^*\omega) \cup V_\varepsilon \subset \mathrm{Supp}(a) \subset T^*M_H,$$

and in particular a does not depend on time. There exists $\kappa > 0$ such that, for any $\varepsilon > 0$ and any $T \geqslant \kappa\varepsilon^{-1}$, it holds

$$C\|(u(0), \partial_t u(0))\|^2_{\mathcal{H}_0 \times L^2_0} \leqslant \int_0^T |(\mathrm{Op}(a)\partial_t u, \partial_t u)_{L^2}|\, dt + \|(u(0), \partial_t u(0))\|^2_{L^2_0 \times \mathcal{H}_0'} \tag{11}$$

for some $C = C(\varepsilon, T) > 0$ and for any solution $u \in C^0(\mathbb{R}; D((-\Delta_H)^{1/2})) \cap C^1(\mathbb{R}; L^2_0)$ of (9).

The term $\|(u_0, u_1)\|^2_{L^2 \times \mathcal{H}_0'}$ in the right-hand side of (11) cannot be removed; i.e., our statement only consists of a *weak* observability inequality. Indeed, the usual way to remove such terms is to use a unique continuation argument for eigenfunctions φ of Δ, but here it does not work since $\mathrm{Op}(a)\varphi = 0$ does not imply in general that $\varphi \equiv 0$ in the whole manifold, even if the support of a contains $j(T^*\omega)$ for some nonempty open set ω: in some sense, there is no "pseudodifferential unique continuation argument".

1.6. *Comments on the existing literature.*

Elliptic and subelliptic waves. The exact controllability/observability of the elliptic wave equation is known to be almost equivalent to the so-called geometric control condition (GCC) (see [Bardos et al. 1992]) that any geodesic enters the control set ω within time T. In some sense, our main result is that GCC is not satisfied in the subelliptic setting, as soon as $M\backslash\omega$ contains in its interior a point x at which Δ is "truly subelliptic". For the elliptic wave equation, in many geometrical situations, there exists a minimal time $T_0 > 0$ such that observability holds only for $T \geqslant T_0$: when there exists a geodesic $\gamma : (0, T_0) \to M$ traveled at speed 1 which does not meet $\bar{\omega}$, one constructs a sequence of initial data $(u_0^k, u_1^k)_{k\in\mathbb{N}^*}$ of the wave equation whose associated microlocal defect measure is concentrated on $(x_0, \xi_0) \in S^*M$ taken to be the initial conditions for the null-bicharacteristic projecting onto γ. Then, the associated sequence of solutions $(u^k)_{k\in\mathbb{N}^*}$ of the wave equation has an associated microlocal defect measure ν which is invariant under the geodesic flow: $\vec{p}\nu = 0$, where $\vec{p}$ is the Hamiltonian flow associated to the principal symbol p of the wave operator. In particular, denoting by $\pi : T^*M \to M$ the canonical projection, $\pi_*\nu$ gives no mass to ω since γ is contained in $M \setminus \bar{\omega}$, and this proves that observability cannot hold.

In the subelliptic setting, the invariance property $\vec{p}\nu = 0$ does not give any information on ν on the characteristic manifold Σ, since $\vec{p} = -2\tau\partial_t + \vec{g}^*$ vanishes on Σ. This is related to the lack of information on propagation of singularities in this characteristic manifold; see the main theorem of [Lascar 1982]. If one instead tries to use the propagation of the microlocal defect measure for subelliptic half-wave equations, one is immediately confronted with the fact that $\sqrt{-\Delta}$ is not a pseudodifferential operator near Σ.

This is why, in this paper, we used only the elliptic part of the symbol g^* (or, equivalently, the strictly hyperbolic part of p_2), where the propagation properties can be established, and then the problem is reduced to proving geometric results on normal geodesics.

Subelliptic Schrödinger equations. The recent article [Burq and Sun 2019] deals with the same observability problem, but for subelliptic Schrödinger equations: namely, the authors consider the Baouendi–Grushin Schrödinger equation $i\partial_t u - \Delta_G u = 0$, where $u \in L^2((0,T)\times M_G)$, $M_G = (-1,1)_x \times \mathbb{T}_y$ and $\Delta_G = \partial_x^2 + x^2\partial_y^2$ is the Baouendi–Grushin Laplacian. Given a control set of the form $\omega = (-1,1)_x \times \omega_y$, where ω_y is an open subset of $\mathbb{T}$, the authors prove the existence of a minimal time of control $\mathcal{L}(\omega)$ related to the maximal height of a horizontal strip contained in $M_G\backslash\omega$. The intuition is that there are solutions of the Baouendi–Grushin Schrödinger equation which travel along the degenerate line $x=0$ at a finite speed; in some sense, along this line, the Schrödinger equation behaves like a classical (half)-wave equation. What we want here is to explain in a few words why there is a minimal time of observability for the Schrödinger equation, while the wave equation is never observable in finite time as shown by Theorem 2.

The plane $\mathbb{R}^2_{x,y}$ endowed with the vector fields ∂_x and $x\partial_y$ also admits normal geodesics similar to the 1-parameter family q_ε, namely, for $\varepsilon > 0$,

$$\begin{aligned} x(t) &= \varepsilon \sin(t/\varepsilon), \\ y(t) &= \varepsilon(t/2 - \varepsilon \sin(2t/\varepsilon)/4). \end{aligned}$$

These normal geodesics, denoted by γ_ε, also “spiral” around the line $x=0$ more and more quickly as $\varepsilon \to 0$, and so we might expect to construct solutions of the Baouendi–Grushin Schrödinger equation with energy concentrated along γ_ε, which would contradict observability when $\varepsilon \to 0$ as above for the Heisenberg wave equation.

However, we can convince ourselves that it is not possible to construct such solutions: in some sense, the dispersion phenomena of the Schrödinger equation exactly compensate for the lengthening of the normal geodesics γ_ε as $\varepsilon \to 0$ and explain that even these Gaussian beams may be observed in ω from a certain minimal time $\mathcal{L}(\omega) > 0$ which is uniform in ε.

To put this argument into a more formal form, we consider the solutions of the bicharacteristic equations for the Baouendi–Grushin Schrödinger equation $i\partial_t u - \Delta_G u = 0$ given by

$$\begin{aligned} x(t) &= \varepsilon \sin(\xi_y t), & \xi_x(t) &= \varepsilon \xi_y \cos(\xi_y t), \\ y(t) &= \varepsilon^2 \xi_y \left(\frac{t}{2} - \frac{\sin(2\xi_y t)}{4\xi_y}\right), & \xi_y(t) &= \xi_y. \end{aligned}$$

It follows from the hypoellipticity of Δ_G (see [Burq and Sun 2019, Section 3] for a proof) that

$$|\xi_y|^{1/2} \lesssim \sqrt{-\Delta_G} = (|\xi_x|^2 + x^2|\xi_y|^2)^{1/2} = \varepsilon|\xi_y|.$$

Therefore $\varepsilon^2|\xi_y| \gtrsim 1$, and hence $|y(t)| \gtrsim t$, independently from ε and ξ_y. This heuristic gives the intuition that a minimal time $\mathcal{L}(\omega)$ is required to detect all solutions of the Baouendi–Grushin Schrödinger equation from ω, but that for $T_0 > \mathcal{L}(\omega)$, no solution is localized enough to stay in $M\backslash\omega$ during the time interval $(0, T_0)$. Roughly speaking, the frequencies of order ξ_y travel at speed $\sim \xi_y$, which is typical for a dispersion phenomenon. This picture is very different from the one for the wave equation (which we consider in this paper) for which no dispersion occurs.

With similar ideas, in [Letrouit and Sun 2021], the interplay between the subellipticity effects measured by the nonholonomic order of the distribution $\mathcal{D}$ (see Section 3.1) and the strength of dispersion of Schrödinger-type equations was investigated. More precisely, for $\Delta_\gamma = \partial_x^2 + |x|^{2\gamma}\partial_y^2$ on $M = (-1, 1)_x \times \mathbb{T}_y$, and for $s \in \mathbb{N}$, the observability properties of the Schrödinger-type equation $(i\partial_t - (-\Delta_\gamma)^s)u = 0$ were shown to depend on the value $\kappa = 2s/(\gamma + 1)$. In particular it is proved that, for $\kappa < 1$, observability fails for any time, which is consistent with the present result, and that for $\kappa = 1$, observability holds only for sufficiently large times, which is consistent with the result of [Burq and Sun 2019]. The results of [Letrouit and Sun 2021] are somehow Schrödinger analogues of the results of [Beauchard et al. 2014] which deal with a similar problem for the Baouendi–Grushin heat equation.

General bibliographical comments. Control of subelliptic PDEs has attracted much attention in the last decade. Most results in the literature deal with subelliptic parabolic equations, either the Baouendi–Grushin heat equation [Koenig 2017; Duprez and Koenig 2020; Beauchard et al. 2020] or the heat equation in the Heisenberg group [Beauchard and Cannarsa 2017]. The paper [Burq and Sun 2019] was the first to deal with a subelliptic Schrödinger equation and the present work is the first to handle exact controllability of subelliptic wave equations.

A slightly different problem is the *approximate* controllability of hypoelliptic PDEs, which was studied in [Laurent and Léautaud 2022] for hypoelliptic wave and heat equations. Approximate controllability is weaker than exact controllability, and it amounts to proving "quantitative" unique continuation results for hypoelliptic operators. For the hypoelliptic wave equation, it is proved in [Laurent and Léautaud 2022] that for $T > 2\sup_{x\in M}(\mathrm{dist}(x, \omega))$ (here, dist is the sub-Riemannian distance), the observation of the solution on $(0, T) \times \omega$ determines the initial data, and therefore the whole solution.

1.7. ***Organization of the paper.*** In Section 2, we construct exact solutions of the subelliptic wave equation (1) concentrating on any given normal geodesic. First, in Section 2.1, we show that, given any normal geodesic $t \mapsto x(t)$ which does not hit ∂M in the time interval $(0, T)$, it is possible to construct a sequence $(v_k)_{k\in\mathbb{N}}$ of *approximate solutions* of (1) whose energy concentrates along $t \mapsto x(t)$ during the time interval $(0, T)$ as $k \to +\infty$. By "approximate", we mean here that $\partial_{tt}^2 v_k - \Delta v_k$ is small, but not necessarily exactly equal to 0. In Section 2.1, we provide a first proof for this construction using the classical propagation of complex Lagrangian spaces. Another proof using a Gaussian beam approach is provided in Appendix B. Then, in Section 2.2, using this sequence $(v_k)_{k\in\mathbb{N}}$, we explain how to construct a sequence $(u_k)_{k\in\mathbb{N}}$ of *exact* solutions of $(\partial_{tt}^2 - \Delta)u = 0$ in M with the same concentration property along the normal geodesic $t \mapsto x(t)$.

In Section 3, we prove the existence of normal geodesics which spiral in M, spending an arbitrarily large time in $M\backslash\omega$. These normal geodesics generalize the example described in Section 1.3 for the Heisenberg manifold with boundary. The proof proceeds in two steps: first, we show that it is sufficient to prove the result in the so-called "nilpotent case" (Section 3.2), and then we prove it in the nilpotent case (Section 3.3).

In Section 4.1, we use the results of Sections 2 and 3 to conclude the proof of Theorem 2. In Section 4.2, we deduce Corollary 4 by a duality argument. Finally, in Section 4.3, we prove Theorem 11.

2. Gaussian beams along normal geodesics

2.1. ***Construction of sequences of approximate solutions.*** We consider a solution $(x(t), \xi(t))_{t\in[0,T]}$ of (8) on M. We shall describe the construction of solutions of

$$\partial_{tt}^2 u - \Delta u = 0 \tag{12}$$

on $[0, T] \times M$ with energy

$$E(u(t,\cdot)) := \tfrac{1}{2}(\|\partial_t u(t,\cdot)\|^2_{L^2(M,\mu)} + \|u(t,\cdot)\|^2_{\mathcal{H}})$$

concentrated along $x(t)$ for $t \in [0, T]$. The following proposition, which is inspired by [Ralston 1982; Macià and Zuazua 2002], shows that it is possible, at least for approximate solutions of (12).

Proposition 12. *Fix $T > 0$ and let $(x(t), \xi(t))_{t\in[0,T]}$ be a solution of (8) $\big($in particular $g^*(x(0), \xi(0)) = \frac{1}{4}\big)$ which does not hit the boundary ∂M in the time interval $(0, T)$. Then there exist $a_0, \psi \in C^2((0, T) \times M)$ such that, setting, for $k \in \mathbb{N}$,*

$$v_k(t, x) = k^{\frac{n}{4}-1} a_0(t, x) e^{ik\psi(t,x)}$$

the following properties hold:

- *v_k is an approximate solution of (12), meaning that*

$$\|\partial_{tt}^2 v_k - \Delta v_k\|_{L^1((0,T);L^2(M))} \leqslant Ck^{-\frac{1}{2}}. \tag{13}$$

- *The energy of v_k is bounded below with respect to k and $t \in [0, T]$:*

$$\text{there exists } A > 0 \text{ such that, for all } t \in [0, T], \quad \liminf_{k\to+\infty} E(v_k(t,\cdot)) \geqslant A. \tag{14}$$

- *The energy of v_k is small off $x(t)$: For any $t \in [0, T]$, we fix V_t an open subset of M for the initial topology of M, containing $x(t)$, so that the mapping $t \mapsto V_t$ is continuous (V_t is chosen sufficiently small so that this makes sense in a chart). Then*

$$\sup_{t\in[0,T]} \int_{M\setminus V_t} \bigg(|\partial_t v_k(t,x)|^2 + \sum_{j=1}^m (X_j v_k(t,x))^2 \bigg) d\mu(x) \xrightarrow[k\to+\infty]{} 0. \tag{15}$$

Remark 13. The construction of approximate solutions such as the ones provided by Proposition 12 is usually done for strictly hyperbolic operators, that is, operators with a principal symbol p_m of order m such that the polynomial $f(s) = p_m(t, q, s, \xi)$ has m distinct real roots when $\xi \neq 0$; see for example [Ralston 1982]. The operator $\partial_{tt}^2 - \Delta$ is not strictly hyperbolic because g^* is degenerate, but our proof shows that the same construction may be adapted without difficulty to this operator *along normal bicharacteristics*. This is due to the fact that along normal bicharacteristics, $\partial_{tt}^2 - \Delta$ is indeed strictly hyperbolic (or equivalently, Δ is elliptic). It was already noted by [Ralston 1982] that the construction of Gaussian beams could be done for more general operators than strictly hyperbolic ones, and that the differences between the strictly hyperbolic case and more general cases arise while dealing with propagation of singularities. Also, in [Hörmander 1985, Chapter 24.2], it was noticed that "since only microlocal properties of p_2 are important, it is easy to see that hyperbolicity may be replaced by $\nabla_\xi p_2 \neq 0$."

Hereafter we provide two proofs of Proposition 12. The first proof is short and is actually quite straightforward for readers acquainted with the theory of propagation of complex Lagrangian spaces, once one has noticed that the solutions of (8) which we consider live in the *elliptic part* of the principal symbol of $-\Delta$. For the sake of completeness, and because this also has its own interest, we provide in Appendix B a second proof, longer but more elementary and accessible without any knowledge of complex Lagrangian spaces; it relies on the construction of Gaussian beams in the subelliptic context. The two proofs follow parallel paths, and indeed, the computations which are only sketched in the first proof are written in full detail in the second proof, given in Appendix B.

First proof of Proposition 12. The construction of Gaussian beams, or more generally of a WKB approximation, is related to the transport of complex Lagrangian spaces along bicharacteristics, as reported for example in [Hörmander 1985, Chapter 24.2; Ivrii 2019, Volume I, Part I, Chapter 1.2]. Our proof follows the lines of [Hörmander 1985, pages 426–428].

A usual way to solve (at least approximately) evolution equations of the form

$$Pu = 0, \tag{16}$$

where P is a hyperbolic second-order differential operator with real principal symbol and C^∞ coefficients, is to search for oscillatory solutions

$$v_k(x) = k^{\frac{n}{4}-1} a_0(x) e^{ik\psi(x)}. \tag{17}$$

In this expression as in the rest of the proof, we suppress the time variable t. Thus, we use $x = (x_0, x_1, \dots, x_n)$, where $x_0 = t$ in the earlier notation, and we set $x' = (x_1, \dots, x_n)$. Similarly, we take the notation $\xi = (\xi_0, \xi_1, \dots, \xi_n)$, where $\xi_0 = \tau$ previously, and $\xi' = (\xi_1, \dots, \xi_n)$. The bicharacteristics are parametrized by s as in (7), and without loss of generality, we only consider bicharacteristics with $x(0) = 0$ at $s = 0$, which implies in particular $x_0(s) = s$ because of our choice $\tau^2(s) = g^*(x(s), \xi(s)) = \frac{1}{4}$.

Taking charts of M, we can assume $M \subset \mathbb{R}^n$. The precise argument for reducing to this case is written at the end of Appendix B. Also, in the sequel, $P = \partial_{tt}^2 - \Delta$.

Plugging the ansatz (17) into (16), we get

$$Pv_k = (k^{\frac{n}{4}+1} A_1 + k^{\frac{n}{4}} A_2 + k^{\frac{n}{4}-1} A_3) e^{ik\psi}, \tag{18}$$

with

$$A_1(x) = p_2(x, \nabla\psi(x)) a_0(x), \quad A_2(x) = La_0(x), \quad A_3(x) = \partial_{tt}^2 a_0(x) - \Delta a_0(x),$$

and L is a transport operator given by

$$La_0 = \frac{1}{i} \sum_{j=0}^{n} \frac{\partial p_2}{\partial \xi_j}(x, \nabla\psi(x)) \frac{\partial a_0}{\partial x_j} + \frac{1}{2i} \left(\sum_{j,k=0}^{n} \frac{\partial^2 p_2}{\partial \xi_j \partial \xi_k}(x, \nabla\psi(x)) \frac{\partial^2 \psi}{\partial x_j \partial x_k} \right) a_0. \tag{19}$$

For v_k to be an approximate solution of P, we are first led to cancel the higher-order term in (18), i.e.,

$$f(x) := p_2(x, \nabla\psi(x)) = 0, \tag{20}$$

which we solve for initial conditions

$$\psi(0, x') = \psi_0(x'), \quad \nabla\psi_0(0) = \xi'(0) \quad \text{and} \quad \psi_0(0) = 0 \tag{21}$$

(i.e., we fix such a ψ_0, and then we solve (20) for ψ). Indeed, it will be sufficient for our purpose for (20) to be satisfied at second order along the curve $x(s)$; i.e., $D_x^\alpha f(x(s)) = 0$ for any $|\alpha| \leqslant 2$ and any s. For that, we first notice that the choice $\nabla\psi(x(s)) = \xi(s)$ ensures that (20) holds at orders 0 and 1 along the curve $s \mapsto x(s)$ (see Appendix B for detailed computations). Now, we explain how to choose $D^2\psi(x(s))$ adequately in order for (20) to hold at order 2.

We use the decomposition of p_2 into

$$p_2(x_0, x', \xi_0, \xi') = -(\xi_0 - r(x', \xi'))(\xi_0 + r(x', \xi')) + R(x', \xi'),$$

where $r = \sqrt{g^*}$ in a conic neighborhood of $(0, \xi(0))$. Note that $\sqrt{g^*}$ is smooth in small conic neighborhoods of $(0, \xi(0))$ since $g^*(0, \xi(0)) = \frac{1}{4} \neq 0$. Indeed, g^* is elliptic along the whole bicharacteristic since $g^*(x(t), \xi(t)) = \frac{1}{4}$ is preserved by the bicharacteristic flow. The rest term $R(x', \xi')$ is smooth and microlocally supported far from the bicharacteristic; i.e., $R(x', \xi') = 0$ for any $(x', \xi') \in T^*M$ in a conic neighborhood of $(x'(s), \xi'(s))$ for $s \in [0, T]$.

We consider the bicharacteristic γ_+ starting at $(0, 0, r(0, \xi'(0)), \xi'(0))$ and the bicharacteristic γ_- starting at $(0, 0, -r(0, \xi'(0)), \xi'(0))$.

We denote by $\Phi^\pm(x_0, y', \eta')$ the solution of the Hamilton equations with Hamiltonian $H_\pm(x_0, x', \xi') = \xi_0 \mp r(x', \xi')$ and initial datum $(x', \xi') = (y', \eta')$ at $x_0 = 0$. In other words, $\Phi^\pm(x_0, y', \eta') = e^{x_0 \vec{H}_\pm}(0, y', \eta')$. Then, for any s, $\Phi(s, \cdot)$ is well-defined and symplectic from a neighborhood of $(0, \xi'(0))$ to a neighborhood of $H_\pm(s, 0, \xi'(0))$.

The solution $\psi(s, \cdot)$ of (20) and (21) is equal to 0 on $\gamma_\pm$ and $\nabla\psi(s, \cdot)$ is obtained by the transport of the values of $\nabla\psi_0$ by $\Phi^\pm(s, \cdot)$. In other words, to compute $\nabla\psi(s, \cdot)$, one transports the Lagrangian subspace $\Lambda_0 = \{(x', \nabla\psi_0(x'))\}$ along the Hamiltonian flow $\vec{H}_\pm$ during a time s, which yields $\Lambda_s \subset T^*M$, and then, if possible, one writes Λ_s under the form $\{(x', \nabla_{x'}\psi(s, x'))\}$, which gives $\nabla_{x'}\psi(s, x')$. The trouble is that the solution is only local in time: when $x' \mapsto \pi(\Phi^\pm(s, x', \nabla\psi_0(x')))$ ceases to be a diffeomorphism (conjugate point), where $\pi : T^*M \to M$ is the canonical projection, we see that the process described above does not work (appearance of caustics). In the language of Lagrangian spaces, $\Lambda_0 = \{(x', \nabla\psi_0(x'))\} \subset T^*M$ is a Lagrangian subspace and, since $\Phi^\pm(s, \cdot)$ is a symplectomorphism, $\Lambda_s = \Phi^\pm(s, \Lambda_0)$ is Lagrangian as well. If $\pi_{|\Lambda_s}$ is a local diffeomorphism, one can locally describe Λ_s by $\Lambda_s = \{(x', \nabla_{x'}\psi(s, x'))\} \subset T^*M$ for some function $\psi(s, \cdot)$, but blow-up happens when $\operatorname{rank}(d\pi_{|\Lambda_s}) < n$ (classical conjugate point theory), and such a $\psi(s, \cdot)$ may not exist.

However, if the phase ψ_0 is complex, quadratic, and satisfies the condition $\operatorname{Im}(D^2\psi_0) > 0$, where $D^2\psi_0$ denotes the Hessian, no blow-up happens, and the solution is global-in-time. Let us explain why. Indeed, $\Lambda_0 = \{(x', \nabla\psi_0(x'))\}$ then lives in the complexification of the tangent space T^*M, which may be thought of as $\mathbb{C}^{2(n+1)}$. We take coordinates (y,η) on $T^*\mathbb{R}^{n+1}$ or $T^*\mathbb{C}^{n+1}$ and we consider the symplectic forms defined by $\sigma=\sum dy_j\wedge d\eta_j$ and $\sigma_{\mathbb{C}}=\sum dy_j\wedge \overline{d\eta_j}$.

Because of the condition $\mathrm{Im}(D^2\psi_0)>0$, Λ_0 is called a "strictly positive Lagrangian space" (see [Hörmander 1985, Definition 21.5.5]), meaning that $i\sigma_{\mathbb{C}}(v,v)>0$ for v in the tangent space to Λ_0. For any s, the symplectic forms σ and $\sigma_{\mathbb{C}}$ are preserved by $\Phi(s,\cdot)$, meaning that $\Phi(s,\cdot)_*\sigma=\sigma$ and $\Phi(s,\cdot)_*\sigma_{\mathbb{C}}=\sigma_{\mathbb{C}}$; therefore $\sigma=0$ on the tangent space to Λ_s, and $i\sigma_{\mathbb{C}}(v,v)>0$ for v tangent to Λ_s. It precisely means that Λ_s is also a strictly positive Lagrangian space. Then, by [Hörmander 1985, Proposition 21.5.9], we know that there exists $\psi(s,\cdot)$ complex and quadratic with $\mathrm{Im}(D^2\psi(s,\cdot))>0$ such that $\Lambda_s=\{(x',\nabla_{x'}\psi(s,x'))\}$ (to apply [Hörmander 1985, Proposition 21.5.9], recall that, for $\varphi(x')=\frac{1}{2}(Ax',x')$, it holds $\nabla\varphi(x')=Ax'$). In other words, the key point in using complex phases is that strictly positive Lagrangian spaces are parametrized by complex quadratic phases φ with $\mathrm{Im}(D^2\varphi)>0$, whereas real Lagrangian spaces were not parametrized by real phases (see explanations above). This parametrization is a diffeomorphism from the Grassmannian of strictly positive Lagrangian spaces to the space of complex quadratic phases with φ with $\mathrm{Im}(D^2\varphi)>0$. Hence, the phase

$$\psi(s,y')=\nabla_{x'}\psi(x(s))\cdot(y'-x'(s))+\tfrac{1}{2}(y'-x'(s))\cdot D^2_{x'}\psi(s,x'(s))(y'-x'(s))$$

for $s\in[0,T]$ and $y'\in\mathbb{R}^n$ is smooth and for this choice (20) is satisfied at second order along $s\mapsto x(s)$ (the rest $R(x',\xi')$ plays no role since it vanishes in a neighborhood of $s\mapsto x(s)$).

Then, we note that A_2 vanishes along the bicharacteristic if and only if $La_0(x(s))=0$ (see also [Hörmander 1985, equation (24.2.9)]). According to (19), this turns out to be a linear transport equation on $a_0(x(s))$, with leading coefficient $\nabla_\xi p_2(x(s),\xi(s))$ different from 0. Given $a\neq 0$ at $(t=0,\ x'=x'(0))$, this transport equation has a solution $a_0(x(s))$ with initial datum a, and, by Cauchy uniqueness, $a_0(x(s))\neq 0$ for any s. We can choose a_0 in a smooth (and arbitrary) way outside the bicharacteristic. We choose it to vanish outside a small neighborhood of this bicharacteristic, so that no boundary effect happens.

With these choices of ψ and a_0, the bound (13) then follows from the following result whose proof is given in [Ralston 1982, Lemma 2.8].

Lemma 14. *Let $c(x)$ be a function on $\mathbb{R}^{n+1}$ which vanishes at order $S-1$ on a curve Γ for some $S\geqslant 1$. Suppose that $\mathrm{Supp}\,c\cap\{|x_0|\leqslant T\}$ is compact and that $\mathrm{Im}\,\psi(x)\geqslant ad(x)^2$ on this set for some constant $a>0$, where $d(x)$ denotes the distance from the point $x\in\mathbb{R}^{d+1}$ to the curve Γ. Then there exists a constant C such that*

$$\int_{|x_0|\leqslant T}|c(x)e^{ik\psi(x)}|^2\,dx\leqslant Ck^{-S-\frac{n}{2}}.$$

Let us now sketch the end of the proof, which is given in Appendix B in full detail. We apply Lemma 14 to $S=3$, $c=A_1$ and to $S=1$, $c=A_2$, and we get

$$\|\partial^2_{tt}v_k-\Delta v_k\|_{L^1(0,T;L^2(M))}\leqslant C(k^{-\frac{1}{2}}+k^{-\frac{1}{2}}+k^{-1}),$$

which implies (13). The bounds (14) and (15) follow from the facts that $\mathrm{Im}(D^2\psi(s,\cdot))>0$ and $v_k(x)=k^{n/4-1}a_0(x)e^{ik\psi(x)}$. ☐

Remark 15. An interesting question would be to understand the delocalization properties of the Gaussian beams constructed along normal geodesics in Proposition 12. Compared with the usual Riemannian case done for example in [Ralston 1982], there is a new phenomenon in the sub-Riemannian case since the normal geodesic $x(t)$ (or, more precisely, its lift to S^*M) may approach the characteristic manifold $\Sigma = \{g^* = 0\}$, which is the set of directions in which Δ is not elliptic. In finite time T as in our case, the lift of the normal geodesic remains far from Σ, but it may happen as $T \to +\infty$ that it goes closer and closer to Σ. The question is then to understand the link between the delocalization properties of the Gaussian beams constructed along such a normal geodesic, and notably the interplay between the time T and the semiclassical parameter $\frac{1}{k}$.

2.2. ***Construction of sequences of exact solutions in M.*** In this section, using the approximate solutions of Section 2.1, we construct *exact* solutions of (12) whose energy concentrates along a given normal geodesic of M which does not meet the boundary ∂M during the time interval $[0, T]$.

Proposition 16. *Let $(x(t), \xi(t))_{t\in[0,T]}$ be a solution of (8) in M $\big($in particular $g^*(x(0), \xi(0)) = \frac{1}{4}\big)$ which does not meet ∂M. Let $\theta \in C_c^\infty([0, T] \times M)$ with $\theta(t, \cdot) \equiv 1$ in a neighborhood of $x(t)$ and such that the support of $\theta(t, \cdot)$ stays at positive distance of ∂M.*

Suppose $(v_k)_{k\in\mathbb{N}}$ is constructed along $x(t)$ as in Proposition 12 and u_k is the solution of the Cauchy problem

$$\begin{cases} (\partial_{tt}^2 - \Delta)u_k = 0 \quad \text{in } (0, T) \times M, \\ u_k = 0 \quad \text{in } (0, T) \times \partial M, \\ u_{k|t=0} = (\theta v_k)_{|t=0}, \quad \partial_t u_{k|t=0} = [\partial_t(\theta v_k)]_{|t=0}. \end{cases}$$

Then:

- *The energy of u_k is bounded below with respect to k and $t \in [0, T]$:*

$$\text{there exists } A > 0 \text{ such that, for all } t \in [0, T], \quad \liminf_{k\to+\infty} E(u_k(t, \cdot)) \geqslant A. \tag{22}$$

- *The energy of u_k is small off $x(t)$: For any $t \in [0, T]$, we fix V_t an open subset of M for the initial topology of M, containing $x(t)$, so that the mapping $t \mapsto V_t$ is continuous (V_t is chosen sufficiently small so that this makes sense in a chart). Then*

$$\sup_{t\in[0,T]} \int_{M\setminus V_t} \Big(|\partial_t u_k(t, x)|^2 + \sum_{j=1}^m (X_j u_k(t, x))^2 \Big) \, d\mu(x) \xrightarrow[k\to+\infty]{} 0. \tag{23}$$

Proof of Proposition 16. Set $h_k = (\partial_{tt}^2 - \Delta)(\theta v_k)$. We consider w_k the solution of the Cauchy problem

$$\begin{cases} (\partial_{tt}^2 - \Delta)w_k = h_k \quad \text{in } (0, T) \times M, \\ w_k = 0 \quad \text{in } (0, T) \times \partial M, \\ (w_{k|t=0}, \partial_t w_{k|t=0}) = (0, 0). \end{cases} \tag{24}$$

Differentiating $E(w_k(t, \cdot))$ and using Gronwall's lemma, we get the energy inequality

$$\sup_{t\in[0,T]} E(w_k(t, \cdot)) \leqslant C(E(w_k(0, \cdot)) + \|h_k\|_{L^1(0,T;L^2(M))}).$$

Therefore, using (13), we get $\sup_{t\in[0,T]} E(w_k(t,\cdot)) \leqslant Ck^{-1}$. Since $u_k = \theta v_k - w_k$, we obtain that

$$\lim_{k\to+\infty} E(u_k(t,\cdot)) = \lim_{k\to+\infty} E((\theta v_k)(t,\cdot)) = \lim_{k\to+\infty} E(v_k(t,\cdot))$$

for every $t \in [0, T]$, where the last equality comes from the fact that θ and its derivatives are bounded and $\|v_k\|_{L^2} \leqslant Ck^{-1}$ when $k \to +\infty$. Using (14), we conclude that (22) holds.

To prove (23), we observe similarly that

$$\sup_{t\in[0,T]} \int_{M\setminus V_t} \left(|\partial_t u_k(t,x)|^2 + \sum_{j=1}^m (X_j u_k(t,x))^2\right) d\mu(x) \\ \leqslant C \sup_{t\in[0,T]} \left(\int_{M\setminus V_t} (|\partial_t v_k(t,x)|^2 + \sum_{j=1}^m (X_j v_k(t,x))^2)\, d\mu(x)\right) + Ck^{-\frac{1}{2}} \to 0$$

as $k \to +\infty$, according to (15). It concludes the proof of Proposition 16. □

3. Existence of spiraling normal geodesics

The goal of this section is to prove the following proposition, which is the second building block of the proof of Theorem 2, after the construction of localized solutions of the subelliptic wave equation (1) done in Section 2.

We say that $X_1, \ldots, X_m$ satisfies the property (P) at $q \in M$ if the following holds:

(P) *For any open neighborhood V of q, for any $T_0 > 0$, there exists a nonstationary normal geodesic $t \mapsto x(t)$, traveled at speed 1, such that $x(t) \in V$ for any $t \in [0, T_0]$.*

Proposition 17. *At any point $q \in M$ such that there exist $1 \leqslant i, j \leqslant m$ with $[X_i, X_j](q) \notin \mathcal{D}_q$, property (P) holds.*

In Section 3.1, we define the so-called nilpotent approximations $\widehat{X}_1^q, \ldots, \widehat{X}_m^q$ at a point $q \in M$, which are first-order approximations of $X_1, \ldots, X_m$ at $q \in M$ such that the associated Lie algebra $\mathrm{Lie}(\widehat{X}_1^q, \ldots, \widehat{X}_m^q)$ is nilpotent. Roughly, we have $\widehat{X}_i^q \approx X_i(q)$, but low-order terms of $X_i(q)$ are not taken into account for defining $\widehat{X}_i^q$, so that the high-order brackets of the $\widehat{X}_i^q$ vanish (which is not generally the case for the X_i). These nilpotent approximations are good local approximations of the vector fields $X_1, \ldots, X_m$, and their study is much simpler.

The proof of Proposition 17 splits into two steps: first, we show that it is sufficient to prove the result in the nilpotent case (Section 3.2), then we handle this simpler case (Section 3.3).

3.1. ***Nilpotent approximation.*** In this section, we recall the construction of the nilpotent approximations $\widehat{X}_1^q, \ldots, \widehat{X}_m^q$. The definitions we give are classical, and the reader can refer to [Agrachev et al. 2020, Chapter 10; Jean 2014, Chapter 2] for more material on this section. This construction is related to the notion of tangent space in the Gromov–Hausdorff sense of a sub-Riemannian structure $(M, \mathcal{D}, g)$ at a point $q \in M$; the tangent space is defined intrinsically (meaning that it does not depend on a choice of coordinates or of local frame) as an equivalence class under the action of sub-Riemannian isometries; see [Bellaïche 1996; Jean 2014].

Sub-Riemannian flag. We define the sub-Riemannian flag as follows: we set $\mathcal{D}^0 = \{0\}$, $\mathcal{D}^1 = \mathcal{D}$, and, for any $j \geqslant 1$, $\mathcal{D}^{j+1} = \mathcal{D}^j + [\mathcal{D}, \mathcal{D}^j]$. For any point $q \in M$, it defines a flag

$$\{0\} = \mathcal{D}_q^0 \subset \mathcal{D}_q^1 \subset \cdots \subset \mathcal{D}_q^{r-1} \subsetneq \mathcal{D}_q^{r(q)} = T_q M.$$

The integer $r(q)$ is called the nonholonomic order of $\mathcal{D}$ at q, and it is equal to 2 everywhere in the Heisenberg manifold for example. Note that it depends on q; see Example 5 in Section 1.2 (the Baouendi–Grushin example).

For $0 \leqslant i \leqslant r(q)$, we set $n_i(q) = \dim \mathcal{D}_q^i$, and the sequence $(n_i(q))_{0 \leqslant i \leqslant r(q)}$ is called the growth vector at point q. We set $\mathcal{Q}(q) = \sum_{i=1}^{r(q)} i(n_i(q) - n_{i-1}(q))$, which is generically the Hausdorff dimension of the metric space given by the sub-Riemannian distance on M; see [Mitchell 1985]. Finally, we define the nondecreasing sequence of weights $w_i(q)$ for $1 \leqslant i \leqslant n$ as follows. Given any $1 \leqslant i \leqslant n$, there exists a unique $1 \leqslant j \leqslant n$ such that $n_{j-1}(q) + 1 \leqslant i \leqslant n_j(q)$. We set $w_i(q) = j$. For example, for any q in the Heisenberg manifold, $w_1(q) = w_2(q) = 1$ and $w_3(q) = 2$; indeed, the coordinates x_1 and x_2 have "weight 1", while the coordinate x_3 has "weight 2" since ∂_{x_3} requires a bracket to be generated.

Regular and singular points. We say that $q \in M$ is regular if the growth vector $(n_i(q'))_{0 \leqslant i \leqslant r(q')}$ at q' is constant for q' in a neighborhood of q. Otherwise, q is said to be singular. If any point $q \in M$ is regular, we say that the structure is equiregular. For example, the Heisenberg manifold is equiregular, but not the Baouendi–Grushin example.

Nonholonomic orders. The nonholonomic order of a smooth germ of function is given by the formula

$$\operatorname{ord}_q(f) = \min\{s \in \mathbb{N} : \text{there exists } i_1, \ldots, i_s \in \{1, \ldots, m\} \text{ such that } (X_{i_1} \cdots X_{i_s} f)(q) \neq 0\},$$

where we adopt the convention that $\min \varnothing = +\infty$.

The nonholonomic order of a smooth germ of vector field X at q, denoted by $\operatorname{ord}_q(X)$, is the real number defined by

$$\operatorname{ord}_q(X) = \sup\{\sigma \in \mathbb{R} : \operatorname{ord}_q(Xf) \geqslant \sigma + \operatorname{ord}_q(f) \text{ for all } f \in C^\infty(q)\}.$$

For example, it holds $\operatorname{ord}_q([X, Y]) \geqslant \operatorname{ord}_q(X) + \operatorname{ord}_q(Y)$ and $\operatorname{ord}_q(fX) \geqslant \operatorname{ord}_q(f) + \operatorname{ord}_q(X)$. As a consequence, every X which has the property that $X(q') \in \mathcal{D}_{q'}^i$ for any q' in a neighborhood of q is of nonholonomic order $\geqslant -i$.

Privileged coordinates. Locally around $q \in M$, it is possible to define a set of so-called "privileged coordinates" of M; see [Bellaïche 1996].

A family $(Z_1, \ldots, Z_n)$ of n vector fields is said to be adapted to the sub-Riemannian flag at q if it is a frame of $T_q M$ at q and if $Z_i(q) \in \mathcal{D}_q^{w_i(q)}$ for any $i \in \{1, \ldots, n\}$. In other words, for any $i \in \{1, \ldots, r(q)\}$, the vectors $Z_1, \ldots, Z_{n_i(q)}$ at q span $\mathcal{D}_q^i$.

A system of privileged coordinates at q is a system of local coordinates $(x_1, \ldots, x_n)$ such that

$$\operatorname{ord}_q(x_i) = w_i \quad \text{for } 1 \leqslant i \leqslant n. \tag{25}$$

In particular, for privileged coordinates, we have $\partial_{x_i} \in \mathcal{D}_q^{w_i(q)} \backslash \mathcal{D}_q^{w_i(q)-1}$ at q, meaning that privileged coordinates are adapted to the flag.

Example: *exponential coordinates of the second kind.* Choose an adapted frame $(Z_1, \dots, Z_n)$ at q. It is proved in [Jean 2014, Appendix B] that the inverse of the local diffeomorphism

$$(x_1, \dots, x_n) \mapsto \exp(x_1 Z_1) \circ \cdots \circ \exp(x_n Z_n)(q)$$

defines privileged coordinates at q, called exponential coordinates of the second kind.

Dilations. We consider a chart of privileged coordinates at q given by a smooth mapping $\psi_q : U \to \mathbb{R}^n$, where U is a neighborhood of q in M, with $\psi_q(q) = 0$. For every $\varepsilon \in \mathbb{R}\setminus\{0\}$, we consider the dilation $\delta_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\delta_\varepsilon(x) = (\varepsilon^{w_i(q)} x_1, \dots, \varepsilon^{w_n(q)} x_n)$$

for every $x = (x_1, \dots, x_n)$. A dilation δ_ε acts also on functions and vector fields on $\mathbb{R}^n$ by pull-back: $\delta_\varepsilon^* f = f \circ \delta_\varepsilon$ and $\delta_\varepsilon^* X$ is the vector field such that $(\delta_\varepsilon^* X)(\delta_\varepsilon^* f) = \delta_\varepsilon^*(Xf)$ for any $f \in C^1(\mathbb{R}^n)$. In particular, for any vector field X of nonholonomic order k, it holds $\delta_\varepsilon^* X = \varepsilon^{-k} X$.

Nilpotent approximation. Fix a system of privileged coordinates $(x_1, \dots, x_n)$ at q. Given a sequence of integers $\alpha = (\alpha_1, \dots, \alpha_n)$, we define the weighted degree of $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ to be

$$w(\alpha) = w_1(q)\alpha_1 + \cdots + w_n(q)\alpha_n.$$

Coming back to the vector fields $X_1, \dots, X_m$, we can write the Taylor expansion

$$X_i(x) \sim \sum_{\alpha, j} a_{\alpha,j} x^\alpha \partial_{x_j}. \tag{26}$$

Since $X_i \in \mathcal{D}$, its nonholonomic order is necessarily -1; hence it holds $w(\alpha) \geqslant w_j(q) - 1$ if $a_{\alpha,j} \neq 0$. Therefore, we may write X_i as a formal series

$$X_i = X_i^{(-1)} + X_i^{(0)} + X_i^{(1)} + \cdots,$$

where $X_i^{(s)}$ is a homogeneous vector field of degree s, meaning that

$$\delta_\varepsilon^* (\psi_q)_* X_i^{(s)} = \varepsilon^s (\psi_q)_* X_i^{(s)}.$$

We set $\widehat{X}_i^q = (\psi_q)_* X_i^{(-1)}$ for $1 \leqslant i \leqslant m$. Then $\widehat{X}_i^q$ is homogeneous of degree -1 with respect to dilations, i.e., $\delta_\varepsilon^* \widehat{X}_i^q = \varepsilon^{-1} \widehat{X}_i^q$ for any $\varepsilon \neq 0$. Each $\widehat{X}_i^q$ may be seen as a vector field on $\mathbb{R}^n$ thanks to the coordinates $(x_1, \dots, x_n)$. Moreover,

$$\widehat{X}_i^q = \lim_{\varepsilon \to 0} \varepsilon \delta_\varepsilon^* (\psi_q)_* X_i$$

in the C^∞ topology; all derivatives uniformly converge on compact subsets. For $\varepsilon > 0$ small enough we have

$$X_i^\varepsilon := \varepsilon \delta_\varepsilon^* (\psi_q)_* X_i = \widehat{X}_i^q + \varepsilon R_i^\varepsilon,$$

where R_i^ε depends smoothly on ε for the C^∞ topology; see also [Agrachev et al. 2020, Lemma 10.58]. An important property is that $(\widehat{X}_1^q, \dots, \widehat{X}_m^q)$ generates a nilpotent Lie algebra of step $r(q)$; see [Jean 2014, Proposition 2.3].

The nilpotent approximation of $X_1, \dots, X_m$ at q is then defined as $\widehat{M}^q \simeq \mathbb{R}^n$ endowed with the vector fields $\widehat{X}_1^q, \dots, \widehat{X}_m^q$. It is important to note that the nilpotent approximation depends on the initial choice of privileged coordinates. For an explicit example of computation of nilpotent approximation; see [Jean 2014, Example 2.8].

3.2. *Reduction to the nilpotent case.* In this section, we show the following.

Lemma 18. *Let $X_1, \dots, X_m$ be smooth vector fields on M satisfying Hörmander's condition, and let $q \in M$. If the property (P) holds at point $0 \in \mathbb{R}^n$ for the nilpotent approximation $\widehat{X}_1^q, \dots, \widehat{X}_m^q$, then the property (P) holds at point q for $X_1, \dots, X_m$.*

Note that the above lemma is true for any nilpotent approximation $\widehat{X}_1^q, \dots, \widehat{X}_m^q$ at q, i.e., for any choice of privileged coordinates (see Section 3.1).

Proof of Lemma 18. We use the notation h_Z for the momentum map associated with the vector field Z (see Section 1.4). We use the notation of Section 3.1, in particular the coordinate chart ψ_q.

We set $Y_i = (\psi_q)_* X_i$ and $X_i^\varepsilon = \varepsilon \delta_\varepsilon^* Y_i$ which is a vector field on $\mathbb{R}^n$. Recall that

$$X_i^\varepsilon = \widehat{X}_i^q + \varepsilon R_i^\varepsilon,$$

where R_i^ε depends smoothly on ε for the C^∞ topology. Therefore, using the homogeneity of $\widehat{X}_i^q$, we get, for any $\varepsilon > 0$,

$$Y_i = \frac{1}{\varepsilon}(\delta_\varepsilon)_* X_i^\varepsilon = \frac{1}{\varepsilon}(\delta_\varepsilon)_*(\widehat{X}_i^q + \varepsilon R_i^\varepsilon) = \widehat{X}_i^q + (\delta_\varepsilon)_* R_i^\varepsilon. \tag{27}$$

The vector field $(\delta_\varepsilon)_* R_i^\varepsilon(x)$ does not depend on ε and has a size which tends uniformly to 0 as $x \to 0 \in \widehat{M}^q \simeq \mathbb{R}^n$. Recall that the Hamiltonian $\widehat{H}$ associated to the vector fields $\widehat{X}_i^q$ is given by

$$\widehat{H} = \sum_{i=1}^m h^2_{\widehat{X}_i^q}.$$

Similarly, we set

$$H = \sum_{i=1}^m h^2_{Y_i}.$$

We note that (27) gives

$$h_{Y_i} = h_{\widehat{X}_i^q} + h_{(\delta_\varepsilon)_* R_i^\varepsilon}.$$

Hence

$$\vec{H} = 2\sum_{i=1}^m h_{Y_i} \vec{h}_{Y_i} = \vec{\widehat{H}} + \vec{\Theta}, \tag{28}$$

where $\vec{\Theta}$ is a smooth vector field on $T^*\mathbb{R}^n$ such that

$$\|(d\pi \circ \vec{\Theta})(x, \xi)\| \leqslant C\|x\| \tag{29}$$

when $\|x\| \to 0$ (independently of ξ), where $\pi : T^*\mathbb{R}^n \to \mathbb{R}^n$ is the canonical projection. This last point comes from the smooth dependence of R_i^ε on ε for the C^∞ topology (uniform convergence of all derivatives on compact subsets of $\mathbb{R}^n$).

Given the projection of an integral curve $c(\cdot)$ of $\vec{H}$, we denote by $\hat{c}(\cdot)$ the projection of the integral curve of $\vec{\widehat{H}}$ with same initial covector. Combining (28) and (29), and using Gronwall's lemma, we obtain the following result:

Fix $T_0 > 0$. For any neighborhood V of 0 in $\mathbb{R}^n$,

there exists another neighborhood V' of 0 such that if $c_{|[0,T_0]} \subset V'$, then $\hat{c}_{|[0,T_0]} \subset V$.

Therefore, if the property (P) holds at $0 \in \mathbb{R}^n$ for $\widehat{X}_1^q, \dots, \widehat{X}_m^q$, then it holds also at $0 \in \mathbb{R}^n$ for the vector fields $Y_1, \dots, Y_m$.

Using that $X_i = \psi_q^* Y_i$, we can pull back the result to M and obtain that the property (P) holds at point q for $X_1, \dots, X_m$, which concludes the proof of Proposition 17. $\square$

Thanks to Lemma 18, it is sufficient to prove the property (P) under the additional assumption that

$$M \subset \mathbb{R}^n \text{ and } \mathrm{Lie}(X_1, \dots, X_m) \text{ is nilpotent.} \tag{30}$$

In all that follows, we assume that this is the case.

3.3. ***End of the proof of Proposition 17.*** Let us finish the proof of Proposition 17. Our ideas are inspired by [Agrachev and Gauthier 2001, Section 6].

First step: *reduction to the constant Goh matrix case.* We consider an adapted frame $Y_1, \dots, Y_n$ at q. We take exponential coordinates of the second kind at q; we consider the inverse ψ_q of the diffeomorphism

$$(x_1, \dots, x_n) \mapsto \exp(x_1 Y_1) \cdots \exp(x_n Y_n)(q).$$

Then we write the Taylor expansion (26) of $X_1, \dots, X_m$ in these coordinates. Thanks to Lemma 18, we can assume that all terms in these Taylor expansions have nonholonomic order -1. We denote by ξ_i the dual variable of x_i. We use the notation $n_1, n_2, \dots$ introduced in Section 3.1, and we make a strong use of (25).

Claim 1. *If a normal geodesic $(x(t), \xi(t))_{t\in\mathbb{R}}$ has initial momentum satisfying $\xi_k(0) = 0$ for any $k \geqslant n_2+1$, then $\dot{\xi}_k \equiv 0$ for any $k \geqslant n_1 + 1$, and in particular $\xi_k \equiv 0$ for any $k \geqslant n_2 + 1$.*

Proof. We write

$$X_j(x) = \sum_{i=1}^{n} a_{ij}(x)\partial_{x_i}, \qquad j = 1, \dots, m,$$

where the a_{ij} are homogeneous polynomials. We have

$$g^*(x, \xi) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ij}(x)\xi_i \right)^2. \tag{31}$$

Let $k \geqslant n_2+1$, which means that x_k has nonholonomic order $\geqslant 3$. If $a_{ij}(x)$ depends on x_k, then necessarily $i \geqslant n_3 + 1$, since $a_{ij}(x)\partial_{x_i}$ has nonholonomic order -1. Thus, writing explicitly $\dot{\xi}_k = -\partial g^*/\partial x_k$ thanks to (31), there is in front of each term a factor ξ_i for some i which is in particular $\geqslant n_2 + 1$. By Cauchy uniqueness, we deduce that $\xi_k \equiv 0$ for any $k \geqslant n_2 + 1$.

Now, let $k \geqslant n_1+1$, which means that x_k has nonholonomic order $\geqslant 2$. If $a_{ij}(x)$ depends on x_k, then necessarily $i \geqslant n_2+1$, since $a_{ij}(x)\partial_{x_i}$ has nonholonomic order -1. Thus, writing explicitly $\dot{\xi}_k = -\partial g^*/\partial x_k$ thanks to (31), there is in front of each term a factor ξ_i for some i which is $\geqslant n_2+1$. It is null by the previous conclusion; hence $\dot{\xi}_k \equiv 0$. □

The previous claim will help us to reduce the complexity of the vector fields X_i once again (after the first reduction provided by Lemma 18). Let us consider, for any $1 \leqslant j \leqslant m$, the vector field

$$X_j^{\mathrm{red}} = \sum_{i=1}^{n_2} a_{ij}(x)\partial_{x_i}, \tag{32}$$

where the sum is taken only up to n_2. We also consider the reduced Hamiltonian on T^*M

$$g^*_{\mathrm{red}} = \sum_{j=1}^{m} h^2_{X_j^{\mathrm{red}}}.$$

Claim 2. *If $X_1^{\mathrm{red}}, \dots, X_m^{\mathrm{red}}$ satisfy property (P) at q, then $X_1, \dots, X_m$ satisfy property (P) at q.*

Proof. Let us assume that $X_1^{\mathrm{red}}, \dots, X_m^{\mathrm{red}}$ satisfy property (P) at q. Let $T_0 > 0$ and let $(x^{\mathrm{red},\varepsilon}(0), \xi^{\mathrm{red},\varepsilon}(0))$ be initial data for the Hamiltonian system associated to g^*_{red} which yield speed-1 normal geodesics $(x^{\mathrm{red},\varepsilon}(t), \xi^{\mathrm{red},\varepsilon}(t))$ such that $x^{\mathrm{red},\varepsilon}(t) \to q$ uniformly over $(0, T_0)$ as $\varepsilon \to 0$.

We can assume without loss of generality that $\xi_i^{\mathrm{red},\varepsilon}(0) = 0$ for any $i \geqslant n_2+1$, since these momenta (preserved under the reduced Hamiltonian evolution) do not change the projection $x^{\mathrm{red},\varepsilon}(t)$ of the normal geodesic. We consider $(x^\varepsilon(0), \xi^\varepsilon(0)) = (x^{\mathrm{red},\varepsilon}(0), \xi^{\mathrm{red},\varepsilon}(0))$ as initial data for the (nonreduced) Hamiltonian evolution associated to g^*. Then we notice that $\xi_k^\varepsilon \equiv 0$ for $k \geqslant n_2+1$ thanks to Claim 1. It follows that when $i \leqslant n_2$, we have $x_i^\varepsilon(t) = x_i^{\mathrm{red},\varepsilon}(t)$; i.e., the coordinate x_i is the same for the reduced and the nonreduced Hamiltonian evolution.

Finally, we take k such that $n_2+1 \leqslant k \leqslant n_3$. Since g^* is given by (31), we have

$$\dot{x}_k^\varepsilon = \frac{\partial g^*}{\partial \xi_k} = 2\sum_{j=1}^{m} a_{kj}(x^\varepsilon)\left(\sum_{i=1}^{n} a_{ij}(x^\varepsilon)\xi_i^\varepsilon\right). \tag{33}$$

But a_{kj} has necessarily nonholonomic order 2 since ∂_{x_k} has nonholonomic order -3. Thus, $a_{kj}(x)$ is a nonconstant homogeneous polynomial in $x_1, \dots, x_{n_2}$. Since $x_1^\varepsilon, \dots, x_{n_2}^\varepsilon$ converge to q uniformly over $(0, T_0)$ as $\varepsilon \to 0$, it is also the case of x_k^ε according to (33), noticing that

$$\left|\sum_{i=1}^{n} a_{ij}(x^\varepsilon)\xi_i^\varepsilon\right| \leqslant (g^*)^{\frac{1}{2}} = \frac{1}{2}$$

for any j. In other words, $x_{n_2+1}^\varepsilon, \dots, x_{n_3}^\varepsilon$ also converge to q uniformly over $(0, T_0)$ as $\varepsilon \to 0$.

We can repeat this argument successively for $k \in \{n_3+1, \dots, n_4\}$, $k \in \{n_4+1, \dots, n_5\}$, etc., and we finally obtain the result: for any $1 \leqslant k \leqslant n$, x_k^ε converges to q uniformly over $(0, T_0)$ as $\varepsilon \to 0$. □

Thanks to the previous claim, we are now reduced to proving Proposition 17 for the vector fields $X_1^{\mathrm{red}}, \dots, X_m^{\mathrm{red}}$. In order to keep notation as simple as possible, we simplify to $X_1, \dots, X_m$; i.e., we drop the upper notation "red". Also, without loss of generality we assume that $q = 0$.

If we choose our normal geodesics so that $x(0)=0$, then $x_i \equiv 0$ for any $i \geqslant n_2+1$ thanks to (32). In other words, we forget the coordinates $x_{n_2+1}, \ldots, x_n$ in the sequel, since they all vanish.[2]

Second step: conclusion of the proof. Now, we write the normal extremal system in its "control" form. We refer the reader to [Agrachev et al. 2020, Chapter 4]. We have

$$\dot{x}(t) = \sum_{i=1}^{m} u_i(t) X_i(x(t)), \tag{34}$$

where the u_i are the controls, explicitly given by

$$u_i(t) = 2h_{X_i}(x(t), \xi(t)) \tag{35}$$

since $(x(t), \xi(t)) = e^{t\vec{g}^*}(0, \xi_0)$. Thanks to (32), we rewrite (34) as

$$\dot{x}(t) = F(x(t))u(t), \tag{36}$$

where $F = (a_{ij})$, which has size $n_2 \times m$, and $u = {}^t(u_1, \ldots, u_m)$. Differentiating (35), we have the complementary equation

$$\dot{u}(t) = G(x(t), \xi(t))u(t),$$

where G is the Goh matrix

$$G = (2\{h_{X_j}, h_{X_i}\})_{1\leqslant i,j\leqslant m}$$

(it differs from the usual Goh matrix by a factor -2 due to the absence of factor $\frac{1}{2}$ in the Hamiltonian g^* in our notation).

Let us prove that $G(t)$ is constant in t. Fix $1 \leqslant j, j' \leqslant m$. We notice that in (32), a_{ij} is a constant (independent of x) as soon as $1 \leqslant i \leqslant n_1$ since ∂_{x_i} has weight -1. This implies

$$[X_j, X_{j'}] \text{ is spanned by the vector fields } \partial_{x_{n_1+1}}, \partial_{x_{n_1+2}}, \ldots, \partial_{x_{n_2}}. \tag{37}$$

Putting this into the relation $\{h_{X_j}, h_{X_{j'}}\} = h_{[X_j, X_{j'}]}$, and using that the dual variables ξ_k for $n_1+1 \leqslant k \leqslant n_2$ are preserved under the Hamiltonian evolution (due to Claim 1), we get that $G(t) \equiv G$ is constant in t.

We know that $G \neq 0$ and that G is antisymmetric. The whole control space $\mathbb{R}^m$ is the direct sum of the image of G and the kernel of G, and G is nondegenerate on its image. We take u_0 in an invariant plane of G; in other words its projection on the kernel of G vanishes (see Remark 19). We denote by $\widetilde{G}$ the restriction of G to this invariant plane. We also assume that u_0, decomposed as $u_0 = (u_{01}, \ldots, u_{0m}) \in \mathbb{R}^m$, satisfies $\sum_{i=1}^m u_{0i}^2 = \frac{1}{4}$. Then $u(t) = e^{t\widetilde{G}}u_0$ and since $e^{t\widetilde{G}}$ is an orthogonal matrix, we have $\|e^{t\widetilde{G}}u_0\| = \|u_0\|$. We have by integration by parts

$$x(t) = \int_0^t F(x(s))e^{s\widetilde{G}}u_0\,ds = F(x(t))\widetilde{G}^{-1}(e^{t\widetilde{G}} - I)u_0 - \int_0^t \frac{d}{ds}(F(x(s)))\widetilde{G}^{-1}(e^{s\widetilde{G}} - I)u_0\,ds. \tag{38}$$

[2] Note that this is the case only because we are now working with the reduced Hamiltonian evolution; otherwise, under the original Hamiltonian evolution associated to (31), the x_i (for $i \geqslant n_2+1$) remain small according to Claim 2, but do not necessarily vanish.

Let us now choose the initial data of our family of normal geodesics (indexed by ε). The starting point $x^\varepsilon(0)=0$ is the same for any ε; we only have to specify the initial covectors $\xi^\varepsilon=\xi^\varepsilon(0)\in T_0^*\mathbb{R}^m$. For any $i=1,\dots,m$, we impose that

$$\langle \xi^\varepsilon, X_i\rangle = u_{0i}. \tag{39}$$

It follows that $g^*(x(0),\xi^\varepsilon(0))=\sum_{i=1}^m u_{0i}^2=\frac{1}{4}$ for any $\varepsilon>0$. Now, we notice that $\mathrm{Span}(X_1,\dots,X_m)$ is in direct sum with the Span of the $[X_i,X_j]$ for i,j running over $1,\dots,m$ (this follows from (37)). Fixing $G^0\neq 0$ an antisymmetric matrix and $\widetilde{G}^0$ its restriction to an invariant plane, we can specify, simultaneously to (39), that

$$\langle \xi^\varepsilon, 2[X_j,X_i]\rangle = \varepsilon^{-1}G^0_{ij}.$$

Then $x^\varepsilon(t)$ is given by (38) applied with $\widetilde{G}=\varepsilon^{-1}\widetilde{G}^0$, which brings a factor ε in front of (38).

Recall finally that the coefficients a_{ij} which compose F have nonholonomic order 0 or 1; thus they are degree-1 (or constant) homogeneous polynomials in $x_1,\dots,x_{n_1}$. Thus $\frac{d}{ds}(F(x(s)))$ is a linear combination of $\dot{x}_i(s)$ which we can rewrite thanks to (36) as a combination with bounded coefficients $\left(\text{since } \sum_{i=1}^m u_i^2=\frac{1}{4}\right)$ of the $x_i(s)$. Hence, applying the Gronwall lemma in (38), we get $\|x^\varepsilon(t)\|\leqslant C\varepsilon$, which concludes the proof.

Remark 19. Let us explain why we choose u_0 to be in an invariant plane of G. If the projection of u_0 to the kernel of G is nonzero then the primitive of the exponential of $e^{(t/\varepsilon)G_0}u_0$ contains a linear term that does not depend on ε. Then the corresponding trajectory follows a singular curve; see [Agrachev et al. 2020, Chapter 4] for a definition. This means we find normal geodesics which spiral around a singular curve and do not remain close to their initial point over $(0,T_0)$, although their initial covector is "high in the cylinder bundle U^*M". For example, for the Hamiltonian $\xi_1^2+(\xi_2+x_1^2\xi_3)^2$ associated to the "Martinet" vector fields $X_1=\partial_{x_1}$, $X_2=\partial_{x_2}+x_1^2\partial_{x_3}$ in $\mathbb{R}^3$, there exist normal geodesics which spiral around the singular curve $(t,0,0)$.

Remark 20. The normal geodesics constructed above lose their optimality quickly, in the sense that their first conjugate point and their cut-point are close to q.

4. Proofs

4.1. *Proof of Theorem 2.*

In this section, we conclude the proof of Theorem 2.

Fix a point q in the interior of $M\setminus\omega$ and $1\leqslant i,j\leqslant m$ such that $[X_i,X_j](q)\notin\mathcal{D}_q$. Fix also an open neighborhood V of q in M such that $V\subset M\setminus\omega$. Fix V' an open neighborhood of q in M such that $\overline{V}'\subset V$, and fix also $T_0>0$.

As already explained in Section 1.3, to conclude the proof of Theorem 2, we use Proposition 16 applied to the particular normal geodesics constructed in Proposition 17.

By Proposition 17, we know that there exists a normal geodesic $t\mapsto x(t)$ such that $x(t)\in V'$ for any $t\in(0,T_0)$. It is the projection of a bicharacteristic $(x(t),\xi(t))$ and since it is nonstationary and travels at speed 1, it holds $g^*(x(t),\xi(t))=\frac{1}{4}$. We denote by $(u_k)_{k\in\mathbb{N}}$ a sequence of solutions of (12) as in

Proposition 16 whose energy at time t concentrates on $x(t)$ for $t \in (0, T_0)$. Because of (22), we know that

$$\|(u_k(0), \partial_t u_k(0))\|_{\mathcal{H}\times L^2} \geqslant c > 0$$

uniformly in k.

Therefore, in order to establish Theorem 2, it is sufficient to show that

$$\int_0^{T_0}\int_\omega |\partial_t u_k(t,x)|^2 \, d\mu(x)\, dt \xrightarrow[k\to+\infty]{} 0. \tag{40}$$

Since $x(t) \in V'$ for any $t \in (0, T_0)$, we get that for V_t chosen sufficiently small for any $t \in (0, T_0)$, the inclusion $V_t \subset V$ holds (see Proposition 16 for the definition of V_t). Combining this last remark with (23), we get (40), which concludes the proof of Theorem 2.

4.2. *Proof of Corollary 4.* We endow the topological dual $\mathcal{H}(M)'$ with the norm

$$\|v\|_{\mathcal{H}(M)'} = \|(-\Delta)^{-1/2} v\|_{L^2(M)}.$$

The following proposition is standard; see, e.g., [Tucsnak and Weiss 2009; Le Rousseau et al. 2017].

Lemma 21. *Let $T_0 > 0$ and $\omega \subset M$ be a measurable set. Then the following two observability properties are equivalent*:

(P1) *There exists C_{T_0} such that, for any $(v_0, v_1) \in D((-\Delta)^{1/2}) \times L^2(M)$, the solution*

$$v \in C^0(0, T_0; D((-\Delta)^{\frac{1}{2}})) \cap C^1(0, T_0; L^2(M))$$

of (1) satisfies

$$\int_0^{T_0}\int_\omega |\partial_t v(t,q)|^2 \, d\mu(q)\, dt \geqslant C_{T_0} \|(v_0, v_1)\|_{\mathcal{H}(M)\times L^2(M)}. \tag{41}$$

(P2) *There exists C_{T_0} such that, for any $(v_0, v_1) \in L^2(M) \times D((-\Delta)^{-1/2})$, the solution*

$$v \in C^0(0, T_0; L^2(M)) \cap C^1(0, T_0; D((-\Delta)^{-\frac{1}{2}}))$$

of (1) satisfies

$$\int_0^{T_0}\int_\omega |v(t,q)|^2 \, d\mu(q)\, dt \geqslant C_{T_0} \|(v_0, v_1)\|^2_{L^2\times\mathcal{H}(M)'}. \tag{42}$$

Proof. Let us assume that (P2) holds. Let u be a solution of (1) with initial conditions $(u_0, u_1) \in D((-\Delta)^{1/2}) \times L^2(M)$. We set $v = \partial_t u$, which is a solution of (1) with initial data $v_{|t=0} = u_1 \in L^2(M)$ and $\partial_t v_{|t=0} = \Delta u_0 \in D((-\Delta)^{-1/2})$. Since $\|(v_0, v_1)\|_{L^2\times\mathcal{H}(M)'} = \|(u_1, \Delta u_0)\|_{L^2\times\mathcal{H}(M)'} = \|(u_0, u_1)\|_{\mathcal{H}(M)\times L^2}$, applying the observability inequality (42) to $v = \partial_t u$, we obtain (41). The proof of the other implication is similar. □

Finally, using Theorem 2, Lemma 21 and the standard HUM method [Lions 1988], we get Corollary 4.

4.3. ***Proof of Theorem 11.*** We consider the space of functions $u \in C^\infty([0,T] \times M_H)$ such that

$$\int_{M_H} u(t,\cdot)\,d\mu = 0$$

for any $t \in [0,T]$, and we denote by $\mathcal{H}_T$ its completion for the norm $\|\cdot\|_{\mathcal{H}_T}$ induced by the scalar product

$$(u,v)_{\mathcal{H}_T} = \int_0^T \int_{M_H} (\partial_t u \partial_t v + (X_1 u)(X_1 v) + (X_2 u)(X_2 v))\,d\mu\,dt.$$

We consider also the topological dual $\mathcal{H}_0'$ of the space $\mathcal{H}_0$ (see Section 1.5).

Lemma 22. *The injections $\mathcal{H}_0 \hookrightarrow L^2(M_H)$, $L^2(M_H) \hookrightarrow \mathcal{H}_0'$ and $\mathcal{H}_T \hookrightarrow L^2((0,T)\times M_H)$ are compact.*

Proof. Let $(\varphi_k)_{k\in\mathbb{N}}$ be an orthonormal basis of real eigenfunctions of $L^2(M_H)$, labeled with increasing eigenvalues $0 = \lambda_0 < \lambda_1 \leqslant \cdots \leqslant \lambda_k \to +\infty$, so that $-\Delta_H \varphi_k = \lambda_k \varphi_k$. The fact that $\lambda_1 > 0$, which will be used in the sequel, can be proved as follows: If $-\Delta_H \varphi = 0$ then $\int_{M_H}((X_1\varphi)^2 + (X_2\varphi)^2)\,d\mu = 0$ and, since $\varphi \in C^\infty(M_H)$ by hypoelliptic regularity, we get $X_1\varphi(x) = X_2\varphi(x) = 0$ for any $x \in M_H$. Hence, $[X_1, X_2]\varphi \equiv 0$, and all together, this proves that φ is constant; thus $\lambda_1 > 0$.

We prove the last injection. Let $u \in \mathcal{H}_T$. Writing $u(t,\cdot) = \sum_{k=1}^\infty a_k(t)\varphi_k(\cdot)$ (note that there is no 0-mode since $u(t,\cdot)$ has null average), we see that

$$\|u\|^2_{\mathcal{H}_T} \geqslant (-\Delta_H u, u)_{L^2((0,T)\times M_H)} = \sum_{k=1}^\infty \lambda_k \|a_k\|^2_{L^2((0,T))} \geqslant \lambda_1 \sum_{k=1}^\infty \|a_k\|^2_{L^2((0,T))} = \lambda_1 \|u\|^2_{L^2((0,T)\times M_H)},$$

and thus $\mathcal{H}_T$ embeds continuously into $L^2((0,T)\times M_H)$. Then, using a classical subelliptic estimate (see [Hörmander 1967; Rothschild and Stein 1976, Theorem 17]), we know that there exists $C > 0$ such that

$$\|u\|_{H^{1/2}((0,T)\times M_H)} \leqslant C(\|u\|_{L^2((0,T)\times M_H)} + \|u\|_{\mathcal{H}_T}).$$

Together with the previous estimate, we obtain that, for any $u \in \mathcal{H}_T$, $\|u\|_{H^{1/2}((0,T)\times M_H)} \leqslant C\|u\|_{\mathcal{H}_T}$. Then, the result follows from the fact that the injection $H^{1/2}((0,T)\times M_H) \hookrightarrow L^2((0,T)\times M_H)$ is compact.

The proof of the compact injection $\mathcal{H}_0 \hookrightarrow L^2(M_H)$ is similar, and the compact injection $L^2(M_H) \hookrightarrow \mathcal{H}_0'$ follows by duality. □

Proof of Theorem 11. In this proof, we use the notation $P = \partial_{tt}^2 - \Delta_H$. For the sake of a contradiction, suppose that there exists a sequence $(u^k)_{k\in\mathbb{N}}$ of solutions of the wave equation such that $\|(u_0^k, u_1^k)\|_{\mathcal{H}\times L^2} = 1$ for any $k \in \mathbb{N}$ and

$$\|(u_0^k, u_1^k)\|_{L^2\times\mathcal{H}_0'} \to 0, \qquad \int_0^T |(\mathrm{Op}(a)\partial_t u^k, \partial_t u^k)_{L^2(M_H,\mu)}|\,dt \to 0 \tag{43}$$

as $k \to +\infty$. Following the strategy of [Tartar 1990; Gérard 1991], our goal is to associate a defect measure to the sequence $(u^k)_{k\in\mathbb{N}}$. Since the functional spaces involved in our result are unusual, we give the argument in detail.

First, up to extraction of a subsequence which we omit, (u_0^k, u_1^k) converges weakly in $\mathcal{H}_0 \times L^2(M_H)$ and, using the first convergence in (43) and the compact embedding $\mathcal{H}_0 \times L^2(M_H) \hookrightarrow L^2(M_H) \times \mathcal{H}_0'$, we get that $(u_0^k, u_1^k) \rightharpoonup 0$ in $\mathcal{H}_0 \times L_0^2$. Using the continuity of the solution with respect to the initial data, we obtain that $u^k \rightharpoonup 0$ weakly in $\mathcal{H}_T$. Using Lemma 22, we obtain $u^k \to 0$ strongly in $L^2((0,T)\times M_H)$.

Fix $B \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$. We have

$$\begin{aligned}(Bu^k, u^k)_{\mathcal{H}_T} &= \int_0^T\int_{M_H} \big((\partial_t Bu^k)(\partial_t u^k) + (X_1Bu^k)(X_1u^k) + (X_2Bu^k)(X_2u^k)\big)\, d\mu(q)\, dt \\ &= \int_0^T\int_{M_H} \big(([\partial_t, B]u^k)(\partial_t u^k) + ([X_1,B]u^k)(X_1u^k) + ([X_2,B]u^k)(X_2u^k)\big)\, d\mu(q)\, dt \\ &\qquad + \int_0^T\int_{M_H} \big((B\partial_t u^k)(\partial_t u^k) + (BX_1u^k)(X_1u^k) + (BX_2u^k)(X_2u^k)\big)\, d\mu(q)\, dt. \end{aligned} \tag{44}$$

Since $[\partial_t, B] \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$, $[X_j, B] \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$ and $u^k \to 0$ strongly in $L^2((0,T)\times M_H)$, the first of the two lines in (44) converges to 0 as $k \to +\infty$. Moreover, the last line is bounded uniformly in k since $B \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$. Hence $(Bu^k, u^k)_{\mathcal{H}_T}$ is uniformly bounded. By a standard diagonal extraction argument (see [Gérard 1991] for example), there exists a subsequence, which we still denote by $(u^k)_{k\in\mathbb{N}}$ such that (Bu^k, u^k) converges for any B of principal symbol b in a countable dense subset of $C_c^\infty((0,T)\times M_H)$. Moreover, the limit only depends on the principal symbol b, and not on the full symbol.

Let us now prove that

$$\liminf_{k\to+\infty} (Bu^k, u^k)_{\mathcal{H}_T} \geqslant 0 \tag{45}$$

when $b \geqslant 0$. With a bracket argument as in (44), we see that it is equivalent to proving that the liminf as $k \to +\infty$ of the quantity

$$Q_k(B) = (B\partial_t u^k, \partial_t u^k)_{L^2} + (BX_1u^k, X_1u^k)_{L^2} + (BX_2u^k, X_2u^k)_{L^2} \tag{46}$$

is $\geqslant 0$. But there exists $B' \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$ such that $B' - B \in \Psi^{-1}_{\mathrm{phg}}((0,T)\times M_H)$ and B' is positive (this is the so-called Friedrichs quantization, see for example [Taylor 1974, Chapter VII]). Then, $\liminf_{k\to+\infty} Q_k(B') \geqslant 0$, and $Q_k(B'-B) \to 0$ since $(B'-B)\partial_t \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$ and $u^k \to 0$ strongly in $L^2((0,T)\times M_H)$. It immediately implies that (45) holds.

Therefore, setting $p = \sigma_p(P)$ and denoting by $\mathcal{C}(p)$ the characteristic manifold $\mathcal{C}(p) = \{p=0\}$, there exists a nonnegative Radon measure ν on $S^*(\mathcal{C}(p)) = \mathcal{C}(p)/(0,+\infty)$ such that

$$(\mathrm{Op}(b)u^k, u^k)_{\mathcal{H}_T} \to \int_{S^*(\mathcal{C}(p))} b\, d\nu$$

for any $b \in S^0_{\mathrm{phg}}((0,T)\times M_H)$.

Let $C \in \Psi^{-1}_{\mathrm{phg}}((0,T)\times M_H)$ of principal symbol c. We have $\vec{p}c = \{p, c\} \in S^0_{\mathrm{phg}}((0,T)\times M_H)$ and, for any $k \in \mathbb{N}$,

$$((CP - PC)u^k, u^k)_{\mathcal{H}_T} = (CPu^k, u^k)_{\mathcal{H}_T} - (Cu^k, Pu^k)_{\mathcal{H}_T} = 0 \tag{47}$$

since $Pu^k = 0$. To be fully rigorous, the identity of the previous line, which holds for any solution $u \in \mathcal{H}_T$ of the wave equation, is first proved for smooth initial data since $Pu \notin \mathcal{H}_T$ in general, and then extended to general solutions $u \in \mathcal{H}_T$. Taking principal symbols in (47), we get $\langle \nu, \vec{p}c\rangle = 0$.

Therefore, denoting by $(\psi_s)_{s\in\mathbb{R}}$ the maximal solutions of

$$\frac{d}{ds}\psi_s(\rho) = \vec{p}(\psi_s(\rho)), \qquad \rho \in T^*(\mathbb{R}\times M_H)$$

(see (7)), we get that, for any $s \in (0, T)$,

$$0 = \langle \nu, \vec{p}c \circ \psi_s\rangle = \left\langle \nu, \frac{d}{ds} c \circ \psi_s \right\rangle = \frac{d}{ds}\langle \nu, c \circ \psi_s\rangle$$

and hence

$$\langle \nu, c\rangle = \langle \nu, c \circ \psi_s\rangle. \tag{48}$$

We note here that the precise homogeneity of c (namely $c \in S^{-1}_{\text{phg}}((0, T)\times M_H)$) does not matter since ν is a measure on the sphere bundle $S^*(\mathcal{C}(p))$. The identity (48) means that ν is invariant under the flow $\vec{p}$.

From the second convergence in (43), we can deduce that

$$\nu = 0 \quad \text{in } S^*(\mathcal{C}(p)) \cap T^*((0, T)\times \operatorname{Supp}(a)). \tag{49}$$

The proof of this fact, which is standard (see for example [Burq and Sun 2022, Section 6.2]), is given in Appendix C.

Let us prove that any normal geodesic of M_H with momentum $\xi \in V_\varepsilon^c$ enters ω in time at most $\kappa\varepsilon^{-1}$ for some $\kappa > 0$, which does not depend on ε. Indeed, the solutions of the bicharacteristic equations (10) with $g^* = \frac{1}{4}$ and $\xi_3 \neq 0$ are given by

$$x_1(t) = \frac{1}{2\xi_3}\cos(2\xi_3 t + \phi) + \frac{\xi_2}{\xi_3}, \qquad x_2(t) = B - \frac{1}{2\xi_3}\sin(2\xi_3 t + \phi),$$
$$x_3(t) = C + \frac{t}{4\xi_3} + \frac{1}{16\xi_3^2}\sin(2(2\xi_3 t + \phi)) + \frac{\xi_2}{2\xi_3^2}\sin(2\xi_3 t + \phi),$$

where B, C, ξ_2, ξ_3 are constants. Since $\xi \in V_\varepsilon^c$ and $g^* = \frac{1}{4}$, it holds

$$\frac{1}{4|\xi_3|} \geqslant \frac{\varepsilon}{2}.$$

Hence, we can conclude using the expression for x_3 (whose derivative is roughly $(4|\xi_3|)^{-1}$) and the fact that $\omega = M_H \backslash B$ contains a horizontal strip. Note that if $\xi_3 = 0$, the expressions of $x_1(t), x_2(t), x_3(t)$ are much simpler and we can conclude similarly.

Hence, together with (49), the propagation property (48) implies that $\nu \equiv 0$. It follows that $\|u^k\|_{\mathcal{H}_T} \to 0$. By conservation of energy, it is a contradiction with the normalization $\|(u_0^k, u_1^k)\|_{\mathcal{H}\times L^2} = 1$. Hence, (11) holds. □

Appendix A: Pseudodifferential calculus

We denote by Ω an open set of a d-dimensional manifold (typically $d = n$ or $d = n+1$ with the notation of this paper) equipped with a smooth volume μ. We denote by q the variable in Ω, typically $q = x$ or $q = (t, x)$ with our notation.

Let $\omega_0 = dp \wedge dq$ be the canonical symplectic form on $T^*\Omega$ written in canonical coordinates (q, p). The Hamiltonian vector field $\vec{f}$ of a function $f \in C^\infty(T^*\Omega)$ is defined by the relation

$$\omega_0(\vec{f}, \cdot) = -df(\cdot).$$

In the coordinates (q, p), it reads

$$\vec{f} = \sum_{j=1}^{d} (\partial_{p_j} f)\partial_{q_j} - (\partial_{q_j} f)\partial_{p_j}.$$

In these coordinates, the Poisson bracket is

$$\{f, g\} = \omega_0(\vec{f}, \vec{g}) = \sum_{j=1}^{d} (\partial_{p_j} f)(\partial_{q_j} g) - (\partial_{q_j} f)(\partial_{p_j} g),$$

which is also equal to $\vec{f} g$ and $-\vec{g} f$.

Let $\pi : T^*\Omega \to \Omega$ be the canonical projection. We recall briefly some facts concerning pseudodifferential calculus, following [Hörmander 1985, Chapter 18].

We denote by $S^m_{\text{hom}}(T^*\Omega)$ the set of homogeneous symbols of degree m with compact support in Ω. We also write $S^m_{\text{phg}}(T^*\Omega)$ for the set of polyhomogeneous symbols of degree m with compact support in Ω. Hence, $a \in S^m_{\text{phg}}(T^*\Omega)$ if $a \in C^\infty(T^*\Omega)$, $\pi(\text{Supp}(a))$ is a compact of Ω, and there exist $a_j \in S^{m-j}_{\text{hom}}(T^*\Omega)$ such that, for all $N \in \mathbb{N}$, $a - \sum_{j=0}^{N} a_j \in S^{m-N-1}_{\text{phg}}(T^*\Omega)$. We denote by $\Psi^m_{\text{phg}}(T^*\Omega)$ the space of polyhomogeneous pseudodifferential operators of order m on Ω, with a compactly supported kernel in $\Omega \times \Omega$. For $A \in \Psi^m_{\text{phg}}(\Omega)$, we denote by $\sigma_p(A) \in S^m_{\text{phg}}(T^*\Omega)$ the principal symbol of A. The subprincipal symbol is characterized by the action of pseudodifferential operators on oscillating functions: if $A \in \Psi^m_{\text{phg}}(\Omega)$ and $f(q) = b(q)e^{ikS(q)}$ with b, S smooth and real-valued, then

$$\int_\Omega A(f)\bar{f}\, d\mu = k^m \int_\Omega \Big(\sigma_p(A)(q, S'(q)) + \frac{1}{k}\sigma_{\text{sub}}(A)(q, S'(q))\Big)|f(q)|^2\, d\mu(q) + O(k^{m-2}).$$

A quantization is a continuous linear mapping

$$\text{Op} : S^m_{\text{phg}}(T^*\Omega) \to \Psi^m_{\text{phg}}(\Omega)$$

satisfying $\sigma_p(\text{Op}(a)) = a$. An example of quantization is obtained by using partitions of unity and, locally, the Weyl quantization, which is given in local coordinates by

$$\text{Op}^W(a) f(q) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d_{q'} \times \mathbb{R}^d_p} e^{i\langle q-q', p\rangle} a\Big(\frac{q+q'}{2}, p\Big) f(q')\, dq'\, dp.$$

We have the following properties:

(1) If $A \in \Psi^l_{\text{phg}}(\Omega)$ and $B \in \Psi^m_{\text{phg}}(\Omega)$, then $[A, B] \in \Psi^{l+m-1}_{\text{phg}}(\Omega)$ and $\sigma_p([A, B]) = \frac{1}{i}\{\sigma_p(a), \sigma_p(b)\}$.

(2) If X is a vector field on Ω and X^* is its formal adjoint in $L^2(\Omega, \mu)$, then $X^*X \in \Psi^2_{\text{phg}}(\Omega)$, $\sigma_p(X^*X) = h_X^2$ and $\sigma_{\text{sub}}(X^*X) = 0$.

(3) If $A \in \Psi^m_{\text{phg}}(\Omega)$, then A maps continuously the space $H^s(\Omega)$ to the space $H^{s-m}(\Omega)$.

Appendix B: Proof of Proposition 12

In this appendix, we give a second proof of Proposition 12 written in a more elementary form than the one of Section 2.1. Let us first prove the result when $M \subset \mathbb{R}^n$, following the proof of [Ralston 1982]. The general case is addressed at the end of this section.

As in the proof of Section 2.1, we suppress the time variable t. Thus we use $x = (x_0, x_1, \dots, x_n)$, where $x_0 = t$. Similarly, $\xi = (\xi_0, \xi_1, \dots, \xi_n)$, where $\xi_0 = \tau$ previously. Let Γ be the curve given by $x(s) \in \mathbb{R}^{n+1}$. We insist on the fact that in the proof the bicharacteristics are parametrized by s, as in (7). We consider functions of the form

$$v_k(x) = k^{\frac{n}{4}-1} a_0(x) e^{ik\psi(x)}.$$

We would like to choose $\psi(x)$ such that for all $s \in \mathbb{R}$, $\psi(x(s))$ is real-valued and

$$\operatorname{Im} \frac{\partial^2 \psi}{\partial x_i \partial x_j}(x(s))$$

is positive definite on vectors orthogonal to $\dot{x}(s)$. Roughly speaking, $|e^{ik\psi(x)}|$ will then look like a Gaussian distribution on planes perpendicular to Γ in $\mathbb{R}^{n+1}$.

We first observe that $\partial^2_{tt} v_k - \Delta v_k$ can be decomposed as

$$\partial^2_{tt} v_k - \Delta v_k = (k^{\frac{n}{4}+1} A_1 + k^{\frac{n}{4}} A_2 + k^{\frac{n}{4}-1} A_3) e^{ik\psi}, \tag{50}$$

with

$$A_1(x) = p_2(x, \nabla\psi(x)) a_0(x), \quad A_2(x) = L a_0(x), \quad A_3(x) = \partial^2_{tt} a_0(x) - \Delta a_0(x).$$

Here we have set

$$L a_0 = \frac{1}{i} \sum_{j=0}^{n} \frac{\partial p_2}{\partial \xi_j}(x, \nabla\psi(x)) \frac{\partial a_0}{\partial x_j} + \frac{1}{2i} \left(\sum_{j,k=0}^{n} \frac{\partial^2 p_2}{\partial \xi_j \partial \xi_k}(x, \nabla\psi(x)) \frac{\partial^2 \psi}{\partial x_j \partial x_k} \right) a_0. \tag{51}$$

(For general strictly hyperbolic operators, L contains a term with the subprincipal symbol of the operator, but here it is null; see Appendix A.)

In what follows, we construct a_0 and ψ so that $A_1(x)$ vanishes at order 2 along Γ and $A_2(x)$ vanishes at order 0 along the same curve. We will then be able to use Lemma 14 with $S = 3$ and $S = 1$ respectively.

Analysis of $A_1(x)$. Our goal is to show that, if we choose ψ adequately, we can make the quantity

$$f(x) = p_2(x, \nabla\psi(x)) \tag{52}$$

vanish at order 2 on Γ. For the vanishing at order 0, we prescribe that ψ satisfies $\nabla\psi(x(s)) = \xi(s)$, and then $f(x(s)) = 0$ since $(x(s), \xi(s))$ is a null-bicharacteristic. Note that this is possible since $x(s) \neq x(s')$ for any $s \neq s'$, due to $\dot{x}_0(s) = 1$ (bicharacteristics are traveled at speed 1; see Section 1.4). For the vanishing at order 1, using (52) and (7), we remark that, for any $0 \leqslant j \leqslant n$,

$$\begin{aligned}\frac{\partial f}{\partial x_j}(x(s)) &= \frac{\partial p_2}{\partial x_j}(x(s)) + \sum_{k=0}^{n} \frac{\partial p_2}{\partial \xi_k}(x(s)) \frac{\partial \psi}{\partial x_j \partial x_k}(x(s)) \\ &= -\dot{\xi}_j(s) + \sum_{k=0}^{n} \dot{x}_k(s) \frac{\partial \psi}{\partial x_j \partial x_k}(x(s)) \\ &= -\frac{d}{ds}\left(\frac{\partial \psi}{\partial x_j}(x(s))\right) + \sum_{k=0}^{n} \dot{x}_k(s) \frac{\partial \psi}{\partial x_j \partial x_k}(x(s)) = 0.\end{aligned} \tag{53}$$

Therefore, f vanishes automatically at order 1 along Γ (without making any particular choice for ψ): it just follows from (52) and the bicharacteristic equations (7). But for $f(x)$ to vanish at order 2 along Γ, it is required to choose a particular ψ. In the end, we will find that if ψ is given by the formula (59) below, with M being a solution of (54), then f vanishes at order 2 along Γ. Let us explain why.

Using the Einstein summation notation, we want that, for any $0 \leqslant i, j \leqslant n$, it holds

$$0 = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 p_2}{\partial x_j \partial x_i} + \frac{\partial^2 p_2}{\partial \xi_k \partial x_i} \frac{\partial^2 \psi}{\partial x_j \partial x_k} + \frac{\partial^2 p_2}{\partial x_j \partial \xi_k} \frac{\partial^2 \psi}{\partial x_i \partial x_k} + \frac{\partial^2 p_2}{\partial \xi_l \partial \xi_k} \frac{\partial^2 \psi}{\partial x_i \partial x_k} \frac{\partial^2 \psi}{\partial x_j \partial x_l} + \frac{\partial p_2}{\partial \xi_k} \frac{\partial^3 \psi}{\partial x_j \partial x_k \partial x_i}$$

along Γ. Introducing the matrices

$$\begin{aligned}(M(s))_{ij} &= \frac{\partial^2 \psi}{\partial x_i \partial x_j}(x(s)), & (A(s))_{ij} &= \frac{\partial^2 p_2}{\partial x_i \partial x_j}(x(s), \xi(s)), \\ (B(s))_{ij} &= \frac{\partial^2 p_2}{\partial \xi_i \partial x_j}(x(s), \xi(s)), & (C(s))_{ij} &= \frac{\partial^2 p_2}{\partial \xi_i \partial \xi_j}(x(s), \xi(s)),\end{aligned}$$

this amounts to solving the matricial Riccati equation

$$\frac{dM}{ds} + MCM + B^T M + MB + A = 0 \tag{54}$$

on a finite-length time interval. While solving (54), we also require $M(s)$ to be symmetric, $\mathrm{Im}(M(s))$ to be positive definite on the orthogonal complement of $\dot{x}(s)$, and $M(s)\dot{x}(s) = \dot{\xi}(s)$ to hold for all s due to (53).

Let M_0 be a symmetric $(n+1) \times (n+1)$ matrix with $\mathrm{Im}(M_0) > 0$ on the orthogonal complement of $\dot{x}(0)$ and $M_0\dot{x}(0) = \dot{\xi}(0)$ (in particular $\mathrm{Im}(M_0)\dot{x}(0) = 0$). It is shown in [Ralston 1982] that there exists a global solution $M(s)$ on $[0, T]$ of (54) which satisfies all the above conditions and such that $M(0) = M_0$. The proof just requires that A, C are symmetric, but does not need anything special about p_2 (in particular, it applies to our sub-Riemannian case where p_2 is degenerate). For the sake of completeness, we recall the proof here.

We consider $(Y(s), N(s))$ the matrix solution with initial data $(Y(0), N(0)) = (\mathrm{Id}, M_0)$ (where Id is the $(n+1) \times (n+1)$ identity matrix) to the linear system

$$\begin{cases}\dot{Y} = BY + CN, \\ \dot{N} = -AY - B^T N.\end{cases} \tag{55}$$

We note that $(Y(s)\dot{x}(0), N(s)\dot{x}(0))$ then also solves (55), with Y and N being this time vectorial. One can check that $(\dot{x}(s), \dot{\xi}(s))$ is the solution of the same linear system with same initial data, and therefore, for any $s \in \mathbb{R}$,

$$\dot{x}(s) = Y(s)\dot{x}(0), \quad \dot{\xi}(s) = N(s)\dot{x}(0). \tag{56}$$

All the coefficients in (55) are real and A and C are symmetric, and it follows that the flow defined by (55) *on vectors* preserves both the real symplectic form acting on pairs $(y, \eta) \in (\mathbb{R}^{n+1})^2$ and $(y', \eta') \in (\mathbb{R}^{n+1})^2$ given by

$$\sigma((y, \eta), (y', \eta')) = y \cdot \eta' - \eta \cdot y'$$

and the complexified form $\sigma_{\mathbb{C}}((y, \eta), (y', \eta')) = \sigma((y, \eta), (\bar{y}', \bar{\eta}'))$ for $(y, \eta) \in (\mathbb{C}^{n+1})^2$ and $(y', \eta') \in (\mathbb{C}^{n+1})^2$. When we say that $\sigma_{\mathbb{C}}$ is invariant under (55), it means that we allow complex vectorial initial data in (55).

Let us prove that $Y(s)$ is invertible for any s. Let $v \in \mathbb{C}^{n+1}$ and $s_0 \in \mathbb{R}$ be such that $Y(s_0)v = 0$. We set $y(s_0) = Y(s_0)v$ and $\eta(s_0) = N(s_0)v$ and consider $\chi(s_0) = (y(s_0), \eta(s_0))$. From the conservation of $\sigma_{\mathbb{C}}$, we get

$$0 = \sigma_{\mathbb{C}}(\chi(s_0), \chi(s_0)) = \sigma_{\mathbb{C}}(\chi(0), \chi(0)) = v \cdot \overline{M_0 v} - \bar{v} \cdot M_0 v = -2i\bar{v} \cdot (\mathrm{Im}(M_0))v.$$

Since $\mathrm{Im}(M_0)$ is positive definite on the orthogonal complement to $\dot{x}(0)$, it holds $v = \lambda \dot{x}(0)$ for some $\lambda \in \mathbb{C}$. Hence

$$0 = Y(s_0)v = \lambda Y(s_0)\dot{x}(0) = \lambda \dot{x}(s_0),$$

where the last equality comes from (56). Since $\dot{x}_0(s_0) = (\partial p_2/\partial \xi_0)(s_0) = -2\xi_0(s_0) = 1$, it holds $\dot{x}(s_0) \neq 0$; hence $\lambda = 0$. It follows that $v = 0$ and $Y(s_0)$ is invertible.

Now, for any $s \in \mathbb{R}$, we set

$$M(s) = N(s)Y(s)^{-1},$$

which is a solution of (54) with $M(0) = M_0$. It satisfies $M(s)\dot{x}(s) = \dot{\xi}(s)$ thanks to (56). Moreover, it is symmetric: if we denote by $y^i(s)$ and $\eta^i(s)$ the column vectors of Y and N, by preservation of σ, for any $0 \leqslant i, j \leqslant n$, the quantity

$$\sigma((y^i(s), \eta^i(s)), (y^j(s), \eta^j(s))) = y^i(s) \cdot M(s)y^j(s) - y^j(s) \cdot M(s)y^i(s)$$

is equal to the same quantity at $s = 0$, which is equal to 0 since M_0 is symmetric.

Let us finally prove that, for any $s \in \mathbb{R}$, $\mathrm{Im}(M(s))$ is positive definite on the orthogonal complement of $\dot{x}(s)$. Let $y(s_0) \in \mathbb{C}^{n+1}$ be in the orthogonal complement of $\dot{x}(s_0)$. We decompose $y(s_0)$ on the column vectors of $Y(s_0)$:

$$y(s_0) = \sum_{i=0}^{n} b_i y^i(s_0), \quad b_i \in \mathbb{C}.$$

For $s \in \mathbb{R}$, we consider $y(s) = \sum_{i=0}^{n} b_i y^i(s)$ and we set $\chi(s) = \sum_{i=0}^{n} b_i (y^i(s), \eta^i(s))$. Then,

$$\sigma_{\mathbb{C}}(\chi(s), \chi(s)) = -2i\overline{y(s)} \cdot \mathrm{Im}(M(s))y(s). \tag{57}$$

By preservation of $\sigma_{\mathbb{C}}$ and using (57), we get that

$$\overline{y(s_0)} \cdot \mathrm{Im}(M(s_0))y(s_0) = \overline{y(0)} \cdot \mathrm{Im}(M_0)y(0). \tag{58}$$

But $y(0)$ cannot be proportional to $\dot{x}(0)$; otherwise, using (56), we would get that $y(s_0)$ is proportional to $\dot{x}(s_0)$. Hence, the right-hand side in (58) is > 0, which implies that $\mathrm{Im}(M(s_0))$ is positive definite on the orthogonal complement to $\dot{x}(s_0)$.

Therefore, we found a choice for the second-order derivatives of ψ along Γ which meets all our conditions. For $x = (t, x') \in \mathbb{R} \times \mathbb{R}^n$ and s such that $t = t(s)$, we set

$$\psi(x) = \xi'(s) \cdot (x' - x'(s)) + \tfrac{1}{2}(x' - x'(s)) \cdot M(s)(x' - x'(s)), \tag{59}$$

and f vanishes at order 2 along Γ for this choice of ψ.

To sum up, as in the Riemannian (or "strictly hyperbolic") case handled in [Ralston 1982], the key observation is that the invariance of σ and $\sigma_{\mathbb{C}}$ prevents the solutions of (54) with positive imaginary part on the orthogonal complement of $\dot{x}(0)$ from blowing up.

Analysis of $A_2(x)$. We note that A_2 vanishes along Γ if and only if $La_0(x(s)) = 0$. According to (51), this turns out to be a linear transport equation on $a_0(x(s))$. Moreover, the coefficient of the first-order term, namely $\nabla_\xi p_2(x(s), \xi(s))$, is different from 0. Therefore, given $a_0 \neq 0$ at $(t{=}0, x{=}x(0))$, this transport equation has a solution $a_0(x(s))$ with initial datum a_0, and, by Cauchy uniqueness, $a_0(x(s)) \neq 0$ for any s. Note that we have prescribed a_0 only along Γ, and we may choose a_0 in a smooth (and arbitrary) way outside Γ. We choose it to vanish outside a small neighborhood of Γ.

Proof of (13). We use (50) and we apply Lemma 14 to $S = 3$, $c = A_1$ and to $S = 1$, $c = A_2$, and we get

$$\|\partial_{tt}^2 v_k - \Delta v_k\|_{L^1(0,T;L^2(M))} \leqslant C(k^{-\frac{1}{2}} + k^{-\frac{1}{2}} + k^{-1}),$$

which implies (13).

Proof of (14). We first observe that since $\mathrm{Im}(M(s))$ is positive definite on the orthogonal complement of $\dot{x}(s)$ and continuous as a function of s, there exist $\alpha, C > 0$ such that, for any $t(s) \in [0, T]$ and any $x' \in M$,

$$|\partial_t v_k(t(s), x')|^2 + \sum_{j=1}^{m} |X_j v_k(t(s), x')|^2 \geqslant (C|a_0(t(s), x')|^2 k^{\frac{n}{2}} + O(k^{2(\frac{n}{2}-1)}))e^{-\alpha k d(x', x'(s))^2},$$

where $d(\cdot\,, \cdot)$ denotes the Euclidean distance in $\mathbb{R}^n$. We denote by ℓ_n the Lebesgue measure on $\mathbb{R}^n$. Using the observation that, for any function f,

$$\int_M f(x')e^{-\alpha k d(x', x'(s))^2}\, d\mu(x') \sim \frac{\pi^{n/2}}{k^{n/2}\sqrt{\alpha}} f(x'(s)) \frac{d\mu}{d\ell_n}(x'(s)) \tag{60}$$

as $k \to +\infty$, and the fact that $a_0(x(s)) \neq 0$, we obtain (14).

***Proof of* (15).** We observe that since $\mathrm{Im}(M(s))$ is positive definite (uniformly in s) on the orthogonal complement of $\dot{x}(s)$, there exist $C, \alpha' > 0$ such that, for any $t \in [0, T]$, for any $x' \in M$, $|\partial_t v_k(t(s), x')|$ and $|X_j v_k(t(s), x')|$ are both bounded above by $Ck^{n/4} e^{-\alpha' k d(x', x'(s))^2}$. Therefore

$$\begin{aligned}\int_{M \setminus V_{t(s)}} \Big(|\partial_t v_k(t(s), x')|^2 + \sum_{j=1}^m |X_j v_k(t(s), x')|^2 \Big)\, d\mu(x') \\ \leqslant Ck^{n/2} \int_{M \setminus V_{t(s)}} e^{-2\alpha' k d(x', x'(s))^2}\, d\mu(x') \\ \leqslant Ck^{n/2} \int_{M \setminus V_{t(s)}} e^{-2\alpha' k d(x', x'(s))^2}\, d\ell_n(x') + \mathrm{o}(1), \qquad (61)\end{aligned}$$

where, in the last line, we used the fact that $|d\mu/d\ell_n| \leqslant C$ in a fixed compact subset of M (since μ is a smooth volume), and the $\mathrm{o}(1)$ comes from the eventual blowup of μ at the boundary of M.

Now, $M \subset \mathbb{R}^n$, and there exists $r > 0$ such that $B_d(x(s), r) \subset V_{t(s)}$ for any s such that $t(s) \in (0, T)$, where $d(\cdot, \cdot)$ still denotes the Euclidean distance in $\mathbb{R}^n$. Therefore, we bound above the integral in (61) by

$$Ck^{n/2} \int_{\mathbb{R}^n \setminus B_d(x(s), r)} e^{-2\alpha' k d(x', x'(s))^2}\, d\ell_n(x'). \qquad (62)$$

Making the change of variables $y = k^{-1/2}(y - x(s))$, we can bound (62) from above by

$$C \int_{\mathbb{R}^n \setminus B_d(0, rk^{1/2})} e^{-2\alpha' \|y\|^2}\, d\ell_n(y),$$

with $\|\cdot\|$ the Euclidean norm. This last expression is bounded above by

$$Ce^{-\alpha' r^2 k} \int_{\mathbb{R}^n} e^{-\alpha' \|y\|^2}\, d\ell_n(y),$$

which implies (15).

Extension of the result to any manifold M. In the case of a general manifold M, not necessarily included in $\mathbb{R}^n$, we use charts together with the above construction. We cover M by a set of charts $(U_\alpha, \varphi_\alpha)$, where (U_α) is a family of open sets of M covering M and $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ is an homeomorphism U_α onto an open subset of $\mathbb{R}^n$. Take a solution $(x(t), \xi(t))_{t \in [0,T]}$ of (8). It visits a finite number of charts in the order $U_{\alpha_1}, U_{\alpha_2}, \ldots,$ and we choose the charts and a_0 so that $v_k(t, \cdot)$ is supported in a unique chart at each time t. The above construction shows how to construct a_0 and ψ as long as $x(t)$ remains in the same chart. For any $l \geqslant 1$, we choose t_l so that $x(t_l) \in U_{\alpha_l} \cap U_{\alpha_{l+1}}$ and $a_0(t_l, \cdot)$ is supported in $U_{\alpha_l} \cap U_{\alpha_{l+1}}$. Since there is a (local) solution v_k for any choice of initial $a_0(t_l, x(t_l))$ and $\mathrm{Im}(\partial^2 \psi/(\partial x_i\, \partial x_j))(t_l, x(t_l))$ in Proposition 12, we see that v_k may be continued from the chart U_{α_l} to the chart $U_{\alpha_{l+1}}$. This continuation is smooth since the two solutions coincide as long as $a_0(t, \cdot)$ is supported in $U_{\alpha_l} \cap U_{\alpha_{l+1}}$. Patching all solutions on the time intervals $[t_l, t_{l+1}]$ together, it yields a global-in-time solution v_k, as desired.

Appendix C: Proof of (49)

Because of the second convergence in (43) and the nonnegativity of a, it amounts to proving that

$$(X_1 \operatorname{Op}(a) u^k, X_1 u^k)_{L^2((0,T) \times M_H)} + (X_2 \operatorname{Op}(a) u^k, X_2 u^k)_{L^2((0,T) \times M_H)} \to 0.$$

Now, we notice that for any $B \in \Psi^0_{\mathrm{phg}}((0,T)\times M_H)$, it holds

$$(Bu^k, X_1u^k)_{L^2((0,T)\times M_H)} \xrightarrow[k\to+\infty]{} 0 \quad \text{and} \quad (Bu^k, \partial_t u^k)_{L^2((0,T)\times M_H)} \xrightarrow[k\to+\infty]{} 0 \tag{63}$$

since $u^k \to 0$ strongly in $L^2((0,T)\times M_H)$ and both X_1u^k and $\partial_t u^k$ are bounded in $L^2((0,T)\times M_H)$. We apply this to $B = [X_1, \mathrm{Op}(a)]$, and then, also using (63), we see that we can replace $\mathrm{Op}(a)$ by its Friedrichs quantization $\mathrm{Op}^F(a)$, which is positive; see [Taylor 1974, Chapter VII]. In other words, we are reduced to proving

$$(\mathrm{Op}^F(a)X_1u^k, X_1u^k)_{L^2((0,T)\times M_H)} + (\mathrm{Op}^F(a)X_2u^k, X_2u^k)_{L^2((0,T)\times M_H)} \xrightarrow[k\to+\infty]{} 0. \tag{64}$$

Let $\delta > 0$ and $\tilde{a} \in S^0_{\mathrm{phg}}((-\delta, T+\delta)\times M_H)$, $0 \leqslant \tilde{a} \leqslant \sup(a)$, and such that $\tilde{a}(t,\cdot) = a(\cdot)$ for $0 \leqslant t \leqslant T$. Making repeated use of (63) and of integrations by parts (since $\tilde{a}$ is compactly supported in time), we have

$$\begin{aligned}
\sum_{j=1}^{2} (\mathrm{Op}^F(\tilde{a})X_ju^k, X_ju^k)_{L^2((0,T)\times M_H)} &= \sum_{j=1}^{2} (X_j\,\mathrm{Op}^F(\tilde{a})u^k, X_ju^k)_{L^2((0,T)\times M_H)} + o(1)\\
&= -(\mathrm{Op}^F(\tilde{a})u^k, \Delta u^k)_{L^2((0,T)\times M_H)} + \mathrm{o}(1)\\
&= -(\mathrm{Op}^F(\tilde{a})u^k, \partial_t^2 u^k)_{L^2((0,T)\times M_H)} + \mathrm{o}(1)\\
&= (\partial_t\,\mathrm{Op}^F(\tilde{a})u^k, \partial_t u^k)_{L^2((0,T)\times M_H)} + \mathrm{o}(1)\\
&= (\mathrm{Op}^F(\tilde{a})\partial_t u^k, \partial_t u^k)_{L^2((0,T)\times M_H)} + \mathrm{o}(1).
\end{aligned}$$

Finally we note that since Op^F is a positive quantization, we have

$$\begin{aligned}
\sum_{j=1}^{2} (\mathrm{Op}^F(a)X_ju^k, X_ju^k)_{L^2((0,T)\times M_H)} &\leqslant \sum_{j=1}^{2} (\mathrm{Op}^F(\tilde{a})X_ju^k, X_ju^k)_{L^2((0,T)\times M_H)}\\
&= (\mathrm{Op}^F(\tilde{a})\partial_t u^k, \partial_t u^k)_{L^2((0,T)\times M_H)} + \mathrm{o}(1)\\
&\leqslant C\delta + (\mathrm{Op}^F(a)\partial_t u^k, \partial_t u^k)_{L^2((0,T)\times M_H)} + \mathrm{o}(1)\\
&\leqslant C\delta + o(1),
\end{aligned}$$

where C does not depend on δ. Taking $\delta \to 0$ concludes the proof of (64), and consequently (49) holds.

Acknowledgments

I warmly thank my PhD advisor Emmanuel Trélat for mentioning this problem to me, for his constant support and his numerous suggestions during the preparation of this paper. Many thanks also to Andrei Agrachev who helped me correct a flaw in the proof of Proposition 17. I thank Yves Colin de Verdière, Luc Hillairet, Armand Koenig, Luca Rizzi, Clotilde Fermanian-Kammerer, Maciej Zworski, Frédéric Jean, Jean-Paul Gauthier, Matthieu Léautaud and Ludovic Sacchelli for interesting discussions related to this problem. Finally, I am very grateful to an anonymous referee whose questions and suggestions allowed me to considerably improve the readability of the present paper. I was partially supported by the grant ANR-15-CE40-0018 of the ANR (project SRGI).

References

[Agrachev and Gauthier 2001] A. Agrachev and J.-P. Gauthier, "On the subanalyticity of Carnot–Carathéodory distances", *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **18**:3 (2001), 359–382. MR Zbl

[Agrachev et al. 2020] A. Agrachev, D. Barilari, and U. Boscain, *A comprehensive introduction to sub-Riemannian geometry: from the Hamiltonian viewpoint*, Cambridge Stud. Adv. Math. **181**, Cambridge Univ. Press, 2020. MR Zbl

[Bardos et al. 1992] C. Bardos, G. Lebeau, and J. Rauch, "Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary", *SIAM J. Control Optim.* **30**:5 (1992), 1024–1065. MR Zbl

[Beauchard and Cannarsa 2017] K. Beauchard and P. Cannarsa, "Heat equation on the Heisenberg group: observability and applications", *J. Differential Equations* **262**:8 (2017), 4475–4521. MR Zbl

[Beauchard et al. 2014] K. Beauchard, P. Cannarsa, and R. Guglielmi, "Null controllability of Grushin-type operators in dimension two", *J. Eur. Math. Soc.* **16**:1 (2014), 67–101. MR Zbl

[Beauchard et al. 2020] K. Beauchard, J. Dardé, and S. Ervedoza, "Minimal time issues for the observability of Grushin-type equations", *Ann. Inst. Fourier (Grenoble)* **70**:1 (2020), 247–312. MR Zbl

[Bellaïche 1996] A. Bellaïche, "The tangent space in sub-Riemannian geometry", pp. 1–78 in *Sub-Riemannian geometry*, edited by A. Bellaïche and J.-J. Risler, Progr. Math. **144**, Birkhäuser, Basel, 1996. MR Zbl

[Burq and Sun 2019] N. Burq and C. Sun, "Time optimal observability for Grushin Schrödinger equation", 2019. To appear in *Anal. PDE*. arXiv 1910.03691

[Burq and Sun 2022] N. Burq and C. Sun, "Time optimal observability for Grushin Schrödinger equation", *Anal. PDE* **15**:6 (2022), 1487–1530. MR Zbl

[Colin de Verdière et al. 2018] Y. Colin de Verdière, L. Hillairet, and E. Trélat, "Spectral asymptotics for sub-Riemannian Laplacians, I: Quantum ergodicity and quantum limits in the 3-dimensional contact case", *Duke Math. J.* **167**:1 (2018), 109–174. MR Zbl

[Duprez and Koenig 2020] M. Duprez and A. Koenig, "Control of the Grushin equation: non-rectangular control region and minimal time", *ESAIM Control Optim. Calc. Var.* **26** (2020), art. id. 3. MR Zbl

[Engel and Nagel 2000] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Grad. Texts in Math. **194**, Springer, 2000. MR Zbl

[Garetto and Ruzhansky 2015] C. Garetto and M. Ruzhansky, "Wave equation for sums of squares on compact Lie groups", *J. Differential Equations* **258**:12 (2015), 4324–4347. MR Zbl

[Gérard 1991] P. Gérard, "Mesures semi-classiques et ondes de Bloch", exposé 16 in *Séminaire sur les Équations aux Dérivées Partielles*, 1990–1991, École Polytech., Palaiseau, France, 1991. MR Zbl

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

[Hörmander 1985] L. Hörmander, *The analysis of linear partial differential operators, III: Pseudodifferential operators*, Grund. der Math. Wissen. **274**, Springer, 1985. MR Zbl

[Ivrii 2019] V. Ivrii, *Microlocal analysis, sharp spectral asymptotics and applications, I: Semiclassical microlocal analysis and local and microlocal semiclassical asymptotics*, Springer, 2019. MR Zbl

[Jean 2014] F. Jean, *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*, Springer, 2014. MR Zbl

[Koenig 2017] A. Koenig, "Non-null-controllability of the Grushin operator in 2D", *C. R. Math. Acad. Sci. Paris* **355**:12 (2017), 1215–1235. MR Zbl

[Lascar 1982] B. Lascar, "Propagation des singularités pour des équations hyperboliques à caractéristique de multiplicité au plus double et singularités masloviennes", *Amer. J. Math.* **104**:2 (1982), 227–285. MR Zbl

[Laurent and Léautaud 2022] C. Laurent and M. Léautaud, *Tunneling estimates and approximate controllability for hypoelliptic equations*, Mem. Amer. Math. Soc. **1357**, Amer. Math. Soc., Providence, RI, 2022. MR Zbl

[Le Rousseau et al. 2017] J. Le Rousseau, G. Lebeau, P. Terpolilli, and E. Trélat, "Geometric control condition for the wave equation with a time-dependent observation domain", *Anal. PDE* **10**:4 (2017), 983–1015. MR Zbl

[Letrouit and Sun 2021] C. Letrouit and C. Sun, "Observability of Baouendi–Grushin-type equations through resolvent estimates", *J. Inst. Math. Jussieu* (online publication June 2021).

[Lions 1988] J.-L. Lions, *Contrôlabilité exacte*, *perturbations et stabilisation de systèmes distribués*, *I*: *Contrôlabilité exacte*, Rech. Math. Appl. **8**, Masson, Paris, 1988. MR Zbl

[Macià and Zuazua 2002] F. Macià and E. Zuazua, "On the lack of observability for wave equations: a Gaussian beam approach", *Asymptot. Anal.* **32**:1 (2002), 1–26. MR Zbl

[Miller 2012] L. Miller, "Resolvent conditions for the control of unitary groups and their approximations", *J. Spectr. Theory* **2**:1 (2012), 1–55. MR Zbl

[Mitchell 1985] J. Mitchell, "On Carnot–Carathéodory metrics", *J. Differential Geom.* **21**:1 (1985), 35–45. MR Zbl

[Montgomery 2002] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Math. Surv. Monogr. **91**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Ralston 1982] J. Ralston, "Gaussian beams and the propagation of singularities", pp. 206–248 in *Studies in partial differential equations*, edited by W. Littman, MAA Stud. Math. **23**, Math. Assoc. Amer., Washington, DC, 1982. MR Zbl

[Rothschild and Stein 1976] L. P. Rothschild and E. M. Stein, "Hypoelliptic differential operators and nilpotent groups", *Acta Math.* **137**:3-4 (1976), 247–320. MR Zbl

[Strichartz 1986] R. S. Strichartz, "Sub-Riemannian geometry", *J. Differential Geom.* **24**:2 (1986), 221–263. MR Zbl

[Tartar 1990] L. Tartar, "H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations", *Proc. Roy. Soc. Edinburgh Sect. A* **115**:3-4 (1990), 193–230. MR Zbl

[Taylor 1974] M. Taylor, *Pseudo differential operators*, Lecture Notes in Math. **416**, Springer, 1974. MR Zbl

[Tucsnak and Weiss 2009] M. Tucsnak and G. Weiss, *Observation and control for operator semigroups*, Birkhäuser, Basel, 2009. MR Zbl

Received 3 Feb 2020. Revised 4 Sep 2021. Accepted 6 Oct 2021.

CYRIL LETROUIT: letrouit@ljll.math.upmc.fr
Sorbonne Université, Université Paris-Diderot, CNRS, Inria, Laboratoire Jacques-Louis Lions, Paris, France

and

Department of Mathematics and Applications, École Normale Supérieure, CNRS, PSL Research University, Paris, France

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 16 No. 3 2023