ANNALES
HENRI LEBESGUE

YVES COLIN DE VERDIÈRE

LUC HILLAIRET

EMMANUEL TRÉLAT

SMALL-TIME ASYMPTOTICS OF HYPOELLIPTIC HEAT KERNELS NEAR THE DIAGONAL, NILPOTENTIZATION AND RELATED RESULTS

ASYMPTOTIQUE EN TEMPS PETIT DE NOYAUX DE LA CHALEUR HYPOELLIPTIQUES PRÈS DE LA DIAGONALE, NILPOTENTISATION ET RÉSUTATS LIÉS

ABSTRACT. — We establish small-time asymptotic expansions for heat kernels of hypoelliptic Hörmander operators in a neighborhood of the diagonal, generalizing former results obtained in particular by Métivier and by Ben Arous. The coefficients of our expansions are identified in terms of the nilpotentization of the underlying sub-Riemannian structure. Our approach is purely analytic and relies in particular on local and global subelliptic estimates as well as on the local nature of small-time asymptotics of heat kernels. The fact that our expansions are valid not only along the diagonal but in an asymptotic neighborhood of the diagonal is the main novelty, useful in view of deriving Weyl laws for subelliptic Laplacians. Incidentally, we establish a number of other results on hypoelliptic heat kernels that are interesting in themselves, such as Kac's principle of not feeling the boundary, asymptotic results for singular perturbations of hypoelliptic operators, global smoothing properties for selfadjoint heat semigroups.

RÉSUMÉ. — Nous établissons le développement asymptotique en temps petit de noyaux de la chaleur d'opérateurs de Hörmander hypoelliptiques au voisinage de la diagonale, généralisant des résultats précédents obtenus en particulier par Métivier et par Ben Arous. Les coefficients dans nos développements sont identifiés en termes de la nilpotentisation de la structure sous-Riemannienne sous-jacente. Notre approche est purement analytique et repose en particulier sur des estimées sous-elliptiques locales et globales ainsi que sur la nature locale des asymptotiques en temps petit des noyaux de la chaleur. Le fait que nos développements soient valides non seulement le long de la diagonale mais aussi dans un voisinage asymptotique de la diagonale est la nouveauté principale ; cela est utile pour établir des lois de Weyl pour des Laplaciens sous-elliptiques. Incidemment, nous établissons un certain nombre de résultats sur les noyaux de la chaleur hypoelliptiques qui sont intéressants en eux-mêmes, comme le principe de Kac de ne pas sentir l'effet du bord, des résultats asymptotiques de perturbation singulière d'opérateurs hypoelliptiques, des propriétés de régularisation globale pour des semi-groupes de la chaleur dans le cas autoadjoint.

CONTENTS

1. Introduction and main result

1.1. Framework

Let n and m be nontrivial integers. Let M be a smooth connected manifold of dimension n. Let $X_0, X_1, \ldots, X_m$ be smooth vector fields on M and let $\mathbb{V}$ be a smooth function (potential) on M that is bounded below. Setting $X = (X_0, X_1, \ldots, X_m)$, we define the Hörmander operator

$$(1.1) \qquad \triangle = \sum_{i=1}^{m} X_i^2 + X_0 - \mathbb{V}$$

where X_i is seen as a derivation operator and $\mathbb{V}$ is the multiplication by the potential. In view of involving the case of magnetic fields, the drift vector field X_0 can even be assumed to take complex values.[1]

Let μ be an arbitrary smooth (Borel) measure on M. We assume that the operator $\triangle$ on $L^2(M, \mu)$, of domain $D(\triangle) \subset L^2(M, \mu)$ encoding some possible boundary conditions whenever M has a boundary, generates a strongly continuous semigroup $(e^{t\triangle})_{t \geqslant 0}$ on $L^2(M, \mu)$ (see Lemma 3.1 in Section 3.1.2 for some sufficient conditions). We denote by $e_{\triangle, \mu}$ the corresponding heat kernel defined on $(0, +\infty) \times M \times M$, associated with the operator $\triangle$ and with the measure μ (see Appendix A).

We set $D = \mathrm{Span}(X_1, \ldots, X_m)$. Under the strong Hörmander condition

$$(1.2) \qquad \mathrm{Lie}_q(D) = \mathrm{Lie}_q(X_1, \ldots, X_m) = T_q M \quad \forall\, q \in M$$

the operator $\triangle$ is subelliptic[2] (see [Hör67]) and the heat kernel $e_{\triangle, \mu}$ is smooth on $(0, +\infty) \times M \times M$.

The objective of this paper is to establish a small-time asymptotic expansion of the heat kernel $e_{\triangle, \mu}$ at any order *near* the diagonal. Our study is in the line of well known results [BA89, Mét76] establishing such expansions *along* the diagonal. The first main novelty here is that our expansion is valid, not only along the diagonal, but in an asymptotic neighborhood of the diagonal. This fact is actually instrumental in view of deriving local and microlocal Weyl laws for general subelliptic Laplacians, which will be done in the forthcoming paper [CdVHT].

The second main novelty is that we identify the functions in the small-time asymptotic expansion of the heat kernel in terms of the so-called nilpotentization of the sub-Riemannian structure (M, D, g) where g is a metric on D defined thanks to the vector fields $X_1, \ldots, X_m$.

Compared with the approach of [BA89] which is probabilistic (Malliavin calculus), our proof (done in Part 2) is purely analytic and relies in particular on local and global subelliptic estimates, on local and global smoothing properties of heat kernels, on the finite propagation speed property for sR waves and on the Kannai transform, and on the local nature of hypoelliptic heat kernels (Kac's principle). Our paper is entirely selfcontained: even for the several tools (such as uniform local subelliptic

---

[1] This requires some obvious slight changes for instance when considering a scalar product.
[2] Actually, the weaker assumption $\mathrm{Lie}_q(X_0, X_1, \ldots, X_m) = T_q M$ for every $q \in M$ (called weak Hörmander condition) is sufficient to ensure subellipticity. The stronger assumption (1.2) is however required to derive our main result.

estimates) that are straightforward extensions of known results, we provide at least a sketch of proof.

Sub-Riemannian structure. Complete reminders on sub-Riemannian (sR) geometry are given in Section 2. Attached with the m-tuple of vector fields $(X_1, \ldots, X_m)$, there is a canonical sR structure (M, D, g), where $D = \mathrm{Span}(X_1, \ldots, X_m)$ and g is a positive definite quadratic form on D. Given any $q \in M$, the sR flag at q is the sequence of nested vector subspaces $\{0\} = D_q^0 \subset D_q = D_q^1 \subset D_q^2 \subset \ldots \subset D_q^{r(q)-1} \subsetneq D_q^{r(q)} = T_q M$ defined in terms of successive Lie brackets, and $r(q)$ is the degree of nonholonomy at q. Setting $n_i(q) = \dim D_q^i$, the integer $\mathcal{Q}(q) = \sum_{i=1}^r i(n_i(q) - n_{i-1}(q))$ coincides with the Hausdorff dimension when q is regular, i.e., when the integers $n_i(\cdot)$ are constant in an open neighborhood of q. The point q is said singular when it is not regular.

The nilpotentization of (M, D, g) at q is the sR structure $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$ defined as the metric tangent space of M (endowed with the sR distance) in the sense of Gromov–Hausdorff. In a local chart of privileged coordinates around q, $\widehat{M}^q$ is a homogeneous space identified to $\mathbb{R}^n$ (with a sR isometry), endowed with dilations $\delta_\varepsilon(x) = (\varepsilon^{w_1(q)} x_1, \ldots, \varepsilon^{w_n(q)} x_n)$, for $\varepsilon \in \mathbb{R}$ and $x \in \mathbb{R}^n$, where $w_i(q) = i(n_i(q) - n_{i-1}(q))$ (sR weights). We have $\widehat{D}^q = \mathrm{Span}(\widehat{X}_1^q, \ldots, \widehat{X}_m^q)$ where $\widehat{X}_i^q$ is the limit of $\varepsilon \delta_\varepsilon^* X_i$ in C^∞ topology as $\varepsilon \to 0$. We also define the (constant) nilpotentized measure $\widehat{\mu}^q$ on $\mathbb{R}^n$ as the limit of $\frac{1}{|\varepsilon|^{\mathcal{Q}(q)}} \delta_\varepsilon^* \mu$ as $\varepsilon \to 0$. Finally, we define the sR Laplacian

$$\widehat{\triangle}^q = \sum_{i=1}^m (\widehat{X}_i^q)^2$$

and we denote by $\widehat{e}^q = e_{\widehat{\triangle}^q, \widehat{\mu}^q}$ the heat kernel generated by $\widehat{\triangle}^q : D(\widehat{\triangle}^q) \to L^2(\widehat{M}^q, \widehat{\mu}^q)$.

C^∞ topology. We recall that the set $C^\infty(\mathbb{R}^n)$ of smooth functions on $\mathbb{R}^n$ is a Montel space, i.e., a Fréchet space enjoying the Heine–Borel property (meaning that closed bounded subsets are compact), for the topology defined by the seminorms $p_{i,j}(f) = \max\{|\partial^\alpha f(x)| \mid |x| \leqslant i, |\alpha| \leqslant j\}$, $(i, j) \in \mathbb{N}^2$. We speak of the $C^\infty(\mathbb{R}^n)$ topology.

Let W be an arbitrary open subset of the manifold M. Covering W with charts, the set $C^\infty(W)$ of smooth functions on W is endowed with the C^∞ topology (in charts), making it a Montel space.

We say that a sequence $(f_k)_{k \in \mathbb{N}}$ in $C^\infty(W)$ converges to 0 in C^∞ topology if f_k converges uniformly to 0 on any compact subset of W, as well as all its derivatives. A sequence $(X_k)_{k \in \mathbb{N}}$ of smooth vector fields on W converges to 0 in C^∞ topology if all its coefficients (in charts) converge to 0 in C^∞ topology.

Throughout, we denote by $C_c^\infty(W)$ the set of smooth functions of compact support on W.

Notations. Throughout the paper, we use in the various estimates the notation $\mathrm{Cst}(\cdot)$, standing for a generic positive constant depending on the parameters indicated in the parenthesis.

The integral of an integrable function f on M with respect to the smooth measure μ is denoted by $\int_M f(q)\, d\mu(q)$.

1.2. Main result

THEOREM A. — *Let $q \in M$ be arbitrary (regular or not). Let $\psi_q : U \to V$ be a chart of privileged coordinates at q such that $\psi_q(q) = 0$, where U is an open connected neighborhood of q in M and V is an open neighborhood of 0 in $\mathbb{R}^n$. We assume that X_0 is a smooth section of D over M. We also assume that $\sup_{q' \in U} r(q') < +\infty$ (this is always satisfied if $\overline{U}$ is compact).*

Then, given any $N \in \mathbb{N}^*$, in the chart[3] we have the asymptotic expansion in $C^\infty((0, +\infty) \times V \times V)$

$$(1.3) \qquad |\varepsilon|^{\mathcal{Q}(q)}\, e_{\triangle,\,\mu}\left(\varepsilon^2\tau, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) = \widehat{e}^q(\tau, x, x') + \sum_{i=1}^{N} \varepsilon^i f_i^q(\tau, x, x') + \mathrm{o}\left(|\varepsilon|^N\right)$$

as $\varepsilon \to 0$, $\varepsilon \neq 0$, where the functions f_i^q are smooth and satisfy the homogeneity property

$$f_i^q(\tau,\, x,\, x') = \varepsilon^{-i}|\varepsilon|^{\mathcal{Q}(q)} f_i^q\left(\varepsilon^2\tau, \delta_\varepsilon(x), \delta_\varepsilon(x')\right)$$

for all $(\tau, x, x') \in (0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ and for every $\varepsilon \neq 0$.

Taking $\tau = 1$, $\varepsilon = \sqrt{t}$ and setting $a_i^q(x, x') = f_i^q(1, x, x')$, it follows that, given any $N \in \mathbb{N}$, in the chart we have the asymptotic expansion in $C^\infty(V \times V)$

$$(1.4) \qquad t^{\mathcal{Q}(q)/2}\, e_{\triangle,\,\mu}\left(t, \delta_{\sqrt{t}}(x), \delta_{\sqrt{t}}(x')\right) = \widehat{e}^q(1, x, x') + \sum_{i=1}^{N} t^{i/2} a_i^q(x, x') + \mathrm{o}\left(t^{N/2}\right)$$

as $t \to 0$, $t > 0$, where the functions a_i^q are smooth and satisfy $a_{2j-1}^q(0,0) = 0$ for every $j \in \mathbb{N}^*$.

Moreover, if q is regular, then the above convergence and asymptotic expansion are also locally uniform with respect to q, and the functions $\widehat{e}^q$, f_i^q and a_i^q depend smoothly (in C^∞ topology) on q in any open neighborhood of q consisting of regular points. If the manifold M is Whitney stratifiable with strata defined according to the sR flag (i.e., the growth vector $(n_1(q), \ldots, n_{r(q)}(q))$ is constant along each stratum) then the latter property is satisfied along strata.

The fact that the asymptotic expansion (1.3) is in $C^\infty((0, +\infty) \times V \times V)$ means that the asymptotics are uniform with respect to (τ, x, x') on any compact subset of $(0, +\infty) \times V \times V$ as well as for all derivatives. In particular, for all $k \in \mathbb{N}$, $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, we have

$$\lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} |\varepsilon|^{\mathcal{Q}(q)+2k}\varepsilon^{\sum_{i=1}^n (\alpha_i+\beta_i)w_i(q)} \left(\partial_1^k \partial_2^\alpha \partial_3^\beta\, e_{\triangle,\,\mu}\right)\left(\varepsilon^2\tau, \delta_\varepsilon(x), \delta_\varepsilon(x')\right)$$

$$= \left(\partial_1^k \partial_2^\alpha \partial_3^\beta\, \widehat{e}^q\right)(\tau, x, x')$$

---

[3] This means that the left-hand side of (1.3) is $|\varepsilon|^{\mathcal{Q}(q)}\, e_{\triangle,\,\mu}(\varepsilon^2\tau, \psi_q^{-1}(\delta_\varepsilon(x)), \psi_q^{-1}(\delta_\varepsilon(x')))$.

uniformly with respect to (τ, x, x') on any compact subset of $(0, +\infty) \times V \times V$. Here, given a function e depending on three variables (τ, y, y'), the notation ∂_1 (resp., ∂_2, ∂_3) denotes the partial derivative with respect to τ (resp., to y, to y').

As a particular case, take $x = x' = 0$ in the expansion (1.4) given in Theorem A and set $c_j(q) = a_{2j}^q(0, 0)$. Since $a_{2j-1}^q(0, 0) = 0$ and $\psi_q^{-1}(0) = q$, we obtain the following corollary.

COROLLARY 1.1. — *Given any* $N \in \mathbb{N}$, *for every* $q \in M$,

$$t^{\mathcal{Q}(q)/2} e_{\triangle, \mu}(t, q, q) = \widehat{e}^q(1, 0, 0) + c_1(q)t + \cdots + c_N(q)t^N + o\left(t^N\right)$$

as $t \to 0$, $t > 0$. *Moreover if* q *is regular then the functions* c_j *are smooth locally around* q.

We thus recover the main result of [BA89] (see also [Mét76]), which is a small-time expansion of the heat kernel along the diagonal. Here, additionally to those well known results, we provide a geometric interpretation of the coefficients of this expansion, in function of the nilpotentization at q: the main coefficient is $\widehat{e}^q(1, 0, 0)$ > 0, but the other coefficients $c_j(q)$ are also given by convolutions of the heat kernel $\widehat{e}^q = e_{\widehat{\triangle}^q, \widehat{\mu}^q}$, as made precise in the proof of the theorem in Part 2 (see in particular Proposition 6.7 in Section 6.2, and see Section 6.3).

Moreover, the expansion stated in Theorem A is established in an asymptotic neighborhood of the diagonal, which is instrumental to derive the microlocal Weyl law for general equiregular sR structures, or to establish local Weyl laws for singular sR structures (see Section 1.3).

Remark 1.2. — In Theorem A, we have assumed that X_0 is a smooth section of D over M. If $X_0(q) \in D_q$ for every $q \in M$ but cannot be written as a combination of the vector fields X_i, $i = 1, \ldots, m$, with smooth functions a_i, then the asymptotics of the heat kernel in small time along the diagonal may degenerate and be exponentially decreasing (see [BAL91a, BAL91b]).

We have the following generalization if X_0 is not a smooth section of D. We assume that X_0 is a smooth section of D^2 over M, meaning that

$$(1.5) \qquad X_0(q) = \sum_{i=1}^m a_i(q)X_i(q) + \sum_{i, j=1}^m b_{ij}(q)[X_i, X_j](q) \quad \forall q \in M$$

where the a_i and b_{ij} are smooth functions on M. In local privileged coordinates around q, it is then possible to write $X_0 = X_0^{(-2)} + X_0^{(-1)} + X_0^{(0)} + \cdots$, where $X_0^{(k)}$ is homogeneous of degree k (see Section 2.5.3). Therefore, defining on $\mathbb{R}^n$ the vector field $\widehat{X}_0^q = X_0^{(-2)}$ (in the local coordinates), which is homogeneous of order -2 (and not of order -1), we have $\varepsilon^2 \delta_\varepsilon^* X_0 = \widehat{X}_0^q + \varepsilon X_0^{(-1)} + \varepsilon^2 X_0^{(0)} + \cdots + \varepsilon^N X_0^{(N-2)} + o(|\varepsilon|^N)$ at any order N, in C^∞ topology. Then Theorem A is still valid provided that $\widehat{\triangle}^q$ be replaced with $\widehat{\triangle}^q + \widehat{X}_0^q$. In other words, under the assumption (1.5), the asymptotics of the heat kernel of $\triangle$ is given by the heat kernel of the operator $\widehat{\triangle}^q + \widehat{X}_0^q$, which is a first-order perturbation of $\widehat{\triangle}^q$, homogeneous of order -2. This is in accordance with results of [BAL91b].

In contrast, if X_0 is not a smooth section of D^2 then the asymptotics in small time may be completely different, and the heat kernel along the diagonal may even decrease exponentially as $\varepsilon \to 0$ (see [BAL91b]).

Remark 1.3. — We have stated Theorem A in terms of the heat kernels $e_{\triangle,\mu}$ and $\widehat{e}^q = e_{\widehat{\triangle}^q, \widehat{\mu}^q}$. But, as explained in Appendix A, the smooth measure $e_{\triangle,\mu}(t, q, q')\, d\mu(q')$ on M does not depend on μ, but only on the operator $\triangle$. The same remark holds for the smooth measure $\widehat{e}^q(t, x, x')\, d\widehat{\mu}^q(x')$ on $\widehat{M}^q \sim \mathbb{R}^n$. It would therefore be more natural to express Theorem A in terms of Schwartz kernels. We keep however the statement in this form, because the concept of heat kernel is familiar and is probably the most standard in the literature.

Anyway, it is useful to note that using Schwartz kernels would avoid nilpotentizations of measures. Moreover, this explains why small-time expansions of heat kernels along the diagonal have no interesting meaning in singular sR cases: what has to be considered, there, is the small-time asymptotics of $e_{\triangle,\mu}(t, q, q)\, d\mu(q)$, which is related to the trace of $e^{t\triangle} f$ (see Appendix A) and does not depend on μ nor on the nilpotentization of μ at the point q.

1.3. Using Theorem A to obtain Weyl laws

In the forthcoming paper [CdVHT], we will establish local and microlocal Weyl laws for sR Laplacians in regular and in singular cases. Let us give a flavor of these results, thus explaining how Theorem A can be used to reach this objective.

Considering the general framework of Section 1.1, we assume here moreover that M is compact without boundary and that $\triangle = \triangle_{sR}$ is selfadjoint (it is a sR Laplacian: see Sections 2.2 and 3.1.1).

Since $\mathrm{Lie}(D) = TM$, the operator $\triangle$ is subelliptic, has a compact resolvent and thus has a discrete spectrum $0 = \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_k \cdots \to +\infty$. Let $(\phi_k)_{k \in \mathbb{N}^*}$ be an orthonormal eigenbasis of $L^2(M, \mu)$ corresponding to these ordered eigenvalues. The spectral counting function is defined by $N(\lambda) = \#\{k \in \mathbb{N}^* \mid \lambda_k \leqslant \lambda\}$ for every $\lambda \in \mathbb{R}$.

Equiregular cases. We prove in [CdVHT] that, if (M, D, g) is equiregular, i.e., if every point of M is regular, then

$$\int_M f(q)\, e_{\triangle,\mu}(t, q, q)\, d\mu(q) = \frac{\int_M f(q)\, \widehat{e}^q(1, 0, 0)\, d\mu(q)}{t^{\mathcal{Q}/2}} + \mathrm{O}\left(\frac{1}{t^{\mathcal{Q}/2-1}}\right)$$

as $t \to 0$, $t > 0$, for every continuous function f on M, where $\mathcal{Q}$ is the Hausdorff dimension of M (local Weyl law). Actually, we establish a small-time expansion at any order, by using the complete statement of Theorem A (asymptotic expansion at any order). In particular, the spectral counting function has the asymptotics

$$N(\lambda) \underset{\lambda \to +\infty}{\sim} \frac{\int_M \widehat{e}^q(1, 0, 0)\, d\mu(q)}{\Gamma(\mathcal{Q}/2 + 1)} \lambda^{\mathcal{Q}/2}.$$

The above limits are an easy consequence of Theorem A along the diagonal (and thus, already follow from [BA89, Mét76]). Indeed, taking $x = x' = 0$ and $\varepsilon = \sqrt{t}$ in Theorem A, we obtain

$$e_{\triangle,\mu}(t, q, q)\, d\mu(q) \underset{\substack{t \to 0 \\ t > 0}}{\sim} \frac{\widehat{e}^{\,q}(1, 0, 0)}{t^{\mathcal{Q}(q)/2}}\, d\mu(q) = \widehat{e}^{\,q}(t, 0, 0)\, d\mu(q) \quad \forall\, q \in M.$$

The result follows by dominated convergence.

In turn, this result puts in evidence an intrinsic sR measure that we call the *Weyl measure*, of which there exists a local and a microlocal version. The *local Weyl measure* $w_\triangle$ is the probability measure on M defined by

$$\int_M f\, dw_\triangle = \lim_{\lambda \to +\infty} \frac{1}{N(\lambda)} \sum_{\lambda_k \leqslant \lambda} \int_M f |\phi_k|^2\, d\mu \quad \forall\, f \in C^0(M)$$

whenever the limit exists for all continuous functions f, i.e., $w_\triangle$ is the weak limit of the sequence of probability measures $\frac{1}{N(\lambda)} \sum_{\lambda_k \leqslant \lambda} |\phi_k|^2\, \mu$ (Cesàro mean) as $\lambda \to +\infty$. The above argument shows that, in the equiregular case, the local Weyl measure exists, does not depend on μ, is a smooth measure on M and its density with respect to μ is

$$dw_\triangle(q) = \frac{\widehat{e}^{\,q}(1, 0, 0)}{\int_M \widehat{e}^{\,q'}(1, 0, 0)\, d\mu(q')}\, d\mu(q).$$

Accordingly, the *microlocal Weyl law* $W_\triangle$ is the probability measure defined on the co-sphere bundle $S^\star M$ by

$$\int_{S^\star M} a\, dW_\triangle = \lim_{\lambda \to +\infty} \frac{1}{N(\lambda)} \sum_{\lambda_k \leqslant \lambda} \langle \mathrm{Op}(a)\phi_k, \phi_k \rangle_{L^2(M, \mu)}$$

for every symbol a of order 0, whenever the limit exists for all symbols a of order 0 (here, Op denotes any quantization operator). We prove in [CdVHT] that, in the equiregular case, the microlocal Weyl law exists and we provide its explicit expression, showing in particular that $W_\triangle$ is supported on $S(D^{r-1})^\perp$ where r is the degree of nonholonomy. This generalizes to equiregular cases a result obtained in [CdVHT18] in the three-dimensional contact case.

Establishing this result instrumentally relies on the fact that, taking $x' = 0$ in Theorem A, we obtain that, for every $q \in M$, in the chart where $\psi_q(q) = 0$,

$$|\varepsilon|^{\mathcal{Q}(q)} e_{\triangle,\mu}\left(\varepsilon^2, \delta_\varepsilon(x), 0\right) d\widehat{\mu}^q(x) \underset{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}}{\sim} \widehat{e}^{\,q}(1, x, 0)\, d\widehat{\mu}^q(x)$$

uniformly with respect to x on any compact subset of $\psi_q(U)$.

Singular cases. Famous singular sR structures are given by the Grushin case in dimension two or by the Martinet case in dimension three. They are chiefly studied in [CdVHT]. One of the main tools is the fact that, taking $x = x'$ in Theorem A, we obtain that, for every $q \in M$, in the chart where $\psi_q(q) = 0$,

$$\lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} |\varepsilon|^{\mathcal{Q}(q)} e_{\triangle,\mu}\left(\varepsilon^2, \delta_\varepsilon(x), \delta_\varepsilon(x)\right) = \widehat{e}^{\,q}(1, x, x)$$

uniformly with respect to x on any compact subset of $\psi_q(U)$. For instance, in the Grushin (resp., Martinet) case the asymptotics of the spectral counting function is

Cst $\lambda \ln \lambda$ (resp., Cst $\lambda^2 \ln \lambda$) as $\lambda \to +\infty$; this fact is known for the Grushin case (see [MS78]) but is new for the Martinet case. We even obtain two-terms asymptotic expansions of the local Weyl law, with intrinsic coefficients, by using the complete statement of Theorem A (asymptotic expansion). In [CdVHT] we prove that, in some sense, the occurrence of a (power of a) logarithm in the spectral counting function is the highest possible complexity. More precisely, we establish that, given any sR structure whose singular set is Whitney stratifiable according to the sR flag with polynomial singularities,

$$\int_M f(q) \, e_{\triangle, \mu}(t, q, q) \, d\mu(q) \sim \text{Cst} \, \frac{\left| \ln^k t \right|}{t^\gamma}$$

as $t \to 0$, $t > 0$, for every continuous function f on M, with $k \in \{0, 1, \ldots, n\}$ and $\gamma \in \mathbb{Q}$ such that $\gamma \geqslant \frac{\mathcal{Q}_{\text{eq}}}{2}$, where $\mathcal{Q}_{\text{eq}}$ is the Hausdorff dimension of the equiregular region of M. Moreover if $k = n$ then $\gamma = \frac{\mathcal{Q}_{\text{eq}}}{2}$. As a consequence the asymptotics of $N(\lambda)$ is Cst $\lambda^\gamma \ln^k \lambda$ as $\lambda \to +\infty$: this gives the maximal complexity, for instance no term $\ln \ln \lambda$ appears in the asymptotics.

1.4. Structure of the paper

The paper is structured as follows.

We provide in Section 2 some reminders in sub-Riemannian (sR) geometry. In particular, we recall the instrumental concept of nilpotentization, much relying on sR dilations that are used in our main result. Also, for the sake of completeness, in Section 3.1.2 we recall sufficient assumptions ensuring existence of the hypoelliptic heat kernel.

The paper is then split into two parts. The reason is the following.

While our main result, Theorem A, states a small-time asymptotic expansion of the heat kernel near the diagonal at any order, obtaining only the limit, i.e., only the first term, is much less difficult than obtaining the complete expansion.

Actually, the mathematical techniques and results that are required to obtain the limit are purely of a *local* nature, and thus, do not require lengthy developments. Since the result is already interesting (in particular, for obtaining Weyl laws), in Part 1 we establish Theorem B, which is Theorem A at the order zero. Part 2 is then devoted to establishing the complete statement of Theorem A, which is surprisingly much more difficult and requires results of a *global* nature, as explained in detail at the beginning of Part 2.

The two parts, as well as the sections therein, are redacted independently enough one from each other, in order to allow several possible levels of reading. For instance, the reader only interested in the limit result may read only Part 1 (and Appendix B.1 on local subelliptic estimates). Besides, a reader, even though not interested in the complete asymptotic expansion, may however find in Part 2 (and in Appendices B.2 and C on global subelliptic estimates) a number of tools of interest to deal with global issues.

Structure of Part 1. This part is concerned with *local* issues, sufficient to prove Theorem B.

Section 3 in Part 1 is devoted to establishing some general facts on hypoelliptic heat kernels:

- In Section 3.1, we gather some remarks on Hörmander operators (intrinsic formula of integration by parts on a domain with boundary, symmetry properties, with Dirichlet or Neumann boundary conditions), useful in order to state a general result for existence of semigroups (Lemma 3.1).
- In Section 3.2, we establish two general results for parameter-dependent hypoelliptic heat kernels:
  - In Section 3.2.1, we prove that the small time asymptotics of hypoelliptic heat kernels is purely local: this reflects the famous Kac's principle of not feeling the boundary. Our version (hypoelliptic Kac's principle, Theorem 3.2) is moreover uniform with respect to parameters.
  - In Section 3.2.2, we give a general result for singular perturbations of hypoelliptic operators (Theorem 3.5): assuming that the Hörmander operator depends continuously on some parameter τ, we prove that the corresponding heat kernel depends as well continuously on τ.

Section 4 is dedicated to proving Theorem B, which corresponds to proving the limit in Theorem A. As said before, this proof only requires to use local tools. As sketched at the beginning of Part 1, the argument starts by applying the Trotter–Kato theorem to the operator $\triangle^\varepsilon = \varepsilon^2 \delta_\varepsilon^* \triangle (\delta_\varepsilon)_*$, which converges to $\widehat{\triangle}^q$ in C^∞ topology, before using uniform local subellipticity (Appendix B.1) to infer strong convergence properties.

Structure of Part 2. This part is devoted to establishing the complete statement of Theorem A. In contrast to the proof of Theorem B, the proof of Theorem A requires *global* smoothing properties.

Since the proof is quite lengthy, in Section 5 we give the idea of our approach to the proof and we list a number of properties that are required. We point out the main difficulties, in order to motivate the developments done in Section 6 and in Appendix.

Theorem A is proved in full detail in Section 6. The proof starts by applying iteratively the Duhamel formula but, as explained in Section 5, getting the complete asymptotic expansion out of it raises serious difficulties and requires global considerations. To facilitate the reading we have organized this section as follows:

- The operator $\triangle^{\varepsilon,\gamma}$, adequate modification of the operator $\triangle^\varepsilon$ with a "damping" parameter $\gamma > 0$ so as to be (uniformly) at most polynomial at infinity, is defined and analyzed in Section 6.1. In particular, the key Lemmas 6.2 and 6.4 state strong global convergence properties of $\triangle^{\varepsilon,\gamma}$ to $\widehat{\triangle}^q$ that are required in our proof (and which do not hold for $\triangle^\varepsilon$ that converges to $\widehat{\triangle}^q$ only in C^∞ topology).
- In Section 6.2, and more precisely, in Proposition 6.7, we prove that $e^{t\triangle^{\varepsilon,\gamma}}$ has a small-time asymptotic expansion with respect to ε at any order, in the sense of (uniformly) smoothing operators. The proof of this proposition

is delicate and uses in an instrumental way, as explained and motivated in Section 5, the global smoothing properties established for $e^{t\widehat{\triangle}^q}$ in Appendix C, and those established for $e^{t\triangle^{\varepsilon,\gamma}}$ in Appendix B.2 thanks to (uniform) global subelliptic estimates.

- Taking Schwartz kernels in Section 6.3, we obtain the asymptotic expansion of the heat kernel.
- The end of the proof, in Section 6.4, consists of applying the localization theorem (hypoelliptic Kac's principle).

As noted in Section 5.3, in the particular case where $M = \mathbb{R}^n$, all vector fields X_i are polynomial and $\triangle$ is selfadjoint, it is not necessary to resort to the modified operator $\triangle^{\varepsilon,\gamma}$, and the global results established in Appendix C are sufficient to achieve the proof of Theorem A.

Structure of Part 3 (Appendix). We have gathered in the appendix the following material.

In Appendix A, we recall some known facts on Schwartz and heat kernels. In particular, we point out the meaningful fact that Schwartz kernels do not depend on the measure while heat kernels do.

In Appendix B, we establish subelliptic estimates and smoothing properties for hypoelliptic heat semigroups:

- Local estimates in Appendix B.1, uniform with respect to some parameters: although these are standard (adding dependence with respect to parameters is straightforward), we give sketches of proofs, in order to prepare the reader to global estimates.
- Global estimates in Appendix B.2, uniform with respect to some parameters, established for parameter-dependent Hörmander operators whose growth at infinity is at most polynomial, satisfying a uniform polynomial Hörmander condition.

In Appendix C, we derive a number of more precise and stronger global smoothing properties in the case where the operator $\triangle$ is a (selfadjoint) sR Laplacian, in Sobolev spaces or iterated domains with polynomial weight. Our arguments are based on the Kannai transform combined with the finite speed propagation property for sR waves and on upper exponential estimates for the sR heat kernel.

2. Reminders: sub-Riemannian (sR) structure

Attached with the m-tuple of vector fields $(X_1, \ldots, X_m)$, there is a canonical sub-Riemannian (sR) structure. This section consists of reminders in sR geometry (see the textbooks [ABB20, Bel96, Gro96, Jea14, LD10, Mon02, Rif14]), which are useful in our analysis.

2.1. Sub-Riemannian metric

The sR metric g associated with $(X_1, \ldots, X_m)$ is defined as follows: given any $q \in M$ and any $v \in D_q = \text{Span}(X_1(q), \ldots, X_m(q))$, we define the positive definite quadratic form g_q on D_q by

$$g_q(v) = \inf \left\{ \sum_{i=1}^m u_i^2 \ \middle| \ v = \sum_{i=1}^m u_i X_i(q) \right\}.$$

The triple (M, D, g) is called a sub-Riemannian structure (see [Bel96, Rif14]). When D has constant rank m on M with $m \leqslant n$, D is a subbundle of TM endowed with the Riemannian metric g, and the frame $(X_1, \ldots, X_m)$ is g-orthonormal. But the rank of D may vary (i.e., D is a subsheaf of TM) and the above definition encompasses the so-called almost-Riemannian case, for which $m \geqslant n$ and $\text{rank}(D) < n$ at some singular points.

More formally, a sR structure on M can be defined by an Euclidean vector bundle E over M and a smooth vector bundle morphism $\sigma : E \to TM$, with $D_q = \sigma(E_q)$ and $g_q(V) = \inf\{\|u\|_{E_q}^2 \ | \ u \in E_q, \ \sigma(u) = V\}$ for every $q \in M$. When $E = M \times \mathbb{R}^m$ and $\sigma(x, u) = \sum_{i=1}^m u_i X_i(x)$, we recover the definition of a sR structure attached with the m vector fields $X_1, \ldots, X_m$.

A horizontal path is, by definition, an absolutely continuous path $q(\cdot) : [0, 1] \to M$ for which there exist m functions $u_i \in L^1(0, 1)$ such that $\dot{q}(t) = \sum_{i=1}^m u_i(t) X_i(q(t))$ for almost every $t \in [0, 1]$. The metric g induces a length on the set of horizontal paths, and thus a distance d_{sR} on M that is called the sR distance.

The cometric g^* associated with $(X_1, \ldots, X_m)$ is the nonnegative quadratic form on T^*M defined as follows: given any $q \in M$, g_q^* is the nonnegative quadratic form defined on T_q^*M by $g_q^*(\xi) = \sum_{i=1}^m \langle \xi, X_i(q) \rangle^2$. Note that

$$\frac{1}{2} g_q(v) = \sup_{\xi \in T_q^*M} \left(\langle \xi, v \rangle - \frac{1}{2} g_q^*(v) \right)$$

(Legendre transform).

Given any smooth function f on M, the horizontal gradient $\nabla_g f$ of f is the smooth section of D defined by $g(\nabla_g f, X) = df.X$ for every smooth section X of D. We have $\nabla_g f = \sum_{i=1}^m (X_i f) X_i$.

2.2. Sub-Riemannian Laplacian

We denote by div_μ the divergence operator associated with the smooth measure μ on M, defined by $\mathcal{L}_X \mu = \text{div}_\mu(X)\, \mu$ for every smooth vector field X on M. Here, $\mathcal{L}_X$ is the Lie derivative along X. The sR Laplacian $\triangle_{\text{sR}}$ is defined as the differential operator

$$(2.1) \qquad \triangle_{\text{sR}} f = \text{div}_\mu (\nabla_g f) = \sum_{i=1}^m \left(X_i^2 f + \text{div}_\mu(X_i) X_i f \right) \qquad \forall\, f \in C^\infty(M).$$

Its principal symbol is the cometric g^*. The sR Laplacian is a particular instance of a Hörmander operator: we have $\triangle_{\text{sR}} = \triangle$ with $X_0 = \sum_{i=1}^m \text{div}_\mu(X_i) X_i$ and $\mathbb{V} = 0$.

Let Ω be an open subset of M. By integration by parts, we have

$$\langle \triangle_{\mathrm{sR}} u, v \rangle_{L^2(\Omega, \mu)} = - \int_\Omega g\left(\nabla_g u, \nabla_g v\right) d\mu + \int_{\partial\Omega} v\, d\left(\iota_{\nabla_g u}\mu\right) \qquad \forall\, u, v \in C^\infty(M)$$

where $\iota_{\nabla_g u}$ is the interior product of μ and $\nabla_g u$. We infer the sR Green formula

$$\langle \triangle_{\mathrm{sR}} u, v \rangle_{L^2(\Omega, \mu)} = \langle u, \triangle_{\mathrm{sR}} v \rangle_{L^2(\Omega, \mu)} + \int_{\partial\Omega} v\, d\left(\iota_{\nabla_g u}\mu\right) - \int_{\partial\Omega} u\, d\left(\iota_{\nabla_g v}\mu\right)$$

for all $u, v \in C^\infty(M)$. Hence, $\triangle_{\mathrm{sR}}$ is symmetric and dissipative on $C^\infty(\Omega)$ in the two following cases:

- Dirichlet case: $f = 0$ along $\partial\Omega$;
- Neumann case: $\iota_{\nabla_g f}\mu = 0$ along $\partial\Omega$.

In these two cases, $\triangle_{\mathrm{sR}}$ has selfadjoint extensions; moreover, if the manifold Ω endowed with the induced sR distance is complete then $\triangle_{\mathrm{sR}}$ is essentially selfadjoint (see [Str86]) and thus has a unique selfadjoint extension. We speak then of the Dirichlet (resp., Neumann) sR Laplacian, which is defined on the maximal domain that is the completion in $L^2(\Omega, \mu)$ of the subset of $f \in C^\infty(\Omega)$ satisfying $f = 0$ (resp., $\iota_{\nabla_g f}\mu = 0$) along $\partial\Omega$.

Note that, if $\Omega = M$ is compact or if $\Omega = M = \mathbb{R}^n$ or if one considers Dirichlet boundary conditions then the adjoint of X_i in $L^2(\Omega, \mu)$ is $X_i^* = -X_i - \mathrm{div}_\mu(X_i)$ and then $\triangle_{\mathrm{sR}} = -\sum_{i=1}^m X_i^* X_i$.

2.3. Sub-Riemannian flag

We define the sequence of subsheaves D^k of TM by $D^0 = \{0\}$, $D^1 = D = \mathrm{Span}(X_1, \ldots, X_m)$ and $D^{k+1} = D^k + [D, D^k]$ for $k \geqslant 1$. Under the strong Hörmander condition (1.2), given any point $q \in M$, we have the flag

$$\{0\} = D_q^0 \subset D_q = D_q^1 \subset D_q^2 \subset \ldots \subset D_q^{r(q)-1} \subsetneq D_q^{r(q)} = T_q M$$

where $r(q)$ is called the *degree of nonholonomy* at q. We set $n_i(q) = \dim D_q^i$. The $r(q)$-tuple of integers $(n_1(q), \ldots, n_{r(q)}(q))$ is called the *growth vector* at q, and we have $n_{r(q)}(q) = n = \dim M$. By convention, we set $n_0(q) = 0$. We define the nondecreasing sequence of weights $w_i(q)$ as follows: given any $i \in \{1, \ldots, n\}$, there exists a unique $j \in \{1, \ldots, n\}$ such that $n_{j-1}(q) + 1 \leqslant i \leqslant n_j(q)$, and we set $w_i(q) = j$. By definition, we have $w_1(q) = \cdots = w_{n_1}(q) = 1$, and $w_{n_{j-1}+1}(q) = \cdots = w_{n_j}(q) = j$ when $n_j(q) > n_{j-1}(q)$. We also have $w_{n_{r-1}+1}(q) = \cdots = w_{n_r}(q) = r(q)$. Note that $n_{w_{n_j}}(q) = n_j(q)$ for $j = 1, \ldots, r$ and that $w_{n_j}(q) = j$ if (and only if) $n_j(q) > n_{j-1}(q)$.

Given any $q \in M$, we set

$$\mathcal{Q}(q) = \sum_{i=1}^r i\left(n_i(q) - n_{i-1}(q)\right) = \sum_{i=1}^n w_i(q).$$

If q is regular then $\mathcal{Q}(q)$ is the Hausdorff dimension of a small ball in M containing q endowed with the induced corresponding sR distance (see [Gro96]).

A point $q \in M$ is said to be *regular* if the growth vector is constant in a neighborhood of q; otherwise it is said to be *singular*. The sR structure is said to be

equiregular if all points of M are regular; in this case, the weights and the Hausdorff dimension are constant as well on M.

We recall that, at a point q that is regular, with degree of nonholonomy $r(q)$, $\triangle$ is locally hypoelliptic and even subelliptic with a gain of regularity $2/r(q)$, meaning that if $\triangle f = g$ with g of Sobolev class H^s locally at q then f is (at least) of Sobolev class $H^{s+2/r(q)}$ locally at q (see [Hör67]).

2.4. Sub-Riemannian isometries

Given two sR structures (M_1, D_1, g_1) and (M_2, D_2, g_2), of respective cometrics g_1^* and g_2^*, a (local) sR isometry $\phi : M_1 \to M_2$ is a (local) smooth diffeomorphism mapping g_1^* to g_2^*.

Note that, if M is a Lie group equipped with a left-invariant sR structure, then the left action is a sR isometry on M.

2.5. Nilpotentization of the sub-Riemannian structure

2.5.1. First definition

Let $q \in M$ be arbitrary. The nilpotentization of the sR structure (M, D, g) at q is defined as the metric tangent space of M (endowed with its sR distance) in the sense of Gromov-Hausdorff (see [Bel96, Gro96]). It is identified, with a sR isometry, to the sR structure $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$ defined hereafter, where $\widehat{M}^q$ is a smooth connected manifold of dimension n (as a topological space, $\widehat{M}^q$ is the usual tangent space to M at q), $\widehat{D}^q = \mathrm{Span}(\widehat{X}_1^q, \ldots, \widehat{X}_m^q)$ with smooth vector fields $\widehat{X}_1^q, \ldots, \widehat{X}_m^q$ on M (defined hereafter) called nilpotentizations at q of the vector fields $X_1, \ldots, X_m$ at q, and the sR metric $\widehat{g}^q$ is defined, accordingly, by

$$\widehat{g}_x^q(v) = \inf \left\{ \sum_{i=1}^m u_i^2 \; \middle| \; v = \sum_{i=1}^m u_i \widehat{X}_i^q(x) \right\} \qquad \forall \, x \in \widehat{M}^q \qquad \forall \, v \in \widehat{D}_x^q.$$

The metric $\widehat{g}^q$ induces a distance $\widehat{\mathrm{d}}_{\mathrm{sR}}^q$ on $\widehat{M}^q$.

To define $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$, it suffices to define one of the elements of the equivalence class under the action of sR isometries. A standard description consists of using charts in M of so-called *privileged coordinates*, and then to identify $\widehat{M}^q \simeq \mathbb{R}^n$ with a sR isometry, as follows.

2.5.2. Privileged coordinates

We first recall the notion of nonholonomic order (see [Bel96, Jea14, Mon02] for details). Given a germ f of a smooth function at q, given $k \in \mathbb{N}$ and integers $j_1, \ldots, j_k$ in $\{1, \ldots, m\}$, the Lie derivative $\mathcal{L}_{X_{j_1}} \cdots \mathcal{L}_{X_{j_k}} f(q)$ is called a nonholonomic derivative of order k. By definition, the nonholonomic order of f at q, denoted by $\mathrm{ord}_q(f)$, is the smallest integer k for which at least one nonholonomic derivative of f of order

k at q is not equal to zero. Given a germ Y of a smooth vector field at q, the non-holonomic order of Y at q is the largest integer k such that $\mathrm{ord}_q(\mathcal{L}_Y f) \geqslant k + \mathrm{ord}_q(f)$, for every germ f at q.

A family $(Z_1, \ldots, Z_n)$ of n vector fields is said to be adapted to the flag at q if it is a frame of $T_q M$ at q and if $Z_i(q) \in D_q^{w_i(q)}$, for every $i \in \{1, \ldots, n\}$.

A system of privileged coordinates at q is a system of local coordinates $(x_1, \ldots, x_n)$ at q such that $\mathrm{ord}_q(x_j) = w_j(q)$, for every $i \in \{1, \ldots, n\}$. Note that we must have $dx_i(D_q^{w_i(q)}) \neq 0$ and $dx_i(D_q^{w_i(q)-1}) = 0$, meaning that $\partial_{x_i} \in D_q^{w_i(q)} \setminus D_q^{w_i(q)-1}$ at q (i.e., privileged coordinates are always adapted to the flag).

An example of privileged coordinates at q is given by

$$(2.2) \qquad (x_1, \ldots, x_n) \mapsto \exp\left(\sum_{i=1}^{n} x_i Z_i^q\right)(q)$$

where $(Z_i^q)_{1 \leqslant i \leqslant n}$ is a frame of vector fields that is adapted to the flag at q.

Privileged coordinates can be obtained from any system of adapted coordinates by a triangular change of variables (see [Jea14]).

2.5.3. Dilations and nilpotentization of smooth sections of D

We consider a chart of privileged coordinates at q, that is a smooth mapping $\psi_q : U \mapsto \mathbb{R}^n$, where U is a neighborhood of q in M, with $\psi_q(q) = 0$, inducing local coordinates $x = \psi_q(q)$ in which the vector fields $(\psi_q)_* X_i$ ($i = 1, \ldots, m$) have a nilpotent approximation $\widehat{X}_i^q$ with the following precise meaning. In these local coordinates, for every $\varepsilon \in \mathbb{R}$, the dilation δ_ε is defined in $\mathbb{R}^n$, according to the flag at q, by

$$\delta_\varepsilon(x) = \left(\varepsilon^{w_1(q)} x_1, \ldots, \varepsilon^{w_n(q)} x_n\right) \qquad \forall\, x = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Note that, denoting by m the Lebesgue measure on $\mathbb{R}^n$ (given by $dm = dx_1 \cdots dx_n$), we have $\delta_\varepsilon^* m = |\varepsilon|^{\mathcal{Q}(q)} m$ for every $\varepsilon \neq 0$.

Given any vector field X on M that is a smooth section [4] of D (i.e., $X(q) = \sum_{i=1}^{m} a_i(q) X_i(q)$ at any point q, with smooth functions a_i), the nilpotentization $\widehat{X}^q$ at q of X is the (nilpotent and complete) vector field on $\mathbb{R}^n$ defined in the chart by

$$\widehat{X}^q = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \varepsilon \delta_\varepsilon^* X.$$

Actually this convergence is valid in C^∞ topology (uniform convergence of all derivatives on compact subsets of $\mathbb{R}^n$). Note that $\widehat{X}^q$ is homogeneous of order -1 with respect to dilations, i.e., $\lambda \delta_\lambda^* \widehat{X}^q = \widehat{X}^q$ for every $\lambda \neq 0$, and that the nonholonomic order of $X - \widehat{X}^q$ at q is nonnegative. Actually, setting $X^\varepsilon = \varepsilon \delta_\varepsilon^* X$ and writing in C^∞ topology the Taylor expansion $X = X^{(-1)} + X^{(0)} + X^{(1)} + \cdots$ around 0, where $X^{(k)}$ is polynomial and homogeneous of degree k (with respect to dilations), we get that X^ε has a Taylor expansion at any order N with respect to ε, in C^∞ topology:

$$(2.3) \qquad X^\varepsilon = \varepsilon \delta_\varepsilon^* X = \widehat{X}^q + \varepsilon X^{(0)} + \varepsilon^2 X^{(1)} + \cdots + \varepsilon^N X^{(N-1)} + \mathrm{o}\!\left(|\varepsilon|^N\right)$$

---

[4] Note that we consider a smooth section of the subsheaf D, otherwise there are some difficulties: take $M = \mathbb{R}^2$, D spanned by $X_1 = \partial_x$ and $X_2 = x^6 \partial_y$, and the vector field $X = x^2 \partial_y$.

with $\widehat{X}^q = X^{(-1)}$ (see also [Bar13, Lemma 1]), i.e., setting $X^0 = \widehat{X}^q$ for $\varepsilon = 0$, X^ε depends smoothly on ε in C^∞ topology. We also have

$$(2.4) \qquad\qquad\qquad X^\varepsilon = \widehat{X}^q + \varepsilon Z^\varepsilon$$

for every $\varepsilon \in \mathbb{R}$ with $|\varepsilon|$ small enough so that we are in the chart, where Z^ε is a smooth vector field depending smoothly on ε in C^∞ topology.

2.5.4. Definition of the nilpotentization of the sR structure

In the above chart, we define $\widehat{M}^q \simeq \mathbb{R}^n$, endowed with the sR structure (denoted by $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$) induced by the vector fields $\widehat{X}_i^q$, $i = 1, \ldots, m$. This definition does not depend on the choice of privileged coordinates at q because two sets of such coordinates produce two sR-isometric sR structures. This is due to the fact that, since transition maps of charts of privileged coordinates are triangular with respect to the flag, the nilpotentization of any transition map is a sR isometry (see [Bel96, Proposition 5.20]). Note that the nilpotent sR structure $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$ is homogeneous with respect to the above dilations and that the corresponding sR distance is homogeneous of order 1. Moreover, the growth vector of $\widehat{D}^q$ coincides with that of D at q, and $\mathrm{Lie}(\widehat{X}_1^q, \ldots, \widehat{X}_m^q)$ is a nilpotent Lie algebra of step $r(q)$.

It follows from the definition of the sR metric that

$$\widehat{g}^q = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \varepsilon^{-2} \delta_\varepsilon^* g \qquad \text{and} \qquad \widehat{g}_x^q\left(\widehat{X}^q(x), \widehat{Y}^q(x)\right) = g_q\left(X(q), Y(q)\right)$$

for every $x \in \mathbb{R}^n$, for all vector fields X and Y on M that are smooth sections of D.

Another useful geometric identification of $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$ is the following. Fixing a chart of privileged coordinates at q, let $\mathcal{G}_q$ be the (nilpotent) Lie group of diffeomorphisms of $\mathbb{R}^n$ generated by $\exp(t\widehat{X}_i^q)$, for $t \in \mathbb{R}$ and $i = 1, \ldots, m$. Its Lie algebra is

$$\mathfrak{g}_q = \mathrm{Lie}\left(\widehat{X}_1^q, \ldots, \widehat{X}_m^q\right) = \bigoplus_{i=1}^{r(q)} \left(\widehat{D}^q\right)^i / \left(\widehat{D}^q\right)^{i-1}$$

it is nilpotent, graded, and generated by its first component $\widehat{D}^q$. In other words, $\mathcal{G}_q$ is a Carnot group (see [Mon02]). Under the strong Hörmander condition (1.2), $\mathcal{G}_q$ acts transitively on $\mathbb{R}^n$. Defining the isotropy group $H_q = \{\varphi \in \mathcal{G}_q \mid \varphi(0) = 0\}$, of Lie algebra $\mathfrak{h}_q = \{Y \in \mathfrak{g}_q \mid Y(0) = 0\}$, we identify $\widehat{M}^q$ to the homogeneous (coset) space $\mathcal{G}_q/H_q$. If q is regular then $H_q = \{0\}$ and thus $\widehat{M}^q \simeq \mathcal{G}_q$ is a Carnot group endowed with a left-invariant sR structure.

Remark 2.1. — Carnot groups are to sub-Riemannian geometry as Euclidean spaces are to Riemannian geometry. However, there is a major difference, which is of particular importance here. In Riemannian geometry, all tangent spaces are isometric, but this is not the case in sub-Riemannian geometry: given two points q_1 and q_2 of M, the nilpotentizations $(\widehat{M}^{q_1}, \widehat{D}^{q_1}, \widehat{g}^{q_1})$ and $(\widehat{M}^{q_2}, \widehat{D}^{q_2}, \widehat{g}^{q_2})$ of the sR structure respectively at q_1 and q_2 may not be sR-isometric, even though the growth

vectors at q_1 and q_2 coincide.[5] There are many algebraically non-isomorphic (and thus non-isometric) n-dimensional Carnot groups, and even uncountably many for $n \geqslant 5$ (due to moduli in their classification). We refer to [AM05, Mar07] for a complete classification of rigid and semi-rigid Carnot algebras.

Note that, in dimension three, if the growth vector is $(2, 3)$ then we have a unique model that is the Heisenberg flat case in the equivalence class of sR-isometric Carnot groups.

2.5.5. Nilpotentized sR Laplacian

Let $q \in M$ be arbitrary. Associated with the sR structure $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$, we define on $C^\infty(\widehat{M}^q)$ the differential operator

$$(2.5) \qquad \widehat{\triangle}^q = \sum_{i=1}^{m} \left(\widehat{X}_i^q\right)^2.$$

2.5.6. Nilpotentization of measures

Let us define the nilpotentization of a smooth measure μ on M. Let $q \in M$ be arbitrary. Using the bijective correspondence between smooth measures and densities, the measure μ induces a volume form that we consider at the point q. Then, the canonical isomorphism [6]

$$\Lambda^n \left(T_q^\star M\right) \simeq \Lambda^n \left(\bigoplus_{k=1}^{r(q)} D_q^k / D_q^{k-1}\right)^*$$

induces a measure $\widehat{\mu}^q$ on $\widehat{M}^q$. Using a chart ψ_q of privileged coordinates at q, and using the isometric representation $\widehat{M}^q \simeq \mathbb{R}^n$, the measure $\widehat{\mu}^q$ on $\mathbb{R}^n$ is given in the chart by

$$\widehat{\mu}^q = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} \frac{1}{|\varepsilon|^{\mathcal{Q}(q)}} \delta_\varepsilon^* \mu$$

where the convergence is understood in the vague topology (i.e., the weak star topology of $C_c(M)'$, where $C_c(M)$ is the set of continuous functions on M of compact support). According to this definition, if μ and ν are two smooth measures on M, with $\mu = h\nu$, where h is a positive smooth function on M, then $\widehat{\mu}^q = h(q)\widehat{\nu}^q$. Equivalently, this means that

$$(2.6) \qquad h(q) = \frac{d\mu}{d\nu}(q) = \frac{\widehat{\mu}^q}{\widehat{\nu}^q}.$$

---

[5] Actually, the flags of two sR structures coincide at any point if and only if the sR structures are locally Lipschitz equivalent, meaning that the corresponding sR distances satisfy $c_1 d_2(q, q') \leqslant d_1(q, q') \leqslant c_2 d_2(q, q')$ for some uniform constants $c_1 > 0$ and $c_2 > 0$.

[6] Indeed, following [ABGR09], considering a basis $(e_1, \ldots, e_n)$ of $T_q M$ that is adapted to the flag, that is, such that $e_i \in D_q^{w_i(q)}$, the wedge product $e_1 \wedge \cdots \wedge e_n$ depends only on $e_i \mod D_q^{w_i(q)-1}$. This induces the canonical isomorphism.

In particular, the nilpotentizations at q of all smooth measures are proportional to the Lebesgue measure m on $\widehat{M}^q \simeq \mathbb{R}^n$. Note that, if q is regular, then $\widehat{\mu}^q$ is a left-invariant measure on the Carnot group $\widehat{M}^q$; in this case, $\widehat{M}^q$ is a nilpotent Lie group and thus is unimodular, and hence $\widehat{\mu}^q$ coincides with the Haar measure, up to scaling.

In passing, note that, applying (2.6) to the measure $\nu = \mathcal{H}_S$ that is the spherical Hausdorff measure and using the fact (proved in [ABB12]) that $\widehat{\mathcal{H}_S}^q(\widehat{B}^q) = 2^{\mathcal{Q}(q)}$, we obtain that the density at q of μ with respect to the spherical Hausdorff measure is

$$h(q) = \frac{d\mu}{d\mathcal{H}_S}(q) = \frac{\widehat{\mu}^q\left(\widehat{B}^q\right)}{2^{\mathcal{Q}(q)}}.$$

Remark 2.2. — Let $q \in M$ be arbitrary, and let μ be an arbitrary smooth measure on M. Endowing $\widehat{M}^q$ with the measure $\widehat{\mu}^q$, we claim that

$$(2.7) \qquad \operatorname{div}_{\widehat{\mu}^q}\left(\widehat{X}_i^q\right) = 0 \quad \forall\, i \in \{1,\, \ldots,\, m\}.$$

As a consequence, we have $(\widehat{X}_i^q)^* = -\widehat{X}_i^q$, where the transpose is considered in $L^2(\widehat{M}^q, \widehat{\mu}^q)$. It follows that

$$\widehat{\triangle}^q = \sum_{i=1}^m \left(\widehat{X}_i^q\right)^2 = \sum_{i=1}^m -\left(\widehat{X}_i^q\right)^* \widehat{X}_i^q.$$

Due to the cancellation of the divergence term, there are no terms of order one (compare with the general formula for a sR Laplacian, given, e.g., in [CdVHT16]).

Let us prove (2.7). Following [Jea14, page 25], we write $\widehat{X}_i^q(x) = \sum_{j=1}^n a_{ij}(x)\partial_{x_j}$ in a chart of privileged coordinates. By definition, ∂_{x_j} is homogeneous of degree $-w_j(q)$, and since $\widehat{X}_i^q$ is homogeneous of degree -1, it follows that a_{ij} is a homogenous polynomial of degree $w_j(q) - 1$. Since x_k is of weight greater than or equal to $w_j(q)$ for $k \geqslant j$, a_{ij} does not depend on variables x_k for $k \geqslant j$. It follows that $\partial_{x_j} a_{ij} = 0$, and hence $\operatorname{div}_{\widehat{\mu}^q}(\widehat{X}_i^q) = 0$ (recall that $\widehat{\mu}^q$ is equal, up to constant scaling, to the Lebesgue measure on $\mathbb{R}^n$).

Part 1. Proof of the limit in Theorem A

This part is devoted to proving the following result (first term in the complete asymptotic expansion stated in Theorem A).

THEOREM B. — *Let $q \in M$ be arbitrary (regular or not). Let $\psi_q : U \to V$ be a chart of privileged coordinates at q such that $\psi_q(q) = 0$, where U is an open connected neighborhood of q in M and V is an open neighborhood of 0 in $\mathbb{R}^n$. We assume that X_0 is a smooth section of D over M. We also assume that $\sup_{q' \in U} r(q') < +\infty$. Then, in the chart, we have*

$$(2.8) \qquad \lim_{\substack{\varepsilon \to 0 \\ \varepsilon \neq 0}} |\varepsilon|^{\mathcal{Q}(q)}\, e_{\triangle,\,\mu}\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) = \widehat{e}^q(t, x, x')$$

in $C^\infty((0, +\infty) \times V \times V)$ topology.

Moreover, if q is regular, then the above convergence is locally uniform with respect to q, and the function $\widehat{e}^q$ depends smoothly (in C^∞ topology) on q in any

open neighborhood of q consisting of regular points. If the manifold M is Whitney stratifiable with strata defined according to the sR flag (i.e., the growth vector $(n_1(q), \ldots, n_{r(q)}(q))$ is constant along each stratum) then the latter property is satisfied along strata.

When X_0 is a smooth section of D^2 over M, the above statement remains true, replacing $\widehat{\triangle}^q$ with $\widehat{\triangle}^q + \widehat{X}_0^q$ (see Remark 1.2).

Sketch of proof. In few words, the proof goes as follows (see Section 4 for all details).

Assume first that $M = \mathbb{R}^n$. For every $i \in \{0, \ldots, m\}$, the vector field $X_i^\varepsilon = \varepsilon \delta_\varepsilon^* X_i$ converges to $\widehat{X}_i^q$ in C^∞ topology as $\varepsilon \to 0$. Hence the operator

$$\triangle^\varepsilon = \varepsilon^2 \delta_\varepsilon^* \triangle (\delta_\varepsilon)_* = \sum_{i=1}^m (X_i^\varepsilon)^2 + \varepsilon X_0^\varepsilon - \varepsilon^2 (\delta_\varepsilon^* \mathbb{V})$$

converges to $\widehat{\triangle}^q = \sum_{i=1}^m \left(\widehat{X}_i^q \right)^2$ in C^∞ topology. By the Trotter–Kato theorem (see [EN00, Paz83]), the corresponding heat kernel

$$(2.9) \qquad e^\varepsilon(t, x, x') = |\varepsilon|^{\mathcal{Q}(q)} e_{\triangle, \mu} \left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x') \right)$$

converges to $\widehat{e}^q$ in a weak topology, but actually convergence is true as well in $C^\infty((0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$ topology because, by uniform local subellipticity of

$$(\triangle^\varepsilon)_{\varepsilon \in [-\varepsilon_0, \varepsilon_0]},$$

the family $(e^\varepsilon)_{\varepsilon \in [-\varepsilon_0, \varepsilon_0]}$ is uniformly bounded in the Montel space C^∞, for some $\varepsilon_0 > 0$ small enough.

On a general manifold, we follow the above argument in a local chart around q, extending the vector fields X_i by 0 outside of a neighborhood of q. The relation (2.9) is then not exactly true, but, thanks to the fact that the small-time asymptotics of hypoelliptic heat kernels is purely local (Kac's principle), the relation (2.9) remains true with a remainder term $O(|\varepsilon|^\infty)$ as $\varepsilon \to 0$, and we conclude as well.

As one can see, in the above argument, we only use *local* results:

- local subellipticity estimates and local smoothing properties for hypoelliptic heat kernels, uniform with respect to some parameters: these are well known results, but for completeness (and in order to prepare global estimates), we give statements and proofs in Appendix B.1;
- local nature of the small-time asymptotics of hypoelliptic heat kernels (Kac's principle), uniform with respect to parameters: this is established in Theorem 3.2 in Section 3.2.1.

Incidentally, we state in Theorem 3.5 (in Section 3.2.2) a general convergence result for hypoelliptic heat kernels depending on parameters: this is a singular perturbation theorem for hypoelliptic operators, generalizing existing singular perturbation results for elliptic operators.

3. Some general facts for hypoelliptic heat kernels

3.1. Hörmander operators, semigroups and heat kernels

3.1.1. Preliminary remarks on Hörmander operators

In this section, we make some remarks on Hörmander operators, which are useful in view of defining the domains of such operators and then show existence of semigroups, as done in Section 3.1.2. Let Ω be an open subset of M.

Integration by parts with a Hörmander operator. We consider the differential operator $\triangle$ is defined by (1.1). By integration by parts, we compute

$$(3.1) \quad \langle \triangle f, f \rangle_{L^2(\Omega, \mu)} = -\sum_{i=1}^m \|X_i f\|^2_{L^2(\Omega, \mu)} + \int_{\partial \Omega} f \, d\left(\iota_Y \mu\right)$$
$$+ \frac{1}{2} \int_\Omega f^2 \left(\sum_{i=1}^m \left(X_i \left(\operatorname{div}_\mu(X_i)\right) + \left(\operatorname{div}_\mu(X_i)\right)^2 \right) - \operatorname{div}_\mu(X_0) - 2\mathbb{V} \right) d\mu$$

with

$$(3.2) \qquad Y = \sum_{i=1}^m \left(X_i f - \frac{1}{2} f \operatorname{div}_\mu(X_i) \right) X_i + \frac{1}{2} f X_0.$$

Of course, if $\partial \Omega = \emptyset$ then there is no boundary term. When $\partial \Omega \neq \emptyset$, the boundary term $\int_{\partial \Omega} f \, d\left(\iota_Y \mu\right)$ is equal to zero in the two following cases:

- Dirichlet case: $f = 0$ along $\partial \Omega$;
- Neumann case: $\iota_Y \mu = 0$ along $\partial \Omega$.

In particular, we have defined here the Neumann boundary condition for the Hörmander operator $\triangle$: the interior product of μ and of the vector field Y (defined by (3.2)) is zero along $\partial \Omega$. Note that

$$Y = \nabla_g f + \frac{1}{2} f \left(X_0 - \sum_{i=1}^m \operatorname{div}_\mu(X_i) X_i \right)$$

where $\nabla_g f$ is the horizontal gradient of f.

Symmetry properties of Hörmander operators. Recalling that the differential operator $\triangle$ is defined by (1.1) and that $\triangle_{\mathrm{sR}}$ is defined by (2.1), we have

$$\triangle = \triangle_{\mathrm{sR}} + X_0' - \mathbb{V} \qquad \text{with} \qquad X_0' = X_0 - \sum_{i=1}^m \operatorname{div}_\mu(X_i) X_i.$$

Integrating by parts, we compute

$$\langle \triangle u, v \rangle_{L^2(\Omega, \mu)} = -\int_\Omega g\left(\nabla_g u, \nabla_g v\right) d\mu - \int_\Omega u(X_0' v) \, d\mu - \int_\Omega uv \left(\operatorname{div}_\mu\left(X_0'\right) + \mathbb{V}\right) d\mu$$
$$+ \int_{\partial \Omega} v \, d\left(\iota_{\nabla_g u} \mu\right) + \int_{\partial \Omega} uv \, d\left(\iota_{X_0' u} \mu\right)$$

and we infer the Green formula for Hörmander operators:

$$\langle \triangle u, v \rangle_{L^2(\Omega,\mu)} = \langle u, \triangle v \rangle_{L^2(\Omega,\mu)} + \int_\Omega \left(v(X_0'u)\, d\mu - u(X_0'v) \right) d\mu$$
$$+ \int_{\partial\Omega} v\, d\left(\iota_{\nabla_g u}\mu \right) - \int_{\partial\Omega} u\, d\left(\iota_{\nabla_g v}\mu \right) \qquad \forall\, u, v \in C^\infty(M).$$

For the operator $\triangle$ to be symmetric on $C^\infty(\Omega)$, there are two necessary conditions:

- The term $\int_\Omega (v(X_0'u)\, d\mu - u(X_0'v))\, d\mu$ must be zero, which is the case if and only if $X_0' = 0$, i.e., $X_0 = \sum_{i=1}^m \operatorname{div}_\mu(X_i) X_i$ on Ω.
- The boundary term must be zero. This is the case for Dirichlet boundary conditions. For Neumann boundary conditions, using the fact that $X_0' = 0$ by the first item and thus that the vector field Y defined by (3.2) coincides with the horizontal gradient, we see then that the Neumann boundary condition for the Hörmander operator coincides with the sR Neumann boundary condition.

Therefore, in the Dirichlet as in the Neumann case, the Hörmander operator $\triangle$ is symmetric on $C^\infty(\Omega)$ if and only if $X_0 = \sum_{i=1}^m \operatorname{div}_\mu(X_i) X_i$ on Ω. In this case, $\triangle = \triangle_{sR} - \mathbb{V}$ has selfadjoint extensions; moreover, if the manifold Ω endowed with the induced sR distance is complete then $\triangle$ is essentially selfadjoint on $C_c^\infty(\Omega)$ and thus has a unique selfadjoint extension.

3.1.2. Hypoelliptic semigroups and heat kernels

We consider the operator $\triangle : D(\triangle) \to L^2(\Omega, \mu)$ defined on a domain $D(\triangle)$ that is assumed to be dense in $L^2(\Omega, \mu)$ and for which $(\triangle, D(\triangle))$ is closed. When $\partial\Omega \neq \emptyset$, the domain $D(\triangle)$ encodes some possible boundary conditions on $\partial\Omega$.

According to Section 3.1.1, given a sufficiently regular function f, we speak of the Dirichlet boundary condition when $f = 0$ along $\partial\Omega$, and of the Neumann boundary condition when $\iota_Y\mu = 0$ (interior product of μ and Y) along $\partial\Omega$, where Y is defined by (3.2).

Let $q \in M$ be arbitrary. Let μ be an arbitrary smooth measure on M. We set $D(\widehat{\triangle}^q) = \{f \in L^2(\widehat{M}^q, \widehat{\mu}^q) \mid \widehat{\triangle}^q f \in L^2(\widehat{M}^q, \widehat{\mu}^q)\}$. According to Section 3.1.1 and Remark 2.2, since $(\widehat{M}^q, \widehat{d}_{sR}^q)$ is complete (indeed, sR balls of small radius are compact, and $\widehat{M}^q$ is invariant under dilations), the operator $\widehat{\triangle}^q : D(\widehat{\triangle}^q) \to L^2(\widehat{M}^q, \widehat{\mu}^q)$ is selfadjoint.

LEMMA 3.1. — *Under any of the following assumptions:*

(A) $\triangle : D(\triangle) \to L^2(\Omega, \mu)$ *is selfadjoint (see Section 3.1.1);*

(B) *the functions* $\operatorname{div}_\mu(X_0)$, $\operatorname{div}_\mu(X_i)$ *and* $X_i(\operatorname{div}_\mu(X_i))$, $i = 1, \ldots, m$, *are bounded on* Ω, *and we have Dirichlet or Neumann boundary conditions whenever* $\partial\Omega \neq \emptyset$;

the operator $(\triangle, D(\triangle))$ *generates a strongly continuous semigroup* $(e^{t\triangle})_{t \geqslant 0}$ *on* $L^2(\Omega, \mu)$.

The operator $(\widehat{\triangle}^q, D(\widehat{\triangle}^q))$ *is selfadjoint and generates a strongly continuous contraction semigroup* $(e^{t\widehat{\triangle}^q})_{t \geqslant 0}$ *on* $L^2(\widehat{M}^q, \widehat{\mu}^q)$.

Proof. — Under Assumption **A**, it follows from Sections 2.2 and 3.1.1 that $\triangle = \triangle_{\mathrm{sR}} - \mathbb{V}$ with $\mathbb{V}$ bounded below and thus there exists $C \geqslant 0$ such that $\triangle - C \operatorname{id}$ is dissipative.

Under Assumption **B**, by integration by parts, we have the formula (3.1). The integral on $\partial\Omega$ is zero in the case of Dirichlet or Neumann boundary conditions. Hence there exists $C \geqslant 0$ (not depending on f) such that $\langle (\triangle - C \operatorname{id})f, f\rangle_{L^2(\Omega,\mu)} \leqslant 0$ for every $f \in D(\triangle)$, and thus $\triangle - C \operatorname{id}$ is dissipative in $L^2(\Omega,\mu)$. The same result holds for its adjoint.

Now, since the operator $\triangle - C \operatorname{id}$ in $L^2(\Omega,\mu)$ is closed and dissipative as well as its adjoint, it follows from the Lumer–Phillips theorem (see, e.g., [EN00, Paz83]) that it generates a strongly continuous contraction semigroup $(e^{t(\triangle - C \operatorname{id})})_{t \geqslant 0}$. Then, the operator $\triangle$ generates a strongly continuous semigroup $(e^{t\triangle})_{t \geqslant 0}$, and we have $e^{t\triangle} = e^{Ct}e^{t(\triangle - C \operatorname{id})}$.

The operator $\widehat{\triangle}^q$ on $L^2(\widehat{M}^q, \widehat{\mu}^q)$ is closed, selfadjoint and dissipative, and thus it generates a strongly continuous contraction semigroup $(e^{t\widehat{\triangle}^q})_{t \geqslant 0}$. $\square$

Hypoelliptic heat kernels. Under the assumptions done in Lemma 3.1, since $\mathrm{Lie}(D) = TM$ (which implies $\mathrm{Lie}(\widehat{D}^q) = T\widehat{M}^q$)[7], both operators $\partial_t - \triangle$ and $\partial_t - \widehat{\triangle}^q$ are hypoelliptic and therefore the corresponding heat kernels exist and are smooth: given any smooth measure μ on M, we consider the heat kernel $e_{\triangle,\mu}$ defined on $(0, +\infty) \times \Omega \times \Omega$, associated with the operator $\triangle$ and with the measure μ, and the heat kernel $\widehat{e}^q = e_{\widehat{\triangle}^q, \widehat{\mu}^q}$ defined on $(0, +\infty) \times \widehat{M}^q \times \widehat{M}^q$, associated with the operator $\widehat{\triangle}^q$ and with the measure $\widehat{\mu}^q$ (see Appendix A for reminders on heat kernels). Smoothness follows from the fact that $e_{\triangle,\mu}$ is solution of $Pe_{\triangle,\mu} = 0$, with $P = 2\partial_t - (\triangle)_q - (\triangle^*)_{q'}$ that is hypoelliptic.

When $\mathbb{V}$ is bounded on M, it follows from the maximum principle for hypoelliptic operators (see [Bon69]) that $e_{\triangle,\mu}$ and $\widehat{e}^q$ are positive functions. Since $-\widehat{\triangle}^q$ is selfadjoint, $\widehat{e}^q$ is also symmetric. Note also that, using the formulas in Appendix A, we have the homogeneity property

$$(3.3) \qquad \widehat{e}^q(t, x, x') = |\varepsilon|^{\mathcal{Q}(q)}\, \widehat{e}^q\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right)$$

for all $(t, x, x') \in (0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ and for every $\varepsilon \neq 0$ (where we have identified $\widehat{M}^q \simeq \mathbb{R}^n$).

In Theorem A further, we will establish an asymptotic relationship between the heat kernel $e_{\triangle,\mu}$ and the nilpotentized heat kernel $\widehat{e}^q$.

Probabilistic interpretation. It can be noted that, when $\Omega = M$ and $\mathbb{V} = 0$, the heat kernel $e_{\triangle,\mu}$ is the density of the probability law of the solution to the stochastic differential equation on M in the Stratonovich sense

$$dx_{t,q} = \sqrt{2} \sum_{i=1}^{m} X_i(x_{t,q}) \circ dw_t^i + X_0(x_{t,q})\, dt$$

---

[7] Actually, the following weaker Hörmander assumption is enough to ensure hypoellipticity of $\partial_t - \triangle$: TM is spanned by the vector fields $(X_i)_{1 \leqslant i \leqslant m}$, $([X_i, X_j])_{0 \leqslant i, j \leqslant m}$, $([X_i, [X_j, X_k]])_{0 \leqslant i, j, k \leqslant m}$, etc.

with $x_{0,q} = q$, where $(w_t^i)_{1 \leqslant i \leqslant m}$ is a m-dimensional Brownian process realized as the coordinate process on $\{u \in C([0,1], \mathbb{R}^m) \mid u(0) = 0\}$ under the Wiener measure (see [BA89]): the solution to $\partial_t u - \triangle u = 0$ for $t > 0$, $u(0, q) = f(q)$ with $f \in L^2(\Omega, \mu)$, is then given by

$$u(t, q) = \int_\Omega e_{\triangle, \mu}(t, q, q') f(q') \, d\mu(q') = \mathbb{E} f(x_{t,q}).$$

3.2. Two general results for parameter-dependent hypoelliptic heat kernels

This section can be read independently of the rest.

Let M be a smooth connected manifold and let Ω be an open subset of M. Let $m \in \mathbb{N}^*$ and let $\mathcal{K}$ be a compact set. For every $\tau \in \mathcal{K}$, let μ^τ be a smooth density on M, let $X_0^\tau, X_1^\tau, \ldots, X_m^\tau$ be smooth vector fields on M and let $\mathbb{V}^\tau$ be a smooth function on M, all of them depending continuously on τ in C^∞ topology. We consider the second-order differential operator

$$\triangle^\tau = \sum_{i=1}^m (X_i^\tau)^2 + X_0^\tau - \mathbb{V}^\tau.$$

Throughout the section, we assume that the Lie algebra $\mathrm{Lie}(X_1^\tau, \ldots, X_m^\tau)$ generated by the vector fields is equal to $T_q M$ at any point $q \in M$, with a degree of nonholonomy that is uniform with respect to $\tau \in \mathcal{K}$ (*uniform strong Hörmander condition*).

3.2.1. Local nature of the small-time asymptotics of hypoelliptic heat kernels

Let Ω_1 and Ω_2 be two arbitrary open subsets of M, assumed to be manifolds with or without boundary. We still denote by μ^τ the volume induced on each Ω_i.

For $i = 1, 2$, we define the operator $\triangle_i^\tau$ on $L^2(\Omega_i, \mu^\tau)$ as follows. Let $D(\triangle_i^\tau)$ be a subset of $\{u \in L^2(\Omega_i, \mu^\tau) \mid (\triangle^\tau u)_{|\Omega_i} \in L^2(\Omega_i, \mu^\tau)\}$, standing for the domain of $\triangle_i^\tau$ and encoding possible boundary conditions on $\partial \Omega_i$ if Ω_i has a nontrivial boundary. We now consider the operator $\triangle_i^\tau : D(\triangle_i^\tau) \to L^2(\Omega_i, \mu^\tau)$ defined by $\triangle_i^\tau u = (\triangle^\tau u)_{|\Omega_i}$ for every $u \in D(\triangle_i^\tau)$.

In other words, we consider here the operator $\triangle^\tau$ on different subsets Ω_i, with some boundary conditions. For instance, one can take $\Omega_1 = M$ and Ω_2 an open subset of M with Dirichlet conditions on $\partial \Omega_2$.

Let us assume that $\triangle_i^\tau$ generates a strongly continuous semigroup $(e^{t\triangle_i^\tau})_{t \geqslant 0}$ on $L^2(\Omega_i, \mu^\tau)$, satisfying the uniform estimate

$$\left\| e^{t\triangle_i^\tau} \right\|_{L(L^2(\Omega_i, \mu^\tau))} \leqslant \mathrm{Cst} \, e^{Ct} \qquad \forall \, t \geqslant 0 \qquad \forall \, \tau \in \mathcal{K}$$

for some $C > 0$. By an obvious (parameter-dependent) generalization of Lemma 3.1, we note that this is the case if $\mathbb{V}^\tau$ is uniformly bounded below on $\Omega_1 \cup \Omega_2$ with respect to τ and if, for $i = 1, 2$:

(A_τ) either $\triangle_i^\tau : D(\triangle_i^\tau) \to L^2(\Omega_i, \mu^\tau)$ is selfadjoint for every $\tau \in \mathcal{K}$,

(B_τ) or the functions $\mathrm{div}_{\mu^\tau}(X_0^\tau)$, $\mathrm{div}_{\mu^\tau}(X_j^\tau)$ and $X_j^\tau(\mathrm{div}_{\mu^\tau}(X_j^\tau))$, $j = 1, \ldots, m$, are bounded on Ω, uniformly with respect to τ, and we have Dirichlet or Neumann boundary conditions (see Lemma 3.1).

Let $e_i^\tau(t, q, q') = e_{\triangle_i^\tau, \mu^\tau}(t, q, q')$ be its heat kernel, defined on $(0, +\infty) \times \Omega_i \times \Omega_i$. Note that e_i^τ is symmetric whenever $\triangle_i^\tau$ is selfadjoint.

The following fact was noticed in [JSC86]: extending the heat kernels by 0 for $t < 0$, by hypoellipticity of the operator $\partial_t - \triangle_i^\tau$ (under the Lie algebra generating assumption), it follows that $e_i^\tau(t, q, q')$ vanishes at infinite order as $t \to 0$ for fixed q and q' such that $q \neq q'$. This observation inspired to us the result below.

Hereafter, given a function e depending on three variables (t, q, q'), the notation ∂_1 (resp., ∂_2, ∂_3) denotes the partial derivative with respect to t (resp., to q, to q').

THEOREM 3.2. — *For all $(k, \alpha, \beta) \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$, we have*

$$\left(\partial_1^k \partial_2^\alpha \partial_3^\beta e_1^\tau\right)(t, q, q') = \left(\partial_1^k \partial_2^\alpha \partial_3^\beta e_2^\tau\right)(t, q, q') + \mathrm{O}(t^\infty)$$

as $t \to 0$, $t > 0$, uniformly with respect to $\tau \in \mathcal{K}$ and to q and q' varying in any compact subset of $\Omega_1 \cap \Omega_2$. This means that, for all $(k, \alpha, \beta) \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$, for every $t_1 > 0$, for every compact subset $K \subset \Omega_1 \cap \Omega_2$, for every $N \in \mathbb{N}^*$, we have

$$\left|\left(\partial_1^k \partial_2^\alpha \partial_3^\beta (e_1^\tau - e_2^\tau)\right)(t, q, q')\right| \leqslant \mathrm{Cst}(k, \alpha, \beta, t_1, K, N) t^N$$
$$\forall\, t \in (0, t_1] \qquad \forall\, (q, q') \in K \times K \qquad \forall\, \tau \in \mathcal{K}.$$

This result reflects Kac's principle of "not feeling the boundary", showing that the small-time asymptotic behavior of the heat kernel is purely local. Moreover, we establish here a uniform parameter-dependent version, which is possible thanks to the uniform subelliptic estimates obtained in Appendix B.

Proof. — Let $\tau \in \mathcal{K}$ be arbitrary. Let Ω be an open subset of $\Omega_1 \cap \Omega_2$. We set $w^\tau(t, q, q') = e_1^\tau(t, q, q') - e_2^\tau(t, q, q')$, for all $t \in \mathbb{R}$ and $(q, q') \in \Omega \times \Omega$. The function w^τ is smooth on $(0, +\infty) \times \Omega \times \Omega$, and actually (extending by 0 for $t < 0$) we are going to prove that it is smooth on $\mathbb{R} \times \Omega \times \Omega$, with uniform estimates with respect to $\tau \in \mathcal{K}$.

On $\Omega \times \Omega$, we consider the differential operator $(\triangle^\tau)_q = \triangle^\tau \otimes \mathrm{id}$, meaning that given any smooth function g on $\Omega \times \Omega$, the function $(\triangle^\tau)_q g$ designates the partial derivative, using the differential operator $\triangle^\tau$, of the function g, with respect to q. Accordingly, we consider the operator $(\triangle^\tau)_{q'}^* = \mathrm{id} \otimes(\triangle^\tau)^*$.

Noticing that the heat kernels have been extended by 0 for $t < 0$, both kernels e_1^τ and e_2^τ are solutions of the same differential equation $(\partial_t - (\triangle^\tau)_q)e = \delta_{(0, q')}(t, q)$ in the sense of distributions, for any fixed $q' \in \Omega$, where the distribution pairing is considered with respect to the measure $dt \times d\mu^\tau(q)$ on $\mathbb{R} \times \Omega$. Hence $(\partial_t - (\triangle^\tau)_q)w^\tau = 0$ on $\mathbb{R} \times \Omega \times \Omega$. Using that $e_{\triangle^\tau, \mu^\tau}(t, q, q') = e_{\triangle_*^\tau, \mu^\tau}(t, q', q)$, both e_1^τ and e_2^τ are also solutions of $(\partial_t - (\triangle^\tau)_{q'}^*)e = \delta_{(0, q)}(t, q')$ in the sense of distributions, for any fixed $q \in \Omega$. Hence $(\partial_t - (\triangle^\tau)_{q'}^*)w^\tau = 0$ on $\mathbb{R} \times \Omega \times \Omega$. Setting $P_\tau = (\triangle^\tau)_q + (\triangle^\tau)_{q'}^* - 2\partial_t$, we infer that $P_\tau w^\tau = 0$ on $\mathbb{R} \times \Omega \times \Omega$. At this step, for any τ fixed, we infer by hypoellipticity of P_τ that w^τ is smooth, and since w^τ vanishes for $t < 0$, it follows that w^τ is flat at $t = 0$. This gives the result, for τ fixed.

But in order to ensure uniform estimates with respect to τ, we have to elaborate further arguments. In order to use the uniform local subelliptic estimates (B.2) established in Section B.1.1, as an initialization, we need to prove that w^τ is bounded, uniformly with respect to τ, for some weak enough Sobolev norm. To this aim, let us first establish a rough norm estimate, valid for both heat kernels, and uniform with respect to τ.

LEMMA 3.3. — *For every $t_1 > 0$, for every open subset $V \subset \Omega$ of compact closure, there exists $p \in \mathbb{N}^*$ such that*

$$\|e_i^\tau(\cdot, \cdot, \cdot)\|_{L_t^\infty(0,t_1) \times H_q^{-p}(V) \times L_{q'}^2(V)} \leqslant \mathrm{Cst}(t_1, V) \qquad \forall\, \tau \in \mathcal{K} \qquad \forall\, i \in \{1, 2\}.$$

In the norm above, and in the proof hereafter, the Sobolev spaces are considered with respect to the density μ^τ.

Proof of Lemma 3.3. Let $t_1 > 0$ be arbitrary. As in the proofs of Lemma 3.1 (Section 3.1.2) and of Corollary B.3 (Appendix B.1.2), we first note that $\|e^{t\triangle_i^\tau}\|_{L(L^2(\Omega_i))} \leqslant \mathrm{Cst}(t_1)$, for every $t \in [0, t_1]$. Besides, we set $\Lambda_a = a(\mathrm{id} - \triangle_R)^{1/2}$, where $\triangle_R$ is any second-order elliptic operator on Ω_i (for instance, a Riemannian Laplacian if M is Riemannian) and a is a smooth positive function on Ω_i chosen such that there exists $p \in \mathbb{N}^*$ large enough so that Λ_a^{-p} is Hilbert–Schmidt (as an operator on $L^2(\Omega_i)$), i.e., $\|\Lambda_a^{-p}\|_{HS} < +\infty$ (see also the proof of Corollary B.3 in Appendix B.1.2 for the existence of such an integer p). It follows that the operator $\Lambda_a^{-p} e^{t\triangle_i^\tau}$ on $L^2(\Omega_i)$ is Hilbert–Schmidt, and its Hilbert–Schmidt norm is bounded uniformly with respect to $\tau \in \mathcal{K}$ and to $t \in [0, t_1]$. Since

$$\left\|\Lambda_a^{-p} e^{t\triangle_i^\tau}\right\|_{HS} = \left\|(\Lambda_a)_q^{-p} e_i^\tau(t, \cdot, \cdot)\right\|_{L_q^2(\Omega_i) \times L_{q'}^2(\Omega_i)},$$

the conclusion follows. □

Now, let ζ and ζ' be arbitrary smooth functions compactly supported in $(-t_1, t_1) \times \Omega \times \Omega$, with $\zeta' = 1$ on the support of ζ. From Lemma 3.3, there exists $s < 0$ such that $\|\zeta' w^\tau\|_{H^s((-t_1,t_1) \times \Omega \times \Omega)} \leqslant \mathrm{Cst}(t_1, \zeta')$. Applying Theorem B.1 to the family of operators P_τ (in particular, applying to w^τ the uniform estimates (B.2) that follow from this theorem), we infer that, for every $k \in \mathbb{N}$, the norm $\|\zeta w^\tau\|_{H^{s+k\sigma}((-t_1,t_1) \times \Omega \times \Omega)}$ is uniformly bounded with respect to $\tau \in \mathcal{K}$. Using Sobolev embeddings, the theorem follows. □

Remark 3.4. — Note that a quite similar result has been established in [Hsu95], without parameter dependence and under completeness assumptions.

3.2.2. A general convergence result for hypoelliptic heat kernels

We keep the notations and assumptions done in Section 3.1.2. We assume that $\triangle^\tau$ generates a strongly continuous semigroup $(e^{t\triangle^\tau})_{t \geqslant 0}$ on $L^2(\Omega, \mu^\tau)$, satisfying uniform estimate

$$\left\|e^{t\triangle^\tau}\right\|_{L(L^2(\Omega, \mu^\tau))} \leqslant \mathrm{Cst}\; e^{Ct} \qquad \forall\, t \geqslant 0 \qquad \forall\, \tau \in \mathcal{K}$$

for some $C > 0$. Like in Section 3.2.1, we note that this is the case if $\mathbb{V}^\tau$ is uniformly bounded below on M and if either A_τ or B_τ is satisfied. We denote by $e^\tau = e_{\triangle^\tau, \mu^\tau}$ the associated heat kernel, defined on $(0, +\infty) \times \Omega \times \Omega$.

THEOREM 3.5. — *The heat kernel e^τ is smooth on $(0, +\infty) \times \Omega \times \Omega$, for every $\tau \in \mathcal{K}$, and depends continuously on $\tau \in \mathcal{K}$ in $C^\infty((0, +\infty) \times \Omega \times \Omega)$ topology.*

Proof. — Let $\tau_0 \in \mathcal{K}$ be arbitrary. The differential operator $\triangle^{\tau_0}$ is the limit of $\triangle^\tau$ in C^∞ topology as $\tau \to \tau_0$, meaning that $\triangle^\tau f \to \triangle^{\tau_0} f$ as $\tau \to \tau_0$ uniformly on any compact subset of M, for every smooth function f on M.

By the Trotter–Kato theorem (see, e.g., [EN00, Chapter III] or [Paz83, Chapter 3]), $e^{t\triangle^\tau} f \to e^{t\triangle^{\tau_0}} f$ in $L^2(\Omega, \mu^\tau)$ as $\tau \to \tau_0$, for every $t \geqslant 0$ and every smooth function f on Ω with compact support, and the convergence is uniform with respect to t on $[0, t_1]$, for every $t_1 > 0$. Taking the Schwartz kernels (see Appendix A), it follows that, given any $0 < t_0 < t_1$ and any compact subset K of Ω, e^τ converges to e^{τ_0} in $C^{-\infty}([t_0, t_1] \times K \times K)$ as $\tau \to \tau_0$ for the weak-star topology. Here, $C^{-\infty}([t_0, t_1] \times K \times K)$ is the topological dual of the Fréchet Montel space $C^\infty([t_0, t_1] \times K \times K)$.

By Corollary B.3 in Appendix B.1.2, applied with $L^\tau = \triangle^\tau$, the family $(e^\tau)_{\tau \in \mathcal{K}}$ is uniformly bounded in $C^\infty([t_0, t_1] \times K \times K)$. Therefore, thanks to the Heine–Borel property, we conclude that e^τ converges to e^{τ_0} in the Fréchet Montel space $C^\infty((0, +\infty) \times \Omega \times \Omega)$ as $\tau \to \tau_0$. $\qquad\square$

Remark 3.6. — We note that $\triangle^\tau$ is a singular perturbation of $\triangle^0$. One can find in [Lio73] (see also [Hue60]) a number of results on singular perturbations of elliptic operators, i.e., when $\triangle^0$ is an elliptic operator. Here, our results can be seen as some singular perturbations of hypoelliptic operators.

For example, Theorem 3.5 can be applied to the situation where $\triangle^\tau = \triangle^0 + \tau \triangle_R$, with $\triangle^0$ being hypoelliptic and $\triangle_R$ being a Riemannian Laplacian if M is Riemannian: this is an elliptic perturbation of a hypoelliptic operator. We thus recover results established in [Rum00, Theorem 7.2] (see also [Ge93]) in the particular case where $\triangle^0$ is a contact sR Laplacian.

Note also that, when the considered operators are selfadjoint and of compact resolvent, using the max-min principle, our results imply convergence of the spectrum of $\triangle^\tau$ (eigenvalues and eigenfunctions) to that of $\triangle^0$ (as in [Fuk87, Ge93, Rum00]). We do not give details. This convergence is of course not uniform in general because the leading term in the short-time asymptotics of heat kernels may differ: for instance when $\triangle^0$ is a 3D contact sub-Riemannian Laplacian then the short-time asymptotics is like $1/t^2$, whereas for $\tau \neq 0$, assuming that $\triangle^\tau = \triangle^0 + \tau \triangle_R$ as above, the short-time asymptotics is like $1/t^{3/2}$ (asymptotics in the Riemannian case).

4. Proof of the limit in Theorem A

In this section, we prove Theorem B, which is Theorem A at the order zero.

4.1. Preliminaries

Throughout, we assume that X_0 is a smooth section of D over M. Each time this is required, we will indicate the modifications that must be done when X_0 is a smooth section of D^2 over M.

Let $\varepsilon_0 > 0$ be small enough such that $\delta_\varepsilon(V) \subset V$ for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. We first extend the vector fields $(\psi_q)_* X_i$ (which are defined in the neighborhood V given by the chart) to $\mathbb{R}^n$. Let W_1 and W_2 be open subsets of $\mathbb{R}^n$ of compact closure such that $\overline{W}_1 \subset W_2 \subset \overline{W}_2 \subset V$ and such that $\delta_\varepsilon(W_1) \subset W_1$ and $\delta_\varepsilon(W_2) \subset W_2$ for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. Let χ be a smooth function of compact support on $\mathbb{R}^n$, such that $0 \leqslant \chi \leqslant 1$, $\chi(x) = 1$ if $x \in \overline{W}_1$ and $\chi(x) = 0$ if $x \in \mathbb{R}^n \setminus W_2$.

Hereafter, we will use the measure $\widehat{\mu}^q$ on $\mathbb{R}^n$, which coincides, up to a constant scaling, with the Lebesgue measure on $\mathbb{R}^n$.

Definition of $\widetilde{\triangle}$ (local version of $\triangle$ in the chart). We define
$$Y_i = \chi(\psi_q)_* X_i, \qquad i = 0, \ldots, m,$$
so that $Y_i = (\psi_q)_* X_i$ on W_1 and $Y_i = 0$ on $\mathbb{R}^n \setminus W_2$. Similarly, we define the function $\mathbf{v}$ on $\mathbb{R}^n$ by $\mathbf{v} = \chi(\psi_q)_* \mathbb{V}$, and the measure ν on $\mathbb{R}^n$ by $\langle \nu, f \rangle = \langle (\psi_q)_* \mu, \chi f \rangle$ for every $f \in C^0(\mathbb{R}^n)$. Setting $Y = (Y_0, Y_1, \ldots, Y_m)$, we consider on $C^\infty(\mathbb{R}^n)$ the differential operator
$$\widetilde{\triangle} = \sum_{i=1}^m Y_i^2 + Y_0 - \mathbf{v}.$$
The operator $\widetilde{\triangle}$ (resp., the measure ν) is the extension to $\mathbb{R}^n$ (by 0) of the local version of $\triangle$ (resp., of μ) in the chart. As we are going to see, in the proof of Theorem B, the way we extend does not have any impact on the local asymptotics of the heat kernel, thanks to the localization result stated in Theorem 3.2 (Section 3.2.1).

Since the vector fields Y_i are of compact support, setting
$$D(\widetilde{\triangle}) = \left\{ f \in L^2(\mathbb{R}^n, \nu) \,\middle|\, \widetilde{\triangle} f \in L^2(\mathbb{R}^n, \nu) \right\},$$
it follows from Lemma 3.1 that the operator $(\widetilde{\triangle}, D(\widetilde{\triangle}))$ generates a strongly continuous semigroup $(e^{t\widetilde{\triangle}})_{t \geqslant 0}$ on $L^2(\mathbb{R}^n, \nu)$. By hypoellipticity (see Corollary B.3), the Schwartz kernel of $e^{t\widetilde{\triangle}}$, restricted to $(0, +\infty) \times W_1 \times W_1 \to (0, +\infty)$, has a continuous density with respect to ν, which is the smooth function
$$\widetilde{e} = e_{\widetilde{\triangle}, \nu} : (0, +\infty) \times W_1 \times W_1 \to (0, +\infty).$$

4.2. Definition of the vector fields Y_i^ε

For every $\varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\}$, we set

$$(4.1) \qquad \nu^\varepsilon = \frac{1}{|\varepsilon|^{\mathcal{Q}(q)}} \delta_\varepsilon^* \nu, \qquad Y_i^\varepsilon = \varepsilon \delta_\varepsilon^* Y_i, \qquad i = 0, \ldots, m.$$

Using that $\delta_\varepsilon = \delta_{\frac{\varepsilon}{\lambda}} \delta_\lambda = \delta_\lambda \delta_{\frac{\varepsilon}{\lambda}}$, we observe that

$$(4.2) \quad Y_i^\varepsilon = \lambda \delta_\lambda^* Y_i^{\varepsilon/\lambda}, \qquad i = 0, \ldots, m, \qquad \forall \lambda > 0 \qquad \forall \varepsilon \in [-\lambda\varepsilon_0, \lambda\varepsilon_0] \setminus \{0\}.$$

When X_0 is a smooth section of D^2 over M, we modify the definition of Y_0^ε by setting $Y_0^\varepsilon = \varepsilon^2 \delta_\varepsilon^* Y_0$, and we have the homogeneity property $Y_0^\varepsilon = \lambda^2 \delta_\lambda^* Y_0^{\varepsilon/\lambda}$.

Note that Y_i^ε is nontrivial on $\delta_\varepsilon^{-1}(W_1) = \delta_{1/\varepsilon}(W_1)$ which is a neighborhood of 0 increasing to $\mathbb{R}^n$ as $\varepsilon \to 0$. For every $i \in \{0, \ldots, m\}$, Y_i^ε converges to $\widehat{X}_i^q$ in

$C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ as $\varepsilon \to 0$ (see Section 2.5 and Remark 1.2). Actually, using (2.3), we have the expansion

$$(4.3) \qquad Y_i^\varepsilon = \varepsilon \delta_\varepsilon^* Y_i = \widehat{X}_i^q + \varepsilon Y_i^{(0)} + \varepsilon^2 Y_i^{(1)} + \cdots + \varepsilon^N Y_i^{(N-1)} + \mathrm{o}\left(|\varepsilon|^N\right)$$

in C^∞ topology, where $Y_i^{(k)}$ is polynomial and homogeneous of degree k (with respect to dilations), and $\widehat{X}_i^q = Y_i^{(-1)}$, i.e., setting $Y_i^0 = \widehat{X}_i^q$ for $\varepsilon = 0$, Y_i^ε depends smoothly on ε in C^∞ topology.

Since Y_i^ε converges to $\widehat{X}_i^q$, as well as all its derivatives, on any compact, since the m-tuple $(\widehat{X}_1^q, \ldots, \widehat{X}_m^q)$ satisfies the Hörmander condition, using that $[Y_i^\varepsilon, Y_j^\varepsilon] = \varepsilon^2 \delta_\varepsilon^* [Y_i, Y_j]$, it is clear that the m-tuple $(Y_1^\varepsilon, \ldots, Y_m^\varepsilon)$ satisfies the uniform strong Hörmander condition (as defined in Section B.1.2) on W_1, for $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, provided that ε_0 be small enough.

Moreover, we have $\nu^\varepsilon \to \hat{\mu}^q$ for the vague topology as $\varepsilon \to 0$. Actually, the density of ν^ε with respect to the Lebesgue measure of $\mathbb{R}^n$ converges in C^∞ topology to the density of $\hat{\mu}^q$ with respect to the Lebesgue measure of $\mathbb{R}^n$ (which is constant).

4.3. Definition of the operator $\triangle^\varepsilon$

Differential operator $\triangle^\varepsilon$. For every $\varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\}$ we define on $C^\infty(\mathbb{R}^n)$ the differential operator

$$(4.4) \qquad \triangle^\varepsilon = \varepsilon^2 \delta_\varepsilon^* \widetilde{\triangle} (\delta_\varepsilon)_*.$$

Using (4.2), we have the homogeneity property

$$(4.5) \qquad \triangle^\varepsilon = \lambda^2 \delta_\lambda^* \triangle^{\frac{\varepsilon}{\lambda}} (\delta_\lambda)_* \qquad \forall\, \lambda > 0 \qquad \forall\, \varepsilon \in [-\lambda \varepsilon_0, \lambda \varepsilon_0] \setminus \{0\}.$$

Using Appendix A, we have

$$(4.6) \qquad \triangle^\varepsilon = \sum_{i=1}^m \left(\varepsilon \delta_\varepsilon^* Y_i\right)^2 + \varepsilon^2 \delta_\varepsilon^* Y_0 - \varepsilon^2 \left(\delta_\varepsilon^* \mathsf{v}\right) = \sum_{i=1}^m \left(Y_i^\varepsilon\right)^2 + \varepsilon Y_0^\varepsilon - \varepsilon^2 \left(\delta_\varepsilon^* \mathsf{v}\right).$$

When X_0 is a smooth section of D^2 over M, the definition of $\triangle^\varepsilon$ is modified as follows:

$$\triangle^\varepsilon = \sum_{i=1}^m \left(Y_i^\varepsilon\right)^2 + Y_0^\varepsilon - \varepsilon^2 \left(\delta_\varepsilon^* \mathsf{v}\right).$$

Convergence of $\triangle^\varepsilon$ to $\widehat{\triangle}^q$. Since $\varepsilon Y_0^\varepsilon \to 0$ in $C^\infty(\mathbb{R}^n, \mathbb{R}^n)$, the differential operator $\widehat{\triangle}^q$ defined by (2.5) is the limit of $\triangle^\varepsilon$ in C^∞ topology as $\varepsilon \to 0$, meaning that $\triangle^\varepsilon f \to \widehat{\triangle}^q f$ in $C^\infty(\mathbb{R}^n)$ as $\varepsilon \to 0$, for every $f \in C^\infty(\mathbb{R}^n)$. Defining $\triangle^0 = \widehat{\triangle}^q$ for $\varepsilon = 0$, $\triangle^\varepsilon$ depends smoothly on ε in C^∞ topology.

We could give an asymptotic expansion of $\triangle^\varepsilon$ in C^∞ topology, as we will do further in Section 6.1.3 for an appropriate modification of $\triangle^\varepsilon$, but we do not give it because it will not be useful. Indeed, we will see further that the C^∞ topology is not strong enough to establish the complete asymptotic expansion stated in Theorem A. Anyway, the limit in C^∞ topology suffices to establish Theorem B.

When X_0 is a smooth section of D^2 over M, we have $Y_0^\varepsilon = \varepsilon^2 \delta_\varepsilon^* Y_0 = \widehat{X}_0^q + \varepsilon Y_0^{(-1)} +$ o$(|\varepsilon|)$ where $\widehat{X}_0^q$ is homogeneous of order -2 (see Remark 1.2 for the details). We obtain in this case that $\triangle^\varepsilon \to \widehat{\triangle}^q + \widehat{X}_0^q$ in C^∞ topology as $\varepsilon \to 0$.

Semigroup generated by $\triangle^\varepsilon$. Setting

$$D(\triangle^\varepsilon) = \left\{ f \in L^2\left(\mathbb{R}^n, \nu^\varepsilon\right) \middle| \triangle^\varepsilon f \in L^2\left(\mathbb{R}^n, \nu^\varepsilon\right) \right\},$$

using (4.4), the operator $(\triangle^\varepsilon, D(\triangle^\varepsilon))$ generates a strongly continuous semigroup $(e^{t\triangle^\varepsilon})_{t \geqslant 0}$ on $L^2(\mathbb{R}^n, \nu^\varepsilon)$, satisfying

$$\delta_\varepsilon^* e^{\varepsilon^2 t \widetilde{\triangle}} (\delta_\varepsilon)_* = e^{t\triangle^\varepsilon} \qquad \forall\, t \geqslant 0 \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\}.$$

Since $\|e^{t\widetilde{\triangle}}\|_{L(L^2(\mathbb{R}^n, \nu))} \leqslant \mathrm{Cst}\ e^{Ct}$ for every $t \geqslant 0$, for some $C > 0$, it follows that the semigroup $(e^{t\triangle^\varepsilon})_{t \geqslant 0}$ satisfies the uniform estimate $\|e^{t\triangle^\varepsilon}\|_{L(L^2(\mathbb{R}^n, \nu^\varepsilon))} \leqslant \mathrm{Cst}\ e^{C\varepsilon^2 t}$.

Heat kernel e^ε of $\triangle^\varepsilon$. By hypoellipticity (see Corollary B.3), the Schwartz kernel of $e^{t\triangle^\varepsilon}$, restricted to $(0, +\infty) \times W_1 \times W_1$, has a continuous density with respect to ν^ε, which is the smooth function

$$e^\varepsilon = e_{\triangle^\varepsilon, \nu^\varepsilon} : (0, +\infty) \times W_1 \times W_1 \to (0, +\infty).$$

Using (4.4) and the formulas (A.1) of Appendix A, we have

$$(4.7) \quad e^\varepsilon(t, x, x') = |\varepsilon|^{\mathcal{Q}(q)} \widetilde{e}\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right)$$

$$\forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\} \qquad \forall\, (t, x, x') \in (0, +\infty) \times W_1 \times W_1.$$

Convergence of e^ε to $\widehat{e}^q$. Recall that the nilpotentized heat kernel is the smooth function $\widehat{e}^q = e_{\widehat{\triangle}^q, \widehat{\mu}^q} : (0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n \to (0, +\infty)$, defined as the continuous density with respect to $\widehat{\mu}^q$ of the Schwartz kernel of $e^{t\widehat{\triangle}^q}$.

Applying the general convergence result stated in Theorem 3.5 (in Section 3.2.2) with $\mathcal{K} = [-\varepsilon_0, \varepsilon_0]$, $\tau = \varepsilon$, $\Omega = W_1$, $\mu^\tau = \nu^\varepsilon$ and $L^\tau = \triangle^\varepsilon$, we obtain that

$$(4.8) \qquad\qquad e^\varepsilon \underset{\varepsilon \to 0}{\longrightarrow} \widehat{e}^q \qquad \text{in} \quad C^\infty\left((0, +\infty) \times W_1 \times W_1\right).$$

When X_0 is a smooth section of D^2 over M, the result remains true provided that $\widehat{\triangle}^q$ be replaced with $\widehat{\triangle}^q + \widehat{X}_0^q$.

Remark 4.1. — The above argument yields a convergence that is much stronger than the convergence on semigroups provided by the Trotter–Kato theorem (which is only pointwise). We do not know if this could have been established by general results on analytic semigroups. Indeed, although the strongly continuous contraction semigroup $(e^{t\widetilde{\triangle}^q})_{t \geqslant 0}$ is analytic of angle $\pi/2$ (this follows, e.g., from [EN00, Chapter II, Corollary 4.7], because the operator $\widehat{\triangle}^q$ on $L^2(\mathbb{R}^n)$ is nonpositive selfadjoint and thus has a real nonpositive spectrum), given any $\varepsilon > 0$, we do not know if the strongly continuous semigroup $(e^{t\triangle^\varepsilon})_{t \geqslant 0}$ on $L^2(\mathbb{R}^n)$ is analytic with an angle that would be uniform with respect to ε in general (unless, of course, we are in the case where $\triangle^\varepsilon$ is selfadjoint). Actually, there are hints (see [EH03]) showing that the

operator $\triangle^\varepsilon$ may fail to be uniformly sectorial.[8] Note also that $\triangle^\varepsilon - \widehat{\triangle}^q$ *is not* $\widehat{\triangle}^q$-bounded in general and that $\widehat{\triangle}^q$ is not an elliptic operator. This is why, instead of using classical integral representations of analytic semigroups, we used the fact (proved in Section B.1.2) that $e^{t\triangle^\varepsilon}$ is uniformly locally smoothing.

4.4. End of the proof of Theorem B

We already know that e^ε is related to $\widetilde{e}$ by the formula (4.7), which gives as well

$$(4.9) \quad \left(\partial_1^k \partial_2^\alpha \partial_3^\beta \, e_{\triangle^\varepsilon, \nu^\varepsilon}\right)(t, x, x')$$
$$= |\varepsilon|^{\mathcal{Q}(q)+2k} \varepsilon^{\sum_{i=1}^n (\alpha_i+\beta_i)\, w_i(q)} \left(\partial_1^k \partial_2^\alpha \partial_3^\beta \, e_{\widetilde{\triangle}, \nu}\right)\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right)$$

for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\}$, for all $(t, x, x') \in (0, +\infty) \times W_1 \times W_1$ and for all $(k, \alpha, \beta) \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$, where we have set $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_n)$. Let us now relate $\widetilde{e}$ with $e_{\triangle, \mu}$. This is done thanks to the localization result stated in Theorem 3.2 (Section 3.2.1). Recalling that $\widetilde{\triangle}$ coincides with $\triangle$ in the chart, Theorem 3.2 gives, in the chart,

$$(4.10) \qquad e_{\triangle, \mu}\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) = e_{\widetilde{\triangle}, \nu}\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) + \mathrm{O}(|\varepsilon|^\infty)$$

in C^∞ topology. The limit (2.8) in C^∞ topology then follows from (4.8), (4.9) and (4.10).

Case q regular. Let us assume that q is regular, meaning that there exists an open subset U of M on which the flag is regular. Let us make vary q in U. We first remark that it is possible to choose, at the beginning of the proof, a chart ψ_q depending smoothly on q: for instance, one may use the map (2.2) that is obtained with a frame of vector fields Z_i^q adapted to the flag and depending smoothly on q. With such a choice, Y_i^ε depends smoothly on q, and the convergence of Y_i^ε to $\widehat{X}_i^q$ is uniform as well with respect to q. Similar properties hold for all convergences under consideration in the proof. Since all our results on subelliptic estimates (Appendix B) and localization of the heat kernels (hypoelliptic Kac's principle, Theorem 3.2) are valid uniformly with respect to q in this regular neighborhood, we can keep track of the regularity with respect to q in the entire proof above, and smoothness with respect to q of all the coefficients of the expansion follows.

A similar argument is developed on strata defined according to the sR flag.

---

[8] We thank Martin Hairer for a discussion on this subject.

Part 2. Proof of the complete asymptotic expansion in Theorem A

Theorem B is a weaker version of Theorem A, in which we have obtained the limit, i.e., the first term of the expansion with respect to ε. The full version of Theorem A states an asymptotic expansion at any order with respect to ε.

5. Idea of the proof

Surprisingly, the proof of Theorem A, done in Section 6, is much more difficult than the (quite easy) one of Theorem B done in Section 4. Deriving the complete expansion indeed requires significant additional work. In particular, as we explain hereafter, it requires to use global smoothing estimates (established in Appendix C and in Appendix B.2) and to consider an adequate modification $\triangle^{\varepsilon,\gamma}$ of the operator $\triangle^{\varepsilon}$, which complicates significantly the analysis.

Hereafter, we explain our proof approach and we point out the main difficulties, in order to motivate some of the developments that will follow.

5.1. Duhamel formula

Consider the operator $\triangle^{\varepsilon}$ defined by (4.4) in Section 4. As in [Bar13], the starting point is the Duhamel formula

$$e^{t\,\triangle^{\varepsilon}} = e^{t\,\widehat{\triangle}^q} + \int_0^t e^{(t-s)\triangle^{\varepsilon}} \left(\triangle^{\varepsilon} - \widehat{\triangle}^q\right) e^{s\widehat{\triangle}^q}\,ds$$

for $t > 0$. Setting

$$\Sigma_i(t) = \left\{ s^{i+1} = (s_1, \ldots, s_{i+1}) \in (0, +\infty)^{i+1} \;\middle|\; \sum_{k=1}^{i+1} s_k = t \right\} \qquad \forall\, i \in \mathbb{N}^*,$$

given any $N \in \mathbb{N}^*$ we obtain by iteration

$$(5.1) \quad e^{t\triangle^{\varepsilon}} = e^{t\widehat{\triangle}^q} + \sum_{i=1}^N \int_{\Sigma_i(t)} e^{s_1\widehat{\triangle}^q} \left(\triangle^{\varepsilon} - \widehat{\triangle}^q\right) e^{s_2\widehat{\triangle}^q} \cdots \left(\triangle^{\varepsilon} - \widehat{\triangle}^q\right) e^{s_{i+1}\widehat{\triangle}^q}\,ds^{i+1}$$

$$+ \int_{\Sigma_{N+1}(t)} e^{s_1\triangle^{\varepsilon}} \left(\triangle^{\varepsilon} - \widehat{\triangle}^q\right) e^{s_2\,\widehat{\triangle}^q} \cdots \left(\triangle^{\varepsilon} - \widehat{\triangle}^q\right) e^{s_{N+2}\widehat{\triangle}^q}\,ds^{N+2}.$$

Besides, using an expansion in homogeneous terms, we have an asymptotic expansion at any order

$$(5.2) \qquad\qquad \triangle^{\varepsilon} = \widehat{\triangle}^q + \varepsilon\mathcal{A}_1 + \cdots + \varepsilon^N \mathcal{A}_N + \varepsilon^{N+1}\mathcal{R}_{N+1}^{\varepsilon}$$

where $\mathcal{A}_i$, $i \in \mathbb{N}^*$, and $\mathcal{R}_{N+1}^{\varepsilon}$ are second-order differential operators. Moreover all derivations $\mathcal{A}_i$, for $i = 1, \ldots, N$, have polynomial coefficients with a degree that is bounded by some power of N. Of course, we must be careful with the topology taken for the convergences and for the asymptotic expansions, and by the way, this is one of the main problems, because the C^{∞} topology, which was considered previously, will not be sufficient. Let us explain why.

Using (5.1) and (5.2), for the moment in a formal way, we obtain

$$(5.3) \qquad e^{t\triangle^\varepsilon} = e^{t\widehat{\triangle}^q} + \varepsilon \mathcal{C}_1(t) + \cdots + \varepsilon^N \mathcal{C}_N(t) + \varepsilon^{N+1} \mathcal{P}^\varepsilon_{N+1}(t)$$

where each operator $\mathcal{C}_j(t)$ is a finite sum of terms $\mathcal{I}_i(t)$ for some $i \in \{1, \ldots, N\}$, where $\mathcal{I}_i(t)$ is defined by

$$(5.4) \qquad \mathcal{I}_i(t) = \int_{\Sigma_i(t)} e^{s_1 \widehat{\triangle}^q} \mathcal{A}_{j_1} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{A}_{j_i} e^{s_{i+1} \widehat{\triangle}^q} \, ds^{i+1}$$

with $j_1, \ldots, j_N \in \{1, \ldots, N\}$, and where the remainder term $\mathcal{P}^\varepsilon_{N+1}(t)$ is a finite sum of terms $\varepsilon^k \mathcal{I}_i(t)$, $\varepsilon^k \mathcal{J}^\varepsilon_i(t)$ and $\varepsilon^k \mathcal{K}^\varepsilon_i(t)$ with $k, i \in \mathbb{N}$, $k \leqslant (N+1)^2$, $1 \leqslant i \leqslant N$, and $\mathcal{J}^\varepsilon_i(t)$ and $\mathcal{K}^\varepsilon_i(t)$ are defined by

$$(5.5) \qquad \mathcal{J}^\varepsilon_i(t) = \int_{\Sigma_i(t)} e^{s_1 \widehat{\triangle}^q} \mathcal{B}^\varepsilon_1 e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{B}^\varepsilon_i e^{s_{i+1} \widehat{\triangle}^q} \, ds_{i+1}$$

$$(5.6) \qquad \mathcal{K}^\varepsilon_i(t) = \int_{\Sigma_i(t)} e^{s_1 \triangle^\varepsilon} \mathcal{B}^\varepsilon_1 e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{B}^\varepsilon_i e^{s_{i+1} \widehat{\triangle}^q} \, ds^{i+1}$$

where each $\mathcal{B}^\varepsilon_j$ is a second-order derivation, either equal to some $\mathcal{A}_i$, $i \in \{1, \ldots, N\}$, or to $\mathcal{R}^\varepsilon_{N+1}$.

All operators above are defined as convolutions, i.e., iterated compositions involving the operators $e^{s_i \widehat{\triangle}^q}$ and $e^{s_1 \triangle^\varepsilon}$, and derivations $\mathcal{A}_i$ and $\mathcal{R}^\varepsilon_{N+1}$ in-between. For instance, we have

$$\mathcal{C}_1(t) = \int_0^t e^{(t-s)\widehat{\triangle}^q} \mathcal{A}_1 e^{s\widehat{\triangle}^q} \, ds, \qquad \mathcal{C}_2(t) = \int_0^t e^{(t-s)\widehat{\triangle}^q} \left(\mathcal{A}_2 e^{s\widehat{\triangle}^q} + \mathcal{A}_1 \mathcal{C}_1(s) \right) ds.$$

Only the terms $\mathcal{K}^\varepsilon_i(t)$ (involved in the remainder $\mathcal{P}^\varepsilon_{N+1}(t)$) contain $e^{s_1 \triangle^\varepsilon}$ as a first term in the convolution. Note that $\widehat{\triangle}^q$ is selfadjoint but $\triangle^\varepsilon$ is not selfadjoint in general.

The basic idea is then to take Schwartz kernels in (5.3), in order to obtain the expansion of the heat kernel e^ε with respect to ε.

Although apparently simple, at least in a formal way, establishing rigorously the expansion at any order appears to be difficult and technical. The main difficulty is to give a sense to the expansion (5.3) with respect to some appropriate topology. The asymptotic expansion (5.3) will be written in the sense of smoothing operators (see Proposition 6.7 in Section 6.2), i.e., operators that map continuously any $H^j_{loc}(\mathbb{R}^n)$ to any $H^k_{loc}(\mathbb{R}^n)$, with a norm that is uniformly bounded with respect to ε. More precisely, let $0 < t_0 < t_1$ be fixed. We would like to prove that

$$(5.7) \quad \|\chi_1 \mathcal{I}_i(t) \chi_2\|_{L\left(H^j(\mathbb{R}^n), H^k(\mathbb{R}^n)\right)} \leqslant \text{Cst}\,(N, \chi_1, \chi_2, j, k, t_0, t_1)$$

$$\forall\, t \in [t_0, t_1] \qquad \forall\, \chi_1, \chi_2 \in C^\infty_c(\mathbb{R}^n) \qquad \forall\, j, k \in \mathbb{Z} \qquad \forall\, i \in \{1, \ldots, N\}$$

and that there exists $\varepsilon_0 > 0$ such that

$$(5.8) \quad \left\|\chi_1 \mathcal{J}^\varepsilon_{N+1}(t) \chi_2\right\|_{L\left(H^j(\mathbb{R}^n), H^k(\mathbb{R}^n)\right)} + \left\|\chi_1 \mathcal{K}^\varepsilon_{N+1}(t) \chi_2\right\|_{L\left(H^j(\mathbb{R}^n), H^k(\mathbb{R}^n)\right)}$$

$$\leqslant \text{Cst}\,(N, \chi_1, \chi_2, j, k, t_0, t_1)$$

$$\forall\, t \in [t_0, t_1] \qquad \forall\, \chi_1, \chi_2 \in C^\infty_c(\mathbb{R}^n) \qquad \forall\, j, k \in \mathbb{Z} \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

To prove these smoothing properties, we will prove that the chain of compositions appearing in the integrals (5.4), (5.5) and (5.6), involving the operators $e^{s_i \widehat{\triangle}^q}$ and $e^{s_1 \triangle^\varepsilon}$ and the derivations $\mathcal{A}_i$ and $\mathcal{R}^\varepsilon_{N+1}$, is performed in the scale of Sobolev spaces with polynomial weight $H^\alpha_\beta(\mathbb{R}^n)$ (whose definition is recalled in Appendix B.2), and that at least one of the operators $e^{s_i \widehat{\triangle}^q}$ and $e^{s_1 \triangle^\varepsilon}$ enjoys a strong smoothing property, able to map continuously any $H^\alpha_\beta(\mathbb{R}^n)$ to $H^{\alpha'}_{\beta'}(\mathbb{R}^n)$ for any $\alpha' \in \mathbb{R}$ and for some appropriate $\beta' \in \mathbb{R}$ (and this, uniformly with respect to ε).

Here, in contrast to Section 4, *local* subelliptic estimates are not enough and *global* subelliptic estimates are required, in order to establish such global smoothing properties. Since the derivations appearing in the chain of compositions can be arbitrary (they have no specific relationship with the operators $\widehat{\triangle}^q$ and $\triangle^\varepsilon$), the use of Sobolev spaces with polynomial weight $H^\alpha_\beta(\mathbb{R}^n)$ appears to be relevant.

Hereafter, we list the properties that we will have to establish in order to prove (5.7) and (5.8). Note that (5.7) involves operators $e^{s_i \widehat{\triangle}^q}$ and derivations $\mathcal{A}_i$, not depending on ε, while (5.8) involves also the operator $e^{s_1 \triangle^\varepsilon}$ and the derivation $\mathcal{R}^\varepsilon_{N+1}$ for which we will have to establish properties that are uniform with respect to ε. As we will see in Section 5.3, this will raise a significant additional difficulty.

5.2. Requirements to prove (5.7)

To prove (5.7), we observe that, inside the integral (5.4), at least one of the real numbers s_p is such that $s_p \geqslant \frac{t_0}{N}$. We want the corresponding operator $e^{s_p \widehat{\triangle}^q}$ to be globally smoothing. To this aim, we will need to establish the following property:

($\mathbf{P}_1$) *Global smoothing property in Sobolev spaces with polynomial weight*: there exists $k_0 \in \mathbb{N}$ such that

$$\left\| e^{\tau \widehat{\triangle}^q} \right\|_{L\left(H^\alpha_\beta(\mathbb{R}^n), H^{\alpha'}_{\beta/k_0 - k_0(|\alpha| + |\alpha'|)}(\mathbb{R}^n)\right)} \leqslant \mathrm{Cst}\,(\alpha, \alpha', \beta, \tau_0)$$

$$\forall\, \alpha, \alpha', \beta \in \mathbb{R} \qquad \forall\, \tau_0 \in (0, 1) \qquad \forall\, \tau \in [\tau_0, 1].$$

In other words, in positive times ($\tau_0 \leqslant \tau \leqslant 1$) the operator $e^{\tau \widehat{\triangle}^q}$ gains differential regularity, with a controlled loss of polynomial weight regularity.

Besides, all other operators in the composition in the integral (5.4) are either $e^{s_i \widehat{\triangle}^q}$, with $0 \leqslant s_i \leqslant 1$ or derivations $\mathcal{A}_i$. The second-order derivations $\mathcal{A}_i$ for $i = 1, \ldots, N$, because they have polynomial coefficients with maximal degree, say, N^ℓ for some $\ell \in \mathbb{N}^*$, and thus map continuously any $H^\alpha_\beta(\mathbb{R}^n)$ to $H^{\alpha-2}_{\beta-N^\ell}(\mathbb{R}^n)$. Concerning $e^{s_i \widehat{\triangle}^q}$, we require the following property:

($\mathbf{P}_2$) *Continuity with controlled loss in the Sobolev spaces with polynomial weight, including time zero*: there exist $\beta_0 > 0$ and $k_1 \in \mathbb{N}$ such that

$$\left\| e^{\tau \widehat{\triangle}^q} \right\|_{L\left(H^\alpha_\beta(\mathbb{R}^n), H^{\alpha-k_1|\alpha|}_{\beta/k_1 - k_1|\alpha| - k_1}(\mathbb{R}^n)\right)} \leqslant \mathrm{Cst}(\alpha, \beta) \quad \forall\, \alpha \in \mathbb{R} \quad \forall\, \beta \geqslant \beta_0 \quad \forall\, \tau \in [0, 1].$$

Note that, in contrast to $(\mathbf{P}_1)$ which is a smoothing property in positive time, $(\mathbf{P}_2)$ is stated on the time interval $[0, 1]$, including time zero. The property $(\mathbf{P}_2)$ is inferred from the following two properties that we will establish:

$(\mathbf{P}_3)$ Let $\widehat{\mathcal{D}}_{k,\beta}^{q,\mathrm{sR}}$ be the completion of $C_c^\infty(\mathbb{R}^n)$ for the norm $\|\langle x \rangle_{\mathrm{sR}}^\beta (\mathrm{id} - \widehat{\triangle}^q)^k u\|_{L^2(\mathbb{R}^n)}$, which is the domain of $(\mathrm{id} - \widehat{\triangle}^q)^k$ polynomially weighted with the power β of the sR Japanese bracket (see a precise definition in Appendix C.2). We have

$$\left\| e^{\tau \widehat{\triangle}^q} \right\|_{L\left(\widehat{\mathcal{D}}_{k,\beta}^{q,\mathrm{sR}}\right)} \leqslant \mathrm{Cst}(k, \beta) \qquad \forall\, \tau \in [0, 1] \qquad \forall\, k \in \mathbb{Z} \qquad \forall\, \beta \geqslant 1.$$

$(\mathbf{P}_4)$ *Continuous embeddings*: there exist $\sigma > 0$ and $N_0 \in \mathbb{N}$ such that

$$H_{\beta + 2kN_0}^{2k}(\mathbb{R}^n) \hookrightarrow \widehat{\mathcal{D}}_{k,\beta}^{q,\mathrm{sR}} \hookrightarrow H_{\beta/r(q) - kN_0}^{k\sigma}(\mathbb{R}^n) \qquad \forall\, k \in \mathbb{N} \qquad \forall\, \beta \in \mathbb{R}$$

and by duality,

$$H_{\beta/r(q) + kN_0}^{-k\sigma}(\mathbb{R}^n) \hookrightarrow \widehat{\mathcal{D}}_{-k,\beta}^{q,\mathrm{sR}} \hookrightarrow H_{\beta - 2kN_0}^{-2k}(\mathbb{R}^n) \qquad \forall\, k \in \mathbb{N} \qquad \forall\, \beta \in \mathbb{R}.$$

The smoothing property of $\mathcal{I}_i(t)$ follows from $(\mathbf{P}_1)$, $(\mathbf{P}_3)$ and $(\mathbf{P}_4)$ (the full detail of the argument will be given in Proposition (6.7), but one can already note the important fact that β must be taken large enough).

The property $(\mathbf{P}_4)$ follows from *global* subelliptic estimates, that we establish in Appendix B.2 for general Hörmander operators whose coefficients (as well as their derivatives) have a growth at infinity that is at most polynomial (of course, there, the polynomial property is crucial). Actually, in Appendix B.2 we will establish $(\mathbf{P}_4)$ for the domains $\widehat{\mathcal{D}}_{k,\beta}^q$ of $(\mathrm{id} - \widehat{\triangle}^q)^k$ polynomially weighted with the power β of the usual (not sR) Japanese bracket (see their definition in Appendix B.2.1). But, since we have the inequality

$$\mathrm{Cst}\, \| \cdot \|_{\widehat{\mathcal{D}}_{j,\alpha/r(q)}^q} \leqslant \| \cdot \|_{\widehat{\mathcal{D}}_{j,\alpha}^{q,\mathrm{sR}}} \leqslant \mathrm{Cst}\, \| \cdot \|_{\widehat{\mathcal{D}}_{j,\alpha}^q}$$

for all $j \in \mathbb{Z}$ and $\alpha > 0$ (see (C.5) in Section C.2), $(\mathbf{P}_4)$ follows.

The property $(\mathbf{P}_1)$ is inferred from $(\mathbf{P}_4)$ and from the following property:

$(\mathbf{P}'_1)$ *Global smoothing property in the iterated domains with polynomial weight*: there exists $k_0 \in \mathbb{N}$ such that

$$\left\| e^{\tau \widehat{\triangle}^q} \right\|_{L\left(\widehat{\mathcal{D}}_{j,\beta}^{q,\mathrm{sR}}, \widehat{\mathcal{D}}_{k,\beta}^{q,\mathrm{sR}}\right)} \leqslant \mathrm{Cst}(j, k, \beta, \tau_0)$$

$$\forall\, j, k \in \mathbb{Z} \qquad \forall\, \beta \geqslant 1 \qquad \forall\, \tau_0 \in (0, 1) \qquad \forall\, \tau \in [\tau_0, 1].$$

While $(\mathbf{P}'_1)$ is a smoothing property (valid for positive times), the property $(\mathbf{P}_3)$, which must hold also at $\tau = 0$ but does not provide any gain of regularity, is equivalent to the fact that $(e^{t\widehat{\triangle}^q})_{t \geqslant 0}$ is a semigroup on the weighted Hilbert space $L_\beta^2(\mathbb{R}^n)$, for every $\beta \geqslant 1$. These facts are not obvious. We prove $(\mathbf{P}'_1)$ and $(\mathbf{P}_3)$ in Appendix C (more precisely, see Proposition C.3 in Appendix C.2) by using upper exponential estimates of the heat kernel of the nilpotent sR Laplacian $\widehat{\triangle}^q$.

At this step, we have realized that, to prove that the operators $\chi_1 \mathcal{I}_i(t) \chi_2$ are smoothing, we have to establish some properties for the operator $\widehat{\triangle}^q$ that are of a global nature. To prove them, we will use the instrumental facts that the operator $\widehat{\triangle}^q$

is selfadjoint and polynomial. This is the objective of Appendix C. Selfadjointness allows us to use the spectral theorem.

Remark 5.1. — It seems unavoidable in general to use Sobolev spaces with polynomial weight $H_\beta^\alpha(\mathbb{R}^n)$, in the above chain of arguments. Indeed, the second-order derivations $\mathcal{A}_i$ can be arbitrary, in the sense that they "cannot be factorized" by $\widehat{\triangle}^q$ and thus they do not map an iterated domain of $\widehat{\triangle}^q$ to another in general.

5.3. Requirements to prove (5.8)

Let us now search what properties are required to prove (5.8). Compared with the operator $\mathcal{J}_i(t)$ defined by (5.4):

- in the definitions (5.5) and (5.6) of $\mathcal{J}_i^\varepsilon(t)$ and $\mathcal{K}_i^\varepsilon(t)$, the derivations in-between can be either $\mathcal{A}_j$ or $\mathcal{R}_{N+1}^\varepsilon$;
- in the definition (5.6) of $\mathcal{K}_i^\varepsilon(t)$, we have a final composition by the operator $e^{s_0\triangle^\varepsilon}$ in the integral.

Note that, in contrast to $\widehat{\triangle}^q$, the operator $\triangle^\varepsilon$ is not selfadjoint in general.

In order to perform a reasoning as above, it would be desirable that there exist $\ell \in \mathbb{N}$ and $\varepsilon_0 > 0$ such that:

(i) $\|\mathcal{R}_{N+1}^\varepsilon\|_{L(H_\beta^\alpha(\mathbb{R}^n), H_{\beta-N\ell}^{\alpha-2}(\mathbb{R}^n))} \leqslant \mathrm{Cst}(\alpha, \beta) \qquad \forall\, \alpha, \beta \in \mathbb{R} \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0]$;

(ii) the operator $\chi_1 e^{\tau\triangle^\varepsilon}$ satisfy the properties ($\mathbf{P}_1$) (global smoothing in the Sobolev spaces with polynomial weight) and ($\mathbf{P}_2$) (continuity with controlled loss, including time zero), uniformly with respect to $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.

However, except in the following particular case, we are going to see that these properties are not satisfied in general, which raises a serious difficulty that we will show how to overcome.

A particular case. It is interesting to note that (i) and (ii) are satisfied under the following additional assumptions:

- $M = \mathbb{R}^n$;
- the vector fields $X_0, X_1, \ldots, X_m$ are polynomial;
- the operator $\triangle$ defined by (1.1) is selfadjoint (see Section 3.1.1).

Indeed, under these assumptions the coefficients of $\triangle^\varepsilon$ are polynomial, and thus the homogeneous expansion (5.2) is exact for N large enough, i.e., $\mathcal{R}_{N+1}^\varepsilon = 0$, hence (i) is satisfied. As for (ii), since $\triangle^\varepsilon$ is also selfadjoint, all results established in Appendix C can straightforwardly be extended to such a one-parameter family of operators. In particular, Proposition C.3 and continuous embeddings give (ii).

Hence, under the above additional assumptions, the results established in Appendix C are sufficient to conclude that $\chi_1 \mathcal{J}_{N+1}^\varepsilon(t)\chi_2$ is smoothing and then complete the proof of Theorem A.

Difficulties in the general case. But in the general case where either the vector fields are not polynomial at infinity, or $\triangle$ is not selfadjoint, the properties (i) and (ii) may fail and we have to proceed differently.

First of all, when $\triangle$ is not selfadjoint, we cannot use the results of Appendix C, but this difficulty is bypassed in Appendix B.2, where we establish some global smoothing properties for general hypoelliptic Hörmander operators (not necessarily selfadjoint) depending on a parameter.

However, this can only be done under the crucial assumption that the growth at infinity of the differential operator is at most polynomial, with a degree that is uniform with respect to the parameter. This requirement of being at most polynomial at infinity is the main constraint. The results of Section B.2 cannot be applied to $e^{t\triangle^\varepsilon}$, because the growth at infinity of the vector fields X_i^ε is *not* at most polynomial uniformly with respect to ε, in general (see Example 6.3 in Section 6.1.1).

This serious flaw is due to the fact that Y_i^ε (defined by (4.1)) converges to $\widehat{X}_i^q$ only in C^∞ topology as $\varepsilon \to 0$, i.e., Y_i^ε converges uniformly, as well as all its derivatives, to $\widehat{X}_i^q$ only on every compact subset of $\mathbb{R}^n$. But this convergence is *not global* in general.

Adding a "damping" parameter γ. To overcome this defect of convergence, we introduce in Section 6.1 a slightly different operator $\triangle^{\varepsilon,\gamma}$, depending on an additional (fixed) parameter $\gamma \in (0,1)$, which satisfies $\triangle^{\varepsilon,\gamma} \to \widehat{\triangle}^q$ as $\varepsilon \to 0$ in a much stronger sense, and whose growth at infinity is uniformly at most polynomial.

To do so, we replace each vector field Y_i^ε with an adequate modification $Y_i^{\varepsilon,\gamma}$: we define the vector field $Y_i^{\varepsilon,\gamma}$ such that, roughly speaking, $Y_i^{\varepsilon,\gamma} = \widehat{X}_i^q$ outside of the sR ball $\widehat{B}_{sR}^q(0, 1/\varepsilon^\gamma)$ and $Y_i^{\varepsilon,\gamma} = Y_i^\varepsilon$ inside the ball. In the key Lemma 6.2, we establish that, while Y_i^ε converges uniformly (as well as all its derivatives) to $\widehat{X}_i^q$ only on every compact, for appropriate (small enough) values of γ, $Y_i^{\varepsilon,\gamma}$ converges to $\widehat{X}_i^q$ globally, on the whole $\mathbb{R}^n$, as $\varepsilon \to 0$, and the convergence is even valid in any space L^p, $p \in [1, +\infty]$, as well as all its derivatives. This much stronger convergence property, obtained thanks to the adequate modification using the parameter γ, is instrumental in our proof. By the way, we think that it could be useful for other purposes.

The parameter γ can be viewed, in some sense, as a "damping" parameter: $Y_i^{\varepsilon,\gamma}$ is an adequate modification of Y_i^ε, which is sufficiently damped (but not too much) to be of polynomial growth at infinity, uniformly with respect to ε.

The construction of $\triangle^{\varepsilon,\gamma}$ is done in Section 6.1. We will prove that the family of vector fields $Y_i^{\varepsilon,\gamma}$ satisfies a uniform polynomial strong Hörmander condition if $\gamma > 0$ is chosen small enough (see Lemma 6.4), thus allowing us to use the uniform global subelliptic estimates established in Section B.2.

Note that, to ensure the validity of the global subelliptic estimates, it could seem sufficient to consider the above truncation on a ball $\widehat{B}_{sR}^q(0,1)$ of fixed radius, rather than on a ball $\widehat{B}_{sR}^q(0, 1/\varepsilon^\gamma)$. But then, the end of the proof of Theorem A would fail; more precisely, the application of the final localization argument (hypoelliptic Kac's principle) would fail (see Section 6.4). So, we underline that it is important

to take $\gamma > 0$ small and thus keep the equality $Y_i^{\varepsilon,\gamma} = Y_i^\varepsilon$ valid on a neighborhood of 0 increasing to $\mathbb{R}^n$ as $\varepsilon \to 0$.

With this modification in mind, we go back to Section 5.1 and we apply the Duhamel formula, replacing $\triangle^\varepsilon$ with $\triangle^{\varepsilon,\gamma}$. This is what we will do in Section 6.2, and we will obtain the expansion (5.3) with $\triangle^{\varepsilon,\gamma}$ instead of $\triangle^\varepsilon$.

We will then want to prove (5.8) with $\triangle^{\varepsilon,\gamma}$ replacing $\triangle^\varepsilon$. Recalling (i) and (ii) as desirable properties, what we will be able to establish is:

($\mathbf{P}_5$) We have $\triangle^{\varepsilon,\gamma} = \widehat{\triangle}^q + \varepsilon \mathcal{A}_1^{\varepsilon,\gamma} + \varepsilon^2 \mathcal{A}_2^{\varepsilon,\gamma} + \cdots + \varepsilon^N \mathcal{A}_N^{\varepsilon,\gamma} + \varepsilon^{N(1-\gamma)+1-\gamma r(q)} \mathcal{R}_{N+1}^{\varepsilon,\gamma}$
and there exist $\beta_0 > 0$ and $\varepsilon_0 > 0$ such that

$$\left\| \mathcal{R}_{N+1}^{\varepsilon,\gamma} \right\|_{L\left(H_\beta^\alpha(\mathbb{R}^n), H_{\beta-\beta_0}^{\alpha-2}(\mathbb{R}^n) \right)} \leqslant \mathrm{Cst}(\alpha,\beta) \qquad \forall\, \alpha, \beta \in \mathbb{R} \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

($\mathbf{P}_6$) *Global-to-local smoothing property*: there exist $k_0 \in \mathbb{N}$ and $\alpha_0 > 0$ such that

$$\left\| \chi_1 e^{\tau \triangle^{\varepsilon,\gamma}} \right\|_{L\left(H_{k_0 \alpha}^{-\alpha}(\mathbb{R}^n), H^\beta(\mathbb{R}^n) \right)} \leqslant \mathrm{Cst}(\alpha, \beta, \tau_0)$$

$$\forall\, \alpha, \beta \geqslant \alpha_0 \qquad \forall\, \tau_0 \in (0,1) \qquad \forall\, \tau \in [\tau_0, 1].$$

We note that, in contrast to the asymptotic expansion (5.2) of $\triangle^\varepsilon$, the introduction of the damping parameter γ implies a loss in the power of ε for the remainder term, in the asymptotic expansion of $\triangle^{\varepsilon,\gamma}$ given in ($\mathbf{P}_5$). The property ($\mathbf{P}_6$) is proved in Appendix B.2.2, thanks to the fact that a uniform polynomial strong Hörmander condition is satisfied as soon as $0 < \gamma < \frac{1}{r(q)(r(q)+1)}$ (see Lemma 6.4).

We show in Proposition 6.7 that the properties ($\mathbf{P}_5$) and ($\mathbf{P}_6$) are sufficient to conclude.

5.4. End of the proof

Finally, we will take Schwartz kernels in Section 6.3, in order to obtain an expansion at any order of the heat kernel $e^{\varepsilon,\gamma}$ associated with $\triangle^{\varepsilon,\gamma}$, in function of the heat kernel $\widehat{e}^q$ of $e^{t\widehat{\triangle}^q}$.

To conclude, it will remain to relate the heat kernels $e^{\varepsilon,\gamma}$ and $e_{\triangle,\mu}$. As in Section 4, the local nature of the small-time asymptotics of heat kernels (hypoelliptic Kac's principle), established in Theorem 3.2 in Section 3, will be instrumental there and will be applied several times (as in the proof of Theorem B). First, by localization, small-time asymptotics of $e_{\triangle,\mu}$ and of the heat kernel of a representation of $\triangle$ in a chart are the same. Second, the heat kernel e^ε is directly related to the kernel of the local representation by homogeneity. Unfortunately, the heat kernel $e^{\varepsilon,\gamma}$ does not satisfy this homogeneity property, and an additional difficulty arises here, which we will solve thanks to the adequate construction of the operator $\triangle^{\varepsilon,\gamma}$. There, the fact that the localization argument can be applied is strongly due to the fact that $0 < \gamma < 1$, more precisely, that $\varepsilon/\varepsilon^\gamma \to 0$ as $\varepsilon \to 0$.

6. Proof of Theorem A

Throughout, we assume that X_0 is a smooth section of D over M. In view of Remark 1.2, each time this is required, we will indicate the modifications that must be done when X_0 is a smooth section of D^2 over M.

We consider the framework and notations introduced in Section 4.

6.1. Construction of $\triangle^{\varepsilon,\gamma}$, with a damping parameter γ

Let $\gamma \in (0,1)$ be fixed, to be chosen later. We assume that $\varepsilon_0 > 0$ is small enough so that $\delta_{\varepsilon^\gamma}(V) \subset V$ for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$.

6.1.1. Definition and properties of the vector fields $Y_i^{\varepsilon,\gamma}$

Considering the subsets W_1, W_2 and the function $\chi \in C^\infty(\mathbb{R}^n)$ introduced in Section 4, for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\}$, we define

$$\nu^{\varepsilon,\gamma} = (\delta_{\varepsilon^\gamma}^* \chi) \nu^\varepsilon + (1 - (\delta_{\varepsilon^\gamma}^* \chi)) \widehat{\mu}^q = \widehat{\mu}^q + (\delta_{\varepsilon^\gamma}^* \chi)(\nu^\varepsilon - \widehat{\mu}^q)$$

and

$$Y_i^{\varepsilon,\gamma} = (\delta_{\varepsilon^\gamma}^* \chi) Y_i^\varepsilon + (1 - \delta_{\varepsilon^\gamma}^* \chi) \widehat{X}_i^q = \widehat{X}_i^q + (\delta_{\varepsilon^\gamma}^* \chi)\left(Y_i^\varepsilon - \widehat{X}_i^q\right), \quad i = 0, \ldots, m.$$

By construction, we have $Y_i^{\varepsilon,\gamma} = Y_i^\varepsilon$ on $\delta_{1/\varepsilon^\gamma}(W_1)$ and $Y_i^{\varepsilon,\gamma} = \widehat{X}_i^q$ on $\mathbb{R}^n \setminus \delta_{1/\varepsilon^\gamma}(W_2)$. Note that $\delta_{1/\varepsilon^\gamma}(W_1) \subset \delta_{1/\varepsilon}(W_1)$ is a neighborhood of 0 increasing to $\mathbb{R}^n$ as $\varepsilon \to 0$. The vector field $Y_i^{\varepsilon,\gamma}$ is thus an adequate restriction of Y_i^ε on a neighborhood increasing to $\mathbb{R}^n$ as $|\varepsilon|$ decreases, and $Y_i^{\varepsilon,\gamma}$ coincides with $\widehat{X}_i^q$ "at infinity", more precisely, outside of the increasing neighborhood $\delta_{1/\varepsilon^\gamma}(W_2)$. The asymptotics in ε^γ, with a $\gamma \in (0,1)$ to be chosen small enough later, will be instrumental several times in the proof. In particular, if we would replace $(\delta_{\varepsilon^\gamma}^* \chi)$ by $(\delta_\lambda^* \chi)$ in the definition of $Y_i^{\varepsilon,\gamma}$, for some $\lambda > 0$ fixed, then, at the very end of the proof, the localization theorem (Theorem 3.2) could not be applied.

Setting $Y_i^{0,\gamma} = \widehat{X}_i^q$ for $\varepsilon = 0$, $Y_i^{\varepsilon,\gamma}$ depends smoothly on ε at $\varepsilon = 0$ in C^∞ topology (this is because Y_i^ε depends smoothly on ε and, on any fixed compact subset $K \subset \mathbb{R}^n$, one has $\delta_{\varepsilon^\gamma}^* \chi = 1$ on K as soon as ε is small enough).

We note that, for $i = 0, \ldots, m$, we have the homogeneity property

$$(6.1) \quad \varepsilon^\beta \delta_{\varepsilon^\beta}^* Y_i^{\varepsilon,\gamma} = Y_i^{\varepsilon^{1+\beta}, \gamma+\beta} \quad \forall \beta \in (-\gamma, 1-\gamma) \quad \forall \varepsilon \in \left[-\varepsilon_0^{1/(1+\beta)}, \varepsilon_0^{1/(1+\beta)}\right] \setminus \{0\}.$$

Remark 6.1. — Given any $i \in \{0, \ldots, m\}$, since $Y_i^{\varepsilon,\gamma}$ coincides with Y_i^ε on the growing (as ε decreases) neighborhood on $\delta_{1/\varepsilon^\gamma}(W_1)$, we obviously have $Y_i^{\varepsilon,\gamma} = Y_i^\varepsilon + O(|\varepsilon|^\infty)$ as $\varepsilon \to 0$ in C^∞ topology.

The parameter $\gamma > 0$ introduced above is used to sufficiently *damp* the growth of the vector field $Y_i^{\varepsilon,\gamma}$ at infinity, with respect to that of the undamped vector field Y_i^ε.

In the next lemma, we prove that, while Y_i^ε converges uniformly (as well as all its derivatives) to $\widehat{X}_i^q$ only on every compact, $Y_i^{\varepsilon,\gamma}$ converges to $\widehat{X}_i^q$ *globally*, on the

whole $\mathbb{R}^n$, as $\varepsilon \to 0$, and the convergence is even valid in any space L^p, $p \in [1, +\infty]$. This much stronger convergence property is due to the adequate modification using the parameter γ. It implies in particular that the growth at infinity of the coefficients of $Y_i^{\varepsilon, \gamma}$, as well as all their derivatives, is at most polynomial, uniformly with respect to ε. This will be crucial in what follows, as motivated and explained in Section 5.3.

LEMMA 6.2. — *Recall that $r(q)$ is the degree of nonholonomy at q (we have $r(q) = w_n(q)$, the largest weight at q, see Section 2.3) and that $\mathcal{Q}(q) = \sum_{i=1}^n w_i(q)$. If $\gamma < \frac{1}{r(q)}$ then:*

(i) *For every $i \in \{0, \dots, m\}$, considering the vector field $Y_i^{\varepsilon, \gamma} - \widehat{X}_i^q$ as a derivation, we have*

$$\left\| \left(Y_i^{\varepsilon, \gamma} - \widehat{X}_i^q \right) f \right\|_{W^{k, p}(\mathbb{R}^n)} \leqslant \mathrm{Cst}(k) |\varepsilon|^{1 - \gamma r(q)} \| f \|_{W^{k+1, p}(\mathbb{R}^n)}$$

$$\forall\, k \in \mathbb{N} \qquad \forall\, p \in [1, +\infty] \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0] \qquad \forall\, f \in C_c^\infty(\mathbb{R}^n).$$

In particular:
   - *Taking $p = +\infty$: $Y_i^{\varepsilon, \gamma}$ converges uniformly on $\mathbb{R}^n$ to $\widehat{X}_i^q$ as $\varepsilon \to 0$ (meaning that all coefficients of the vector field converge uniformly on $\mathbb{R}^n$), as well as all its derivatives.*
   - *Taking $p = 2$: $Y_i^{\varepsilon, \gamma} f$ converges to $\widehat{X}_i^q f$ in $L^2(\mathbb{R}^n)$ as $\varepsilon \to 0$, as well as all its derivatives.*

(ii) *The density of $\nu^{\varepsilon, \gamma}$ with respect to $\widehat{\mu}^q$ (which is a constant times the Lebesgue measure of $\mathbb{R}^n$) converges uniformly on the whole $\mathbb{R}^n$ to 1, and all its derivatives converge uniformly on $\mathbb{R}^n$ to 0. Therefore, for every $s \in \mathbb{R}$ we have $H^s(\mathbb{R}^n, \nu^{\varepsilon, \gamma}) = H^s(\mathbb{R}^n, \widehat{\mu}^q) = H^s(\mathbb{R}^n)$ with respective norms that are equivalent, uniformly with respect to ε.*

(iii) *As in Appendix B.2, we denote by $(Y_i^{\varepsilon, \gamma})_{i \geqslant 1}$ the family of vector fields consisting of the vector fields $Y_1^{\varepsilon, \gamma}, Y_2^{\varepsilon, \gamma}, \dots, Y_m^{\varepsilon, \gamma}$ completed with all their successive Lie brackets. Then (i) is satisfied as well for every $i \geqslant 1$.*
   In particular, if ε_0 is small enough then $\mathrm{Lie}(Y_1^{\varepsilon, \gamma}, \dots, Y_m^{\varepsilon, \gamma}) = \mathbb{R}^n$ for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, with a uniform degree of nonholonomy r.

(iv) *Generalizing (i) (which we recover for $N = 0$), we have the following asymptotic expansion: given any $N \in \mathbb{N}$, for every $i \in \{0, \dots, m\}$, for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$,*

$$(6.2) \qquad Y_i^{\varepsilon, \gamma} = \widehat{X}_i^q + \varepsilon \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(0)} + \cdots + \varepsilon^N \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(N-1)} + \varepsilon^{N(1-\gamma)+1-\gamma r(q)} R_{i, N}^{\varepsilon, \gamma}$$

where $R_{i, N}^{\varepsilon, \gamma}$ is a smooth vector field on $\mathbb{R}^n$, depending smoothly on ε, which satisfies

$$\left\| R_{i, N}^{\varepsilon, \gamma} f \right\|_{W^{k, p}(\mathbb{R}^n)} \leqslant \mathrm{Cst}(k, N) \| f \|_{W^{k+1, p}(\mathbb{R}^n)}$$

$$\forall\, k \in \mathbb{N} \qquad \forall\, p \in [1, +\infty] \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0] \qquad \forall\, f \in C_c^\infty(\mathbb{R}^n).$$

Note that the asymptotic expansion (6.2) is not at order N but is at the order of the floor (integer part) of $N(1 - \gamma) + 1 - \gamma r(q)$. Note also that the asymptotic expansion (6.2) of $Y_i^{\varepsilon, \gamma}$ is *global*, in contrast to the asymptotic expansion (4.3) of Y_i^ε which is in C^∞ topology.

We will mainly use this lemma with $p = 2$ (note that $W^{k,2}(\mathbb{R}^n) = H^k(\mathbb{R}^n)$) or with $p = +\infty$.

Proof. — Using (4.2), we have $Y_i^{\varepsilon} = \varepsilon^{\gamma} \delta_{\varepsilon^{\gamma}}^* Y_i^{\varepsilon^{1-\gamma}}$, and besides, since $\widehat{X}_i^q$ is homogeneous of degree -1, we have $\widehat{X}_i^q = \varepsilon^{\gamma} \delta_{\varepsilon^{\gamma}}^* \widehat{X}_i^q$, therefore,

$$(6.3) \quad Y_i^{\varepsilon, \gamma} = \widehat{X}_i^q + \varepsilon^{\gamma} \delta_{\varepsilon^{\gamma}}^* \left(\chi \left(Y_i^{\varepsilon^{1-\gamma}} - \widehat{X}_i^q \right) \right), \qquad i = 0, \dots, m, \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

Now, using (2.4), we have $Y_i^{\eta} = \widehat{X}_i^q + \eta Z_i^{\eta}$ for every $\eta \in \mathbb{R}$ with Z_i^{η} depending smoothly on η in C^{∞} topology, and we have $\delta_{\lambda}^* Z_i^{\eta} = Z_i^{\lambda \eta}$ for every $\lambda > 0$. Hence

$$Y_i^{\varepsilon, \gamma} = \widehat{X}_i^q + \varepsilon \delta_{\varepsilon^{\gamma}}^* \left(\chi Z_i^{\varepsilon^{1-\gamma}} \right), \qquad i = 0, \dots, m, \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

Recalling that Y_i^{η} has an asymptotic expansion in η around 0, at any order, in C^{∞} topology, $Y_i^{\eta} = \widehat{X}_i^q + \eta Y_i^{(0)} + o(\eta)$ (see (4.3)), we have that Z_i^{η} converges uniformly to $Y_i^{(0)}$ as well as all its derivatives, on any compact, as $\eta \to 0$. Therefore, multiplying by χ that is of compact support, $\chi Z_i^{\varepsilon^{1-\gamma}}$ converges to $\chi Y_i^{(0)}$ uniformly on $\mathbb{R}^n$, as well as all its derivatives, as $\varepsilon \to 0$. In particular, for every $\alpha \in \mathbb{N}^n$, we have

$$(6.4) \qquad \left\| \partial_x^{\alpha} \left(\chi Z_i^{\varepsilon^{1-\gamma}} \right) \right\|_{L^p(\mathbb{R}^n)} \leqslant \mathrm{Cst}(\alpha) \qquad \forall\, p \in [1, +\infty] \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

We note that $\chi(\delta_{\eta}(x))\, d\delta_{1/\eta}(\delta_{\eta}(x))$ is uniformly bounded on $\mathbb{R}^n$ by $\mathrm{Cst}\,/\eta^{r(q)}$, as well as all its derivatives. Now, since

$$\delta_{\varepsilon^{\gamma}}^* \left(\chi Z_i^{\varepsilon^{1-\gamma}} \right)(x) = d\delta_{1/\varepsilon^{\gamma}}(\delta_{\varepsilon^{\gamma}}(x)). \left(\chi(\delta_{\varepsilon^{\gamma}}(x))\, Z_i^{\varepsilon^{1-\gamma}}(\delta_{\varepsilon^{\gamma}}(x)) \right),$$

using (6.4), it follows that the pullback $\delta_{\varepsilon^{\gamma}}^*(\chi Z_i^{\varepsilon^{1-\gamma}})$ is uniformly bounded on $\mathbb{R}^n$ by $\mathrm{Cst}\,/|\varepsilon|^{\gamma r(q)}$, as well as all its derivatives (indeed, differentiation with respect to x can only multiply terms by ε^{γ}), and then, given any $f \in C_c^{\infty}(\mathbb{R}^n), (Y_i^{\varepsilon, \gamma} - \widehat{X}_i^q) f$ is uniformly bounded on $\mathbb{R}^n$ by $\mathrm{Cst}\,|\varepsilon|^{1-\gamma r(q)}$, as well as all its derivatives, in any L^p space and independently of ε. The proof is similar for the measure $\nu^{\varepsilon, \gamma}$. The statements (i) and (ii) follow immediately.

In order to prove (iii), let us for instance prove that $[Y_i^{\varepsilon, \gamma}, Y_j^{\varepsilon, \gamma}]$ converges to $[\widehat{X}_i^q, \widehat{X}_j^q]$. Noting that, by (i), we have $\|Y_i^{\varepsilon, \gamma}\|_{L(W^{1,p}(\mathbb{R}^n), L^p(\mathbb{R}^n))} \leqslant \mathrm{Cst}$ for every $p \in [1, +\infty]$, given any $f \in C_c^{\infty}(\mathbb{R}^n)$, we have

$$\left\| Y_i^{\varepsilon, \gamma} Y_j^{\varepsilon, \gamma} f - \widehat{X}_i^q \widehat{X}_j^q f \right\|_{L^p(\mathbb{R}^n)}$$
$$\leqslant \left\| Y_i^{\varepsilon, \gamma} \left(Y_j^{\varepsilon, \gamma} - \widehat{X}_j^q \right) f \right\|_{L^p(\mathbb{R}^n)} + \left\| \left(Y_i^{\varepsilon, \gamma} - \widehat{X}_i^q \right) \widehat{X}_j^q f \right\|_{L^p(\mathbb{R}^n)}$$
$$\leqslant \mathrm{Cst} \left(\left\| \left(Y_j^{\varepsilon, \gamma} - \widehat{X}_j^q \right) f \right\|_{W^{1,p}(\mathbb{R}^n)} + |\varepsilon|^{1-\gamma r(q)} \|\widehat{X}_j^q f\|_{W^{1,p}(\mathbb{R}^n)} \right)$$
$$\leqslant \mathrm{Cst}\, |\varepsilon|^{1-\gamma r(q)} \|f\|_{W^{2,p}(\mathbb{R}^n)}$$

and the result follows.

Let us finally prove (iv). We generalize the above argument, starting from the expansion at the order N (see (2.3) and (4.3))

$$Y_i^{\eta} = \widehat{X}_i^q + \eta Y_i^{(0)} + \cdots + \eta^N Y_i^{(N-1)} + \eta^{N+1} Z_{i,N}^{\eta}, \qquad i = 0, \dots, m,$$

for every $\eta \in \mathbb{R}$, where $Z_{i,N}^{\eta}$ is a smooth vector field on $\mathbb{R}^n$ depending smoothing on η in C^∞ topology. Hence, using (6.3) and the fact that $Y_i^{(k)}$ is homogeneous of order k (with respect to dilations), we obtain

$$Y_i^{\varepsilon,\gamma} = \widehat{X}_i^q + \varepsilon \left(\delta_{\varepsilon\gamma}^* \chi\right) Y_i^{(0)} + \cdots + \varepsilon^N \left(\delta_{\varepsilon\gamma}^* \chi\right) Y_i^{(N-1)} + \varepsilon^{N(1-\gamma)+1} \delta_{\varepsilon\gamma}^* \left(\chi Z_{i,N}^{\varepsilon^{1-\gamma}}\right)$$

and then, reasoning as above, all coefficients of the vector field $\delta_{\varepsilon\gamma}^*(\chi Z_{i,N}^{\varepsilon^{1-\gamma}})$ are smooth functions of compact support that are uniformly bounded on $\mathbb{R}^n$ by Cst / $|\varepsilon|^{\gamma r(q)}$, as well as all their derivatives. Setting $R_{i,N}^{\varepsilon,\gamma} = \varepsilon^{\gamma r(q)} \delta_{\varepsilon\gamma}^*(\chi Z_{i,N}^{\varepsilon^{1-\gamma}})$, the statement follows. $\qquad\square$

Thanks to the damping parameter γ, chosen such that $1 - \gamma r(q) > 0$, the modified vector fields $Y_i^{\varepsilon,\gamma}$ converge uniformly on $\mathbb{R}^n$ to $\widehat{X}_i^q$, as well as all their derivatives (with rate of convergence $|\varepsilon|^{1-\gamma r(q)}$), while the vector fields Y_i^{ε} *do not* converge uniformly on $\mathbb{R}^n$ to $\widehat{X}_i^q$ in general, although their convergence is true on any compact. Convergence is also established in Sobolev spaces.

To shed light on the above proof, let us take an example, which moreover shows that the estimates derived in Lemma 6.2 (convergence rate $|\varepsilon|^{1-\gamma r(q)}$) are sharp.

Example 6.3. — In $\mathbb{R}^2$, consider the two vector fields

$$X_1(x_1, x_2) = \frac{\partial}{\partial x_1}, \quad X_2(x_1, x_2) = f(x_1, x_2)\frac{\partial}{\partial x_2}$$

where $f \in C^\infty(\mathbb{R}^2)$ is such that

$$f(0,0) = \frac{\partial f}{\partial x_1}(0,0) = \cdots = \frac{\partial^{k-1} f}{\partial x_1^{k-1}}(0,0) = 0 \quad \text{and} \quad \frac{\partial^k f}{\partial x_1^k}(0,0) = 1$$

for some $k \in \mathbb{N}^*$, i.e., $f(x_1, x_2) = x_1^k(1 + \mathrm{O}(x_1, x_2)) + \mathrm{O}(x_2)$ around $q = (0,0)$. The coordinates (x_1, x_2) are privileged around $(0,0)$ and we have $w_1(0,0) = 1$ and $r(0,0) = w_2(0,0) = k + 1$.

For instance, for $k = 1$ one can take $f(x_1, x_2) = x_1$ or $f(x_1, x_2) = x_1 + x_2^2$ (this case is known in the literature as the *Grushin case* and the corresponding Grushin sR structure is singular at $(0,0)$); for $k = 2$ one can take $f(x_1, x_2) = x_1^2$ or $f(x_1, x_2) = x_1^2 + x_2^2$ or $f(x_1, x_2) = x_1^2 - x_2$ (the latter is called *singular Grushin case*).

For every $\varepsilon \in \mathbb{R} \setminus \{0\}$, we have

$$X_1^{\varepsilon}(x_1, x_2) = \frac{\partial}{\partial x_1}, \qquad X_2^{\varepsilon}(x_1, x_2) = \frac{f\left(\varepsilon x_1, \varepsilon^{k+1} x_2\right)}{\varepsilon^k} \frac{\partial}{\partial x_2}$$

and thus

$$\widehat{X}_1^{(0,0)}(x_1, x_2) = \frac{\partial}{\partial x_1}, \qquad \widehat{X}_2^{(0,0)}(x_1, x_2) = x_1^k \frac{\partial}{\partial x_2}.$$

The family of functions $(f^{\varepsilon})_{\varepsilon \in \mathbb{R}\setminus\{0\}}$, defined by $f^{\varepsilon}(x_1, x_2) = \frac{f(\varepsilon x_1, \varepsilon^{k+1} x_2)}{\varepsilon^k}$, converges to the function $\widehat{f}$ defined by $\widehat{f}(x_1, x_2) = x_1^k$, uniformly on any compact subset of $\mathbb{R}^2$, but does not converge uniformly on $\mathbb{R}^2$ to $\widehat{f}$ in general (take for instance $f(x_1, x_2) = x_1 + x_1^2$, or $f(x_1) = e^{x_1} - 1$).

Now, let us fix an arbitrary $\gamma \in (0, \frac{1}{k+1})$ (recall that $r(0,0) = k+1$) and an arbitrary function $\chi \in C_c^\infty(\mathbb{R})$ such that $\chi(x_1, x_2) = 1$ on the sR ball $\mathrm{B_{sR}}(0,1)$ and $\chi(x_1, x_2) = 0$ on $\mathbb{R}^2 \setminus \mathrm{B_{sR}}(0,2)$. Writing, in short, $\chi \simeq \mathbf{1}_{\mathrm{B_{sR}}(0,1)}$, we have $\delta_{\varepsilon^\gamma}^* \chi \simeq \mathbf{1}_{\mathrm{B_{sR}}(0,1/\varepsilon^\gamma)}$, and we compute

$$X_1^{\varepsilon,\gamma}(x_1, x_2) = \frac{\partial}{\partial x_1},$$

$$X_2^{\varepsilon,\gamma}(x_1, x_2) = x_1 \frac{\partial}{\partial x_2} + \chi\left(\varepsilon^\gamma x_1, \varepsilon^{(k+1)\gamma} x_2\right) \left(\frac{f(\varepsilon x_1, \varepsilon^{k+1} x_2)}{\varepsilon^k} - x_1\right) \frac{\partial}{\partial x_2}.$$

Thanks to the introduction of the parameter γ (small enough), now, the family of functions $(f^{\varepsilon,\gamma})_{\varepsilon \in \mathbb{R}\setminus\{0\}}$, defined by $f^{\varepsilon,\gamma}(x_1, x_2) = \frac{f(\varepsilon x_1, \varepsilon^{k+1} x_2)}{\varepsilon^k}$, converges uniformly on $\mathbb{R}^2$ to the function $\widehat{f}$ defined by $\widehat{f}(x_1, x_2) = x_1$. Indeed, thanks to the truncation function, we have to prove this convergence for $|x_1| \leqslant 1/|\varepsilon|^\gamma$ and $|x_2| \leqslant 1/|\varepsilon|^{(k+1)\gamma}$, i.e., $|\varepsilon x_1| \leqslant |\varepsilon|^{1-\gamma}$ and $|\varepsilon^{k+1} x_2| \leqslant |\varepsilon|^{(k+1)(1-\gamma)}$. But then, for such values of (x_1, x_2), we have

$$f\left(\varepsilon x_1, \varepsilon^{k+1} x_2\right) = \varepsilon^k x_1^k \left(1 + \mathrm{O}\left(|\varepsilon|^{1-\gamma}\right)\right) + \mathrm{O}\left(|\varepsilon|^{(k+1)(1-\gamma)}\right)$$

and thus

$$f^{\varepsilon,\gamma}(x_1, x_2) = x_1^k \left(1 + \mathrm{O}\left(|\varepsilon|^{1-\gamma}\right)\right) + \mathrm{O}\left(|\varepsilon|^{1-(k+1)\gamma}\right) = x_1^k + \mathrm{O}\left(|\varepsilon|^{1-(k+1)\gamma}\right)$$

where the remainder term $\mathrm{O}(|\varepsilon|^{1-(k+1)\gamma})$ is uniform with respect to $(x_1, x_2) \in \mathbb{R}^2$. Hence $\|f^{\varepsilon,\gamma} - \widehat{f}\|_{L^\infty(\mathbb{R}^2)} \leqslant \mathrm{Cst}\, |\varepsilon|^{1-(k+1)\gamma}$, as stated in the general lemma above.

Thanks to the introduction of the damping parameter γ, we have seen that $Y_i^{\varepsilon,\gamma}$ converges to $\widehat{X}_i^q$ in a much stronger way than Y_i^ε. We next prove that these modified vector fields $Y_i^{\varepsilon,\gamma}$ satisfy a global strong Hörmander condition on $\mathbb{R}^n$.

6.1.2. Uniform polynomial strong Hörmander condition

As in Appendix B.2, we denote by $(Y_i^{\varepsilon,\gamma})_{i \geqslant 1}$ the family of vector fields consisting of the vector fields $Y_1^{\varepsilon,\gamma}, Y_2^{\varepsilon,\gamma}, \ldots, Y_m^{\varepsilon,\gamma}$ completed with all their successive Lie brackets. In the following lemma, we prove that the family $(Y_i^{\varepsilon,\gamma})_{i \geqslant 1}$ satisfies the uniform polynomial strong Hörmander condition (B.9) (see Appendix B.2.2), whenever $\gamma > 0$ is small enough.

LEMMA 6.4. — *There exist $N_1 \in \mathbb{N}$ such that, if $\gamma < \frac{1}{r(q)(r(q)+1)}$ then*

$$(6.5) \qquad \|y\|_2^2 \leqslant \mathrm{Cst}\, \langle x \rangle^{2r(q)} \sum_{i=1}^{N_1} \langle Y_i^{\varepsilon,\gamma}(x), y \rangle^2 \qquad \forall\, x, y \in \mathbb{R}^n \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

This lemma will allow us to use the global subelliptic estimates established in Appendix B.2, required in order to obtain the asymptotic expansion of the semigroup with respect to ε (see further).

Proof. — Recall that $Y_i^{0,\gamma} = \widehat{X}_i^q$ for $\varepsilon = 0$. Let us first prove that the family of vector fields $(\widehat{X}_i^q)_{i \geqslant 1}$ (consisting of the vector fields $\widehat{X}_1^q, \ldots, \widehat{X}_m^q$, completed with their iterated Lie brackets), satisfies the polynomial strong Hörmander condition,

i.e., that (6.5) is satisfied for $\varepsilon = 0$. Actually, one could apply Remark B.8 (i.e., one could use [Hör58, Lemma 2]), which would give the inequality (6.5) with the term $\langle x \rangle^{2N_0}$ at the right-hand side, for some $N_0 \in \mathbb{N}$. This would be enough for our needs. But actually, hereafter, instead of using the non-obvious result [Hör58, Lemma 2], we make a direct proof by using the homogeneity property of the vector fields $\widehat{X}_i^q$. In turn, we obtain the optimal integer $N_0 = r(q)$ (it is easy to see on examples that it is optimal).

Let $N_1 \in \mathbb{N}$ be an integer, large enough so that the finite family $(\widehat{X}_i^q)_{0 \leqslant i \leqslant N_1}$ is a frame of $\mathbb{R}^n$ at every point: such an integer exists because the m-tuple of polynomial vector fields $(\widehat{X}_1^q, \ldots, \widehat{X}_m^q)$ satisfies the Hörmander condition at every point of $\mathbb{R}^n$, with a uniform degree of nonholonomy. We have to prove that

$$(6.6) \qquad \|y\|_2^2 \leqslant \mathrm{Cst} \langle x \rangle^{2r(q)} \sum_{i=1}^{N_1} \left\langle \widehat{X}_i^q(x), y \right\rangle^2 \qquad \forall\, x, y \in \mathbb{R}^n.$$

Let us introduce the sR pseudo-norm, defined by

$$\|x\|_{\mathrm{sR}} = \sum_{i=1}^{n} |x_i|^{1/w_i(q)}$$

for every $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ in privileged coordinates. Note that $\|\delta_\varepsilon(x)\|_{\mathrm{sR}} = |\varepsilon| \|x\|_{\mathrm{sR}}$ for every $x \in \mathbb{R}^n$ and every $\varepsilon \in \mathbb{R}$. By the Hörmander condition and by compactness, we first note that

$$(6.7) \qquad \|y\|_2^2 \leqslant \mathrm{Cst} \sum_{i=1}^{N_1} \left\langle \widehat{X}_i^q(x), y \right\rangle^2 \qquad \forall\, y \in \mathbb{R}^n \qquad \forall\, x \in \mathbb{R}^n \ \mid\ \|x\|_{\mathrm{sR}} \leqslant 1$$

and thus (6.6) is satisfied as well for x such that $\|x\|_{\mathrm{sR}} \leqslant 1$. Let us now establish (6.6) for every $x \in \mathbb{R}^n$. Let $x \in \mathbb{R}^n \setminus \{0\}$ be arbitrary. Setting $\varepsilon = \frac{1}{\|x\|_{\mathrm{sR}}}$, we have $\|\delta_\varepsilon(x)\|_{\mathrm{sR}} = 1$ and hence, using (6.7), we get

$$\|y\|_2^2 \leqslant \mathrm{Cst} \sum_{i=1}^{N_1} \left\langle \widehat{X}_i^q(\delta_\varepsilon(x)), y \right\rangle^2 \leqslant \frac{\mathrm{Cst}}{\varepsilon^{2r(q)}} \sum_{i=1}^{N_1} \left\langle \widehat{X}_i^q(x), y \right\rangle^2$$

$$= \mathrm{Cst}\, \|x\|_{\mathrm{sR}}^{2r(q)} \sum_{i=1}^{N_1} \left\langle \widehat{X}_i^q(x), y \right\rangle^2$$

for all $x, y \in \mathbb{R}^n$, where we have used that $\varepsilon \widehat{X}_i^q(\delta_\varepsilon(x)) = \widehat{X}_i^q(x)$ for $i \in \{1, \ldots, m\}$ and thus $\varepsilon^{k_i} \widehat{X}_i^q(\delta_\varepsilon(x)) = \widehat{X}_i^q(x)$ for $i \in \{1, \ldots, N_1\}$ with $k_i \leqslant w_n(q) = r(q)$. To obtain (6.6), we use the fact that $\|x\|_{\mathrm{sR}} \leqslant \mathrm{Cst} \langle x \rangle$ for every $x \in \mathbb{R}^n$.

Now, let us establish (6.5) for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, whenever γ is small enough. Since χ has a compact support, there exists $R > 0$ such that $\mathrm{supp}(\chi) \subset \widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, R)$. Here, given any $r > 0$, the set $\widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, r) = \{x \in \mathbb{R}^n \ \mid\ \widehat{\mathrm{d}}_{\mathrm{sR}}^q(0, x) < r\}$ is the sR ball of center 0 and of radius r, for the sR distance $\widehat{\mathrm{d}}_{\mathrm{sR}}^q$ on $\widehat{M}^q \simeq \mathbb{R}^n$. By definition, given any $f \in C^\infty(\mathbb{R}^n)$ we have

$$\mathrm{supp}\left(\left(Y_i^{\varepsilon, \gamma} - \widehat{X}_i^q\right) f\right) \subset \mathrm{supp}\left(\delta_{\varepsilon^\gamma}^* \chi\right) \subset \widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, R/|\varepsilon^\gamma|),$$

i.e., $Y_i^{\varepsilon,\gamma}$ coincides with $\widehat{X}_i^q$ outside of the sR ball $\widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, R/|\varepsilon^\gamma|)$. Since (6.5) is satisfied for the family $(\widehat{X}_i^q)_{1 \leqslant i \leqslant N_1}$, we only have to prove that (6.5) is satisfied for the family $(Y_i^{\varepsilon,\gamma})_{1 \leqslant i \leqslant N_1}$ for $x \in \widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, R/|\varepsilon^\gamma|)$, i.e., defining the n-by-N_1 matrix

$$P^{\varepsilon,\gamma}(x) = \left(Y_1^{\varepsilon,\gamma}(x) \, Y_2^{\varepsilon,\gamma}(x) \, \cdots \, Y_{N_1}^{\varepsilon,\gamma}(x)\right) \qquad \forall \, x \in \mathbb{R}^n,$$

we have to prove that, if γ is small enough then

$$(6.8) \qquad y^\top P^{\varepsilon,\gamma}(x) P^{\varepsilon,\gamma}(x)^\top y \geqslant \mathrm{Cst} \langle x \rangle^{-2r(q)} \|y\|_2^2 \qquad \forall \, x, y \in \mathbb{R}^n.$$

It follows from Lemma 6.2 that

$$P^{\varepsilon,\gamma}(x) P^{\varepsilon,\gamma}(x)^\top = \widehat{P}^q(x) \widehat{P}^q(x)^\top y + \mathrm{O}\left(|\varepsilon|^{1-\gamma r(q)}\right)$$

when $0 < \gamma < \frac{1}{r(q)}$. Therefore, using (6.6), to prove (6.8) it suffices to observe that

$$\langle x \rangle^{2r(q)} |\varepsilon|^{1-\gamma r(q)} \leqslant |\varepsilon|^{1-\gamma r(q)(r(q)+1)} \qquad \forall \, x \in \widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, R/|\varepsilon^\gamma|),$$

which is so because, for $x \in \widehat{\mathrm{B}}_{\mathrm{sR}}^q(0, R/|\varepsilon^\gamma|)$, we have $|x| \leqslant \mathrm{Cst} \, |\varepsilon|^{-\gamma r(q)}$ (actually, in privileged coordinates, we have $|x_i| \leqslant \mathrm{Cst} \, |\varepsilon|^{-\gamma w_i(q)}$). The conclusion follows, by taking $\gamma < \frac{1}{r(q)(r(q)+1)}$. $\qquad \square$

6.1.3. Definition of the operator $\triangle^{\varepsilon,\gamma}$

Differential operator $\triangle^{\varepsilon,\gamma}$. For every $\gamma \in (0,1)$, for every $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, we define on $C^\infty(\mathbb{R}^n)$ the differential operator

$$(6.9) \qquad \triangle^{\varepsilon,\gamma} = \sum_{i=1}^m \left(Y_i^{\varepsilon,\gamma}\right)^2 + \varepsilon \left(\delta_{\varepsilon^\gamma}^* \chi\right) Y_0^\varepsilon - \varepsilon^2 \left(\delta_{\varepsilon^\gamma}^* \chi\right) \left(\delta_\varepsilon^* \mathrm{v}\right)$$

(note that $\triangle^{0,\gamma} = \widehat{\triangle}^q$). By construction, we have $\triangle^{\varepsilon,\gamma} = \triangle^\varepsilon$ on $\delta_{1/\varepsilon^\gamma}(W_1)$ and $\triangle^{\varepsilon,\gamma} = \widehat{\triangle}^q$ on $\mathbb{R}^n \setminus \delta_{1/\varepsilon^\gamma}(W_2)$. We have the homogeneity property (obvious to check, using (6.1)):

$$(6.10) \quad \varepsilon^{2\beta} \delta_{\varepsilon^\beta}^* \triangle^{\varepsilon,\gamma} \left(\delta_{\varepsilon^\beta}\right)_* = \triangle^{\varepsilon^{1+\beta}, \gamma+\beta}$$

$$\forall \, \beta \in (\gamma - 1, \gamma) \qquad \forall \, \varepsilon \in \left[-\varepsilon_0^{1/(1+\beta)}, \varepsilon_0^{1/(1+\beta)}\right] \setminus \{0\}.$$

When X_0 is a smooth section of D^2 over M, we modify the definition of $\triangle^{\varepsilon,\gamma}$ as follows:

$$\triangle^{\varepsilon,\gamma} = \sum_{i=1}^m \left(Y_i^{\varepsilon,\gamma}\right)^2 + Y_0^{\varepsilon,\gamma} - \varepsilon^2 (\delta_{\varepsilon^\gamma}^* \chi) \left(\delta_\varepsilon^* \mathrm{v}\right).$$

Remark 6.5. — As in Remark 6.1, since $\triangle^{\varepsilon,\gamma}$ coincides with $\triangle^\varepsilon$ on the growing neighborhood $\delta_{1/\varepsilon^\gamma}(W_1)$, we have in particular $\triangle^{\varepsilon,\gamma} = \triangle^\varepsilon + \mathrm{O}(|\varepsilon|^\infty)$ as $\varepsilon \to 0$ in C^∞ topology.

The operator $\triangle^{\varepsilon,\gamma}$, defined by (6.9), depends smoothly on $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ in C^∞ topology. In particular, $\triangle^{\varepsilon,\gamma}$ converges to $\widehat{\triangle}^q$ in C^∞ topology as $\varepsilon \to 0$. But, as a consequence of Lemma 6.2, we have the following stronger result.

LEMMA 6.6. — *If $0 < \gamma < 1/r(q)$ then*

$$\left\| \left(\triangle^{\varepsilon,\gamma} - \widehat{\triangle}^q \right) f \right\|_{L^p(\mathbb{R}^n)} \leqslant \operatorname{Cst} |\varepsilon|^{1-\gamma r(q)} \|f\|_{W^{2,p}(\mathbb{R}^n)}$$

$$\forall\, p \in [1, +\infty] \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0] \qquad \forall\, f \in C_c^\infty(\mathbb{R}^n).$$

Asymptotic expansion of $\triangle^{\varepsilon,\gamma}$. Let us first give an asymptotic expansion in C^∞ topology. Using (4.3), (4.6) and Remark 6.5, we get that $\triangle^{\varepsilon,\gamma}$ has an asymptotic expansion in C^∞ topology at any order N with respect to ε,

$$(6.11) \qquad \triangle^{\varepsilon,\gamma} = \widehat{\triangle}^q + \varepsilon \mathcal{A}_1 + \varepsilon^2 \mathcal{A}_2 + \cdots + \varepsilon^N \mathcal{A}_N + \mathrm{o}\left(|\varepsilon|^N\right)$$

where $\mathcal{A}_i$ is a second-order differential operator for every $i \in \mathbb{N}^*$, with

$$(6.12) \qquad \begin{aligned} \mathcal{A}_1 &= \sum_{i=1}^m \left(\widehat{X}_i^q Y_i^{(0)} + Y_i^{(0)} \widehat{X}_i^q \right) + \widehat{X}_0^q \\ \mathcal{A}_2 &= \sum_{i=1}^m \left(\widehat{X}_i^q Y_i^{(1)} + Y_i^{(1)} \widehat{X}_i^q + \left(Y_i^{(0)} \right)^2 \right) + Y_0^{(0)} - \mathtt{v}(0) \end{aligned}$$

etc. Moreover, $\mathcal{A}_i$ has polynomial coefficients of degree less than $(r(q) + i)^2$ (this bound is not optimal). Indeed, given any integer $k \geqslant -1$, the vector field $Y_i^{(k)}$ is a polynomial that is homogeneous of degree k with respect to dilations. In privileged coordinates, its coefficient along $\frac{\partial}{\partial x_n}$ (where x_n has the largest weight, $w_n(q) = r(q)$) must therefore be a polynomial of degree $\leqslant w_n(q) + k = r(q) + k$.

When X_0 is a smooth section of D^2 over M, since $Y_0^{\varepsilon,\gamma} = \widehat{X}_0^q + (\delta_{\varepsilon\gamma}^* \chi)(Y_0^\varepsilon - \widehat{X}_0^q) = \widehat{X}_i^q + (\delta_{\varepsilon\gamma}^* \chi)(\varepsilon Y_0^{(-1)} + \varepsilon^2 Y_0^{(0)} + \cdots + \varepsilon^N Y_0^{(N-2)} + \mathrm{o}(|\varepsilon|^N))$ with $\widehat{X}_0^q$ homogeneous of order -2 (see Remark 1.2), the expansion (6.11) of $\triangle^{\varepsilon,\gamma}$ remains true as well by replacing $\widehat{\triangle}^q$ with $\widehat{\triangle}^q + \widehat{X}_0^q$ and $\mathcal{A}_1$ with $\mathcal{A}_1 = \sum_{i=1}^m \left(\widehat{X}_i^q Y_i^{(0)} + Y_i^{(0)} \widehat{X}_i^q \right) + Y_0^{(-1)}$, etc.

The asymptotic expansion (6.11) is in C^∞ topology. Thanks to (6.2) in Lemma 6.2, we now derive an asymptotic expansion of $\triangle^{\varepsilon,\gamma}$, at any order, valid on the whole $\mathbb{R}^n$ (not only on every compact): in the above definition of $\mathcal{A}_i$, it suffices to replace $Y_i^{(k)}$ by $(\delta_{\varepsilon\gamma}^* \chi) Y_i^{(k)}$, $\widehat{X}_0^q$ by $(\delta_{\varepsilon\gamma}^* \chi) \widehat{X}_0^q$, etc, and we obtain

$$(6.13) \qquad \triangle^{\varepsilon,\gamma} = \widehat{\triangle}^q + \varepsilon \mathcal{A}_1^{\varepsilon,\gamma} + \varepsilon^2 \mathcal{A}_2^{\varepsilon,\gamma} + \cdots + \varepsilon^N \mathcal{A}_N^{\varepsilon,\gamma} + \varepsilon^{N(1-\gamma)+1-\gamma r(q)} \mathcal{R}_{N+1}^{\varepsilon,\gamma}$$

with

$$(6.14) \qquad \begin{aligned} \mathcal{A}_1^{\varepsilon,\gamma} &= \sum_{i=1}^m \left(\widehat{X}_i^q \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(0)} + \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(0)} \widehat{X}_i^q \right) + \left(\delta_{\varepsilon\gamma}^* \chi \right) \widehat{X}_0^q \\ \mathcal{A}_2^{\varepsilon,\gamma} &= \sum_{i=1}^m \left(\widehat{X}_i^q \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(1)} + \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(1)} \widehat{X}_i^q + \left(\left(\delta_{\varepsilon\gamma}^* \chi \right) Y_i^{(0)} \right)^2 \right) \\ &\qquad + \left(\delta_{\varepsilon\gamma}^* \chi \right) Y_0^{(0)} - \left(\delta_{\varepsilon\gamma}^* \chi \right) \mathtt{v}(0) \end{aligned}$$

etc. Because of the damping parameter γ, the growth of the coefficients of the second-order differential operator $\mathcal{A}_i^{\varepsilon,\gamma}$ is at most polynomial of degree $r(q) - 1$ (which is the maximal degree of the coefficients of $\widehat{X}_i^q$), for all $i \in \mathbb{N}^*$ and $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. Note that $\mathcal{A}_i^{\varepsilon,\gamma} = \mathcal{A}_i + \mathrm{O}(|\varepsilon|^\infty)$ as $\varepsilon \to 0$ in C^∞ topology. We infer from (iv) in Lemma 6.2 that, in (6.13), the term $\mathcal{R}_{N+1}^{\varepsilon,\gamma}$ is a finite sum of products of $R_{j,N}^{\varepsilon,\gamma}$ (whose coefficients

are uniformly bounded) with either some $\widehat{X}_i^q$ (whose coefficients are polynomial of degree $\leqslant r(q) - 1$) or some $(\delta_{\varepsilon\gamma}^* \chi) Y_i^{(0)}$ (whose coefficients are uniformly bounded). Therefore $\mathcal{R}_{N+1}^{\varepsilon,\gamma}$ a second-order differential operator, depending smoothly on ε, with coefficients whose growth is at most polynomial of degree $r(q) - 1$, and

$$\|\mathcal{R}_{N+1}^{\varepsilon,\gamma}\|_{L\left(H_\beta^\alpha(\mathbb{R}^n), H_{\beta-r(q)+1}^{\alpha-2}\right)} \leqslant \mathrm{Cst}(\alpha,\beta) \qquad \forall\, \alpha, \beta \in \mathbb{R} \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0]$$

which is exactly the property $(\mathbf{P}_5)$ that we have identified in Section 5.3. This means that the asymptotic expansion (6.13) is in the sense of operators mapping $H_\beta^\alpha(\mathbb{R}^n)$ to $H_{\beta-N\ell}^{\alpha-2}(\mathbb{R}^n)$, uniformly with respect to ε.

Note that we also have $\|\mathcal{R}_{N+1}^{\varepsilon,\gamma} f\|_{W^{k,p}(\mathbb{R}^n)} \leqslant \mathrm{Cst}(k, N) \|f\|_{W^{k+2,p}(\mathbb{R}^n)}$ for all $k \in \mathbb{N}$, $p \in [1, +\infty]$, $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$ and $f \in C_c^\infty(\mathbb{R}^n)$. This means that the asymptotic expansion (6.13) is also in the sense of operators mapping $W^{k+2,p}(\mathbb{R}^n)$ to $W^{k,p}(\mathbb{R}^n)$, uniformly with respect to ε.

Semigroup generated by $\triangle^{\varepsilon,\gamma}$. In the sequel, we assume that $0 < \gamma < 1/r(q)$. By Lemma 6.2, we have $H^s(\mathbb{R}^n, \nu^{\varepsilon,\gamma}) = H^s(\mathbb{R}^n)$, so that, hereafter, we use the Lebesgue measure in all Sobolev spaces that we consider.

Reasoning as in the proof of Lemma 3.1 and using Lemma 6.6 with $p = 2$, we first observe that there exists $C \geqslant 0$, not depending on $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$, such that the operator $\triangle^{\varepsilon,\gamma} - C$ id is closed and dissipative in $L^2(\mathbb{R}^n)$, as well as its adjoint. Setting $D(\triangle^{\varepsilon,\gamma}) = \{f \in L^2(\mathbb{R}^n) \mid \triangle^{\varepsilon,\gamma} f \in L^2(\mathbb{R}^n)\}$, it follows that the operator $(\triangle^{\varepsilon,\gamma}, D(\triangle^{\varepsilon,\gamma}))$ generates a strongly continuous (quasicontraction) semigroup $(e^{t\triangle^{\varepsilon,\gamma}})_{t\geqslant 0}$ on $L^2(\mathbb{R}^n)$, satisfying $\|e^{t\triangle^{\varepsilon,\gamma}}\|_{L(L^2(\mathbb{R}^n))} \leqslant e^{Ct}$ for every $t \geqslant 0$ and every $\varepsilon \in [-\varepsilon_0, \varepsilon_0] \setminus \{0\}$.

Following Section B.2, for every $k \in \mathbb{Z}$, we define the Hilbert space $\mathcal{D}_k^{\varepsilon,\gamma}$ as the completion of $C_c^\infty(\mathbb{R}^n)$ for the norm

$$\|f\|_{\mathcal{D}_k^{\varepsilon,\gamma}} = \left\|(2C \,\mathrm{id} - \triangle^{\varepsilon,\gamma} - (\triangle^{\varepsilon,\gamma})^*)^k f\right\|_{L^2(\mathbb{R}^n)}.$$

For $k = 1$, we have $\mathcal{D}_1^{\varepsilon,\gamma} = D(\triangle^{\varepsilon,\gamma})$.

By applying the semigroup estimate to $(2C \,\mathrm{id} - \triangle^{\varepsilon,\gamma} - (\triangle^{\varepsilon,\gamma})^*)^k f$ and by a classical argument of restriction or extension of semigroups on the Sobolev towers (see [EN00]), we obtain the uniform estimates

$$(6.15) \qquad \left\|e^{t\triangle^{\varepsilon,\gamma}}\right\|_{L\left(\mathcal{D}_k^{\varepsilon,\gamma}\right)} \leqslant \mathrm{Cst}\, e^{Ct} \qquad \forall\, t \geqslant 0 \qquad \forall\, k \in \mathbb{Z} \qquad \forall\, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

Recall that, for $\varepsilon = 0$, we have $\triangle^{0,\gamma} = \widehat{\triangle}^q$. For every $k \in \mathbb{Z}$, the Hilbert space $\widehat{\mathcal{D}}_k^q$ is the completion of $C_c^\infty(\mathbb{R}^n)$ for the (equivalent) norm $\|f\|_{\widehat{\mathcal{D}}_k^q} = \|(\mathrm{id} - \widehat{\triangle}^q)^k f\|_{L^2(\mathbb{R}^n)}$. Actually, for $\varepsilon = 0$ we have the better estimate $\|e^{t\widehat{\triangle}^q}\|_{L(\widehat{\mathcal{D}}_k^q)} \leqslant 1$ for every $k \in \mathbb{Z}$ (contraction semigroup).

Heat kernel $e^{\varepsilon,\gamma}$ of $\triangle^{\varepsilon,\gamma}$. By hypoellipticity (see Corollary B.3), the Schwartz kernel of $e^{t\triangle^{\varepsilon,\gamma}}$ has a continuous density with respect to $\nu^{\varepsilon,\gamma}$, which is the smooth function

$$e^{\varepsilon,\gamma} = e_{\triangle^{\varepsilon,\gamma}, \nu^{\varepsilon,\gamma}} : (0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n \to (0, +\infty).$$

Using (6.10), we have the homogeneity property (which will be useful at the end of the proof):

$$(6.16) \quad e^{\varepsilon,\gamma}(t,x,x') = |\varepsilon|^{\beta\mathcal{Q}(q)} e^{\varepsilon^{1-\beta},\gamma-\beta}\left(\varepsilon^{2\beta}t, \delta_{\varepsilon^\beta}(x), \delta_{\varepsilon^\beta}(x')\right) \qquad \forall\, \beta \in (\gamma, 1-\gamma)$$

$$\forall\, \varepsilon \in \left[-\varepsilon_0^{1/(1-\beta)}, \varepsilon_0^{1/(1-\beta)}\right] \setminus \{0\} \qquad \forall\, (t,x,x') \in (0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n.$$

Convergence of $e^{\varepsilon,\gamma}$ to $\widehat{e}^q$. Applying Theorem 3.5 (in Section 3.2.2), we obtain that $e^{\varepsilon,\gamma}$ converges to $\widehat{e}^q$ in $C^\infty((0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$ topology as $\varepsilon \to 0$ (when X_0 is a smooth section of D^2 over M, one has to replace $\widehat{\triangle}^q$ with $\widehat{\triangle}^q + \widehat{X}_0^q$), but, as announced earlier, we are now going to derive an asymptotic expansion in ε at any order. This is possible thanks to the introduction of the parameter γ.

6.2. Asymptotic expansion in ε of the semigroup

Recall that the operators $\mathcal{C}_i(t)$ have been defined in Section 5.1, as finite sums of operators $\mathcal{I}_i(t)$, defined by the integral (5.4), which is a convolution in which the compositions involve the operators $e^{s\widehat{\triangle}^q}$ and the derivations $\mathcal{A}_j$. The operators $\mathcal{C}_i(t)$ appear in the asymptotic expansion (5.3) of $e^{t\triangle^\varepsilon}$ in C^∞ topology. In the proposition hereafter, we prove that they appear as well in the asymptotic expansion of $e^{t\triangle^{\varepsilon,\gamma}}$ in the sense of uniformly smoothing operators. This key result is the byproduct of the introduction of the damping parameter γ.

PROPOSITION 6.7. — *We assume that*

$$\gamma < \frac{1}{r(q)(r(q)+1)}.$$

Given any $0 < t_0 < t_1 < +\infty$, given any $N \in \mathbb{N}^*$, we have

$$(6.17) \qquad e^{t\triangle^{\varepsilon,\gamma}} = e^{t\widehat{\triangle}^q} + \varepsilon\mathcal{C}_1(t) + \cdots + \varepsilon^N\mathcal{C}_N(t) + \varepsilon^{N+1}\mathcal{Q}_{N+1}^{\varepsilon,\gamma}(t)$$

as $\varepsilon \to 0$, for every $t \in [t_0, t_1]$, where all operators in (6.17) are locally smoothing for $t \in [t_0, t_1]$, i.e.,

$$\sum_{i=1}^N \left\|\chi_1\mathcal{C}_i(t)\chi_2\right\|_{L\left(H^j(\mathbb{R}^n), H^k(\mathbb{R}^n)\right)} + \left\|\chi_1\mathcal{Q}_{N+1}^{\varepsilon,\gamma}(t)\chi_2\right\|_{L\left(H^j(\mathbb{R}^n), H^k(\mathbb{R}^n)\right)}$$

$$\leqslant \mathrm{Cst}\left(\chi_1, \chi_2, j, k, t_0, t_1, N\right) \forall\, t \in [t_0, t_1] \qquad \forall\, \chi_1, \chi_2 \in C_c^\infty(\mathbb{R}^n) \qquad \forall\, j, k \in \mathbb{Z}.$$

Proof. — Following the approach described in Section 5, we start by applying the Duhamel formula, with the operator $\triangle^{\varepsilon,\gamma}$,

$$e^{t\triangle^{\varepsilon,\gamma}} = e^{t\widehat{\triangle}^q} + \int_0^t e^{(t-s)\triangle^{\varepsilon,\gamma}}\left(\triangle^{\varepsilon,\gamma} - \widehat{\triangle}^q\right) e^{s\widehat{\triangle}^q}\,ds,$$

which we iterate N times, thus obtaining the formula (5.1) with $\triangle^\varepsilon$ replaced by $\triangle^{\varepsilon,\gamma}$. Now, using the asymptotic expansion (6.13) of $\triangle^{\varepsilon,\gamma}$, applied at order N_1 such that $N_1(1-\gamma) + 1 - \gamma r(q) \geqslant N+1$, we obtain

$$(6.18) \qquad e^{t\triangle^{\varepsilon,\gamma}} = e^{t\widehat{\triangle}^q} + \varepsilon\mathcal{C}_i^{\varepsilon,\gamma}(t) + \cdots + \varepsilon^N\mathcal{C}_N^{\varepsilon,\gamma}(t) + \varepsilon^{N+1}\mathcal{P}_{N+1}^{\varepsilon,\gamma}(t)$$

where each operator $\mathcal{C}_j^{\varepsilon,\gamma}(t)$ is a finite sum of terms $\mathcal{I}_i^{\varepsilon,\gamma}(t)$ for some $i \in \{1, \ldots, N\}$, where $\mathcal{I}_i^{\varepsilon,\gamma}(t)$ is defined by

$$\tag{6.19} \mathcal{I}_i^{\varepsilon,\gamma}(t) = \int_{\Sigma_i(t)} e^{s_1 \widehat{\triangle}^q} \mathcal{A}_{j_1}^{\varepsilon,\gamma} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{A}_{j_i}^{\varepsilon,\gamma} e^{s_{i+1} \widehat{\triangle}^q} \, ds^{i+1}$$

with $j_1, \ldots, j_N \in \{1, \ldots, N\}$, and where the remainder term $\mathcal{P}_{N+1}^{\varepsilon,\gamma}(t)$ is a finite sum of terms $\varepsilon^k \mathcal{I}_i^{\varepsilon,\gamma}(t)$, $\varepsilon^\alpha \mathcal{J}_i^{\varepsilon,\gamma}(t)$ and $\varepsilon^\alpha \mathcal{K}_i^{\varepsilon,\gamma}(t)$ with $k, i \in \mathbb{N}$, $k \leqslant (N+1)^2$, $1 \leqslant i \leqslant N$, $\alpha \in [0, (N+1)^2]$, and $\mathcal{J}_i^{\varepsilon,\gamma}(t)$ and $\mathcal{K}_i^{\varepsilon,\gamma}(t)$ are defined by

$$\tag{6.20} \mathcal{J}_i^{\varepsilon,\gamma}(t) = \int_{\Sigma_i(t)} e^{s_1 \widehat{\triangle}^q} \mathcal{B}_1^{\varepsilon,\gamma} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{B}_i^{\varepsilon,\gamma} e^{s_{i+1} \widehat{\triangle}^q} \, ds^{i+1}$$

$$\tag{6.21} \mathcal{K}_i^{\varepsilon,\gamma}(t) = \int_{\Sigma_i(t)} e^{s_1 \triangle^{\varepsilon,\gamma}} \mathcal{B}_1^{\varepsilon,\gamma} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{B}_i^{\varepsilon,\gamma} e^{s_{i+1} \widehat{\triangle}^q} \, ds^{i+1}$$

where each $\mathcal{B}_j^{\varepsilon,\gamma}$ is a second-order derivation, either equal to some $\mathcal{A}_i^{\varepsilon,\gamma}$, $i \in \{1, \ldots, N\}$, or to $\mathcal{R}_{N_1+1}^{\varepsilon,\gamma}$, whose coefficients growth at infinity is at most polynomial of degree $r(q)$.

Note that, compared with the operator $\mathcal{I}_i(t)$ defined by (5.4) in Section 5.1, in the definition (6.19) of $\mathcal{I}_i^{\varepsilon,\gamma}(t)$ the derivations $\mathcal{A}_j$ are replaced with $\mathcal{A}_j^{\varepsilon,\gamma}$.

The above expansion is formal. We give it a rigorous meaning in the following proposition, by using, as explained in Section 5, the various global smoothing properties established in Appendix B.2 and in Appendix C.

The arguments that we develop hereafter will show that all operators $\mathcal{C}_i^{\varepsilon,\gamma}(t)$, $\mathcal{C}_i(t)$, for $i \in \mathbb{N}^*$ and the operators $\mathcal{P}_{N+1}^{\varepsilon,\gamma}(t)$ and $\mathcal{Q}_{N+1}^{\varepsilon,\gamma}(t)$ are locally smoothing for $t \in [t_0, t_1]$, and that, actually, $\mathcal{C}_i^{\varepsilon,\gamma}(t) = \mathcal{C}_i(t) + \mathrm{O}(|\varepsilon|^\infty)$ in C^∞ topology, and then (6.17) will follow from (6.18).

We start by noting that:

- by construction, the coefficients of the operator $\triangle^{\varepsilon,\gamma}$, as well as all their derivatives, have a growth at infinity that is at most polynomial, uniformly with respect to $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$;
- the vector fields defining the one-parameter family of Hörmander operators $(\triangle^{\varepsilon,\gamma})_{-\varepsilon_0 \leqslant \varepsilon \leqslant \varepsilon_0}$ satisfy the uniform polynomial strong Hörmander condition established in Lemma 6.4.

Therefore, the global subelliptic estimates established in Appendix B.2 can be applied to $\triangle^{\varepsilon,\gamma}$ (in other words, we have the properties $(\mathbf{P}_5)$ and $(\mathbf{P}_6)$ motivated in Section 5.3).

We are also going to apply to $\widehat{\triangle}^q$ the global smoothing properties established in Appendix C (in other words, we have the properties $(\mathbf{P}_1)$ and $(\mathbf{P}_2)$ motivated in Section 5.2).

Let $\chi_1, \chi_2 \in C_c^\infty(\mathbb{R}^n)$, let $0 < t_0 < t_1$ be arbitrarily fixed and let $j, k \in \mathbb{Z}$ be arbitrary.

LEMMA 6.8. — *All operators $\mathcal{C}_i(t)$ and $\mathcal{C}_i^{\varepsilon,\gamma}(t)$, $i \in \mathbb{N}^*$, are locally smoothing for $t \in [t_0, t_1]$.*

Proof. — It suffices to prove that, for every $i \in \mathbb{N}^*$, the operator $\chi_1 \mathcal{I}_i(t) \chi_2$, defined by (5.4), maps continuously $H^j(\mathbb{R}^n)$ to $H^k(\mathbb{R}^n)$, with a norm that is uniformly bounded with respect to $t \in [t_0, t_1]$ (in other words, we want to prove (5.7)). The proof for the operator $\mathcal{I}_i^{\varepsilon, \gamma}(t)$, defined by (6.19), will be exactly similar, by replacing the derivations $\mathcal{A}_i$ with $\mathcal{A}_i^{\varepsilon, \gamma}$ which are, as well, derivations whose coefficients growth is at most polynomial, uniformly with respect to ε.

Although the argument has been sketched in Section 5.2, we provide hereafter the detailed proof. Inside the integral (5.4) defining $\mathcal{I}_i(t)$, at least one of the real numbers s_p is such that $s_p \geqslant \frac{t_0}{N}$. We write:

$$\chi_1 \mathcal{I}_i(t) \chi_2 = \int_{\Sigma_i(t)} \chi_1 e^{s_1 \widehat{\triangle}^q} \mathcal{A}_{j_1} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{A}_{j_{p-1}} e^{s_p \widehat{\triangle}^q} \mathcal{A}_{j_p} e^{s_{p+1} \widehat{\triangle}^q} \cdots \mathcal{A}_{j_i} e^{s_{i+1} \widehat{\triangle}^q} \chi_2 \, ds^{i+1}.$$

Using (C.5) and the global continuous embeddings (B.6) and (B.7), the localization operator χ_2 maps continuously $H^j(\mathbb{R}^n)$ to $\widehat{\mathcal{D}}^q_{-m, \beta}$ for some $m \in \mathbb{N}$, for every $\beta \in \mathbb{R}$. We note here that β can be chosen arbitrarily large. Then, by Proposition C.3 (and more precisely, by (C.7)), the operator $e^{s_{i+1} \widehat{\triangle}^q}$ maps continuously $\widehat{\mathcal{D}}^q_{-m, \beta}$ to itself for every $s_i \in [0, t_1]$, provided that $\beta \geqslant 1$, with a norm constant not depending on $s_i \in [0, t_1]$. By (B.7), $\widehat{\mathcal{D}}^q_{-m, \beta}$ is continuously embedded into $H^{-2m}_{\beta - 2mN}(\mathbb{R}^n)$. Applying the derivation $\mathcal{A}_{j_i}$ then maps to $H^{-2m-2}_{\beta - 2mN - (r(q)+j_i)^2}(\mathbb{R}^n)$, because $\mathcal{A}_{j_i}$ is a second-order derivation with polynomial coefficients of degree less than $(r(q) + j_i)^2$. This step is repeated $i - p$ times, so that

$$\mathcal{A}_{j_p} e^{s_{p+1} \widehat{\triangle}^q} \cdots \mathcal{A}_{j_i} e^{s_{i+1} \widehat{\triangle}^q} \chi_2$$

maps continuously $H^j(\mathbb{R}^n)$ to $H^\alpha_\beta(\mathbb{R}^n)$, for some $\alpha \leqslant 0$, for every $\beta \geqslant 1$. We stress that, in the latter repeated argument, as well as in the following, it is important to ensure that the weight be greater than 1, in order to be able to apply (C.7). One now applies the operator $e^{s_p \widehat{\triangle}^q}$, which maps continuously $H^\alpha_\beta(\mathbb{R}^n)$ to $H^{\alpha'}_{\beta - k_0(|\alpha| + |\alpha'|)}(\mathbb{R}^n)$, for every $\alpha' \in \mathbb{R}$, by applying the global smoothing property $(\mathbf{P}_1)$ stated in Section 5.2 (itself following from Proposition C.3 in Appendix C.2 and from the embeddings (B.6) and (B.7)). Then we continue by applying again compositions $\mathcal{A}_j e^{s \widehat{\triangle}^q}$ as above, and the final composition by the localization operator χ_1, which maps continuously to $H^k(\mathbb{R}^n)$ if α' and β have been chosen large enough. □

LEMMA 6.9. — *The operator $\mathcal{P}_{N+1}^{\varepsilon, \gamma}(t)$ is locally smoothing for $t \in [t_0, t_1]$.*

Proof. — It suffices to prove that $\chi_1 \mathcal{I}_i^{\varepsilon, \gamma}(t) \chi_2$ and $\chi_1 \mathcal{K}_i^{\varepsilon, \gamma}(t) \chi_2$ (see (6.20) and (6.21)) map continuously $H^j(\mathbb{R}^n)$ to $H^k(\mathbb{R}^n)$, with a norm that is uniformly bounded with respect to $t \in [t_0, t_1]$ and to $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. As pointed out at the beginning of Section 5.3, the argument is similar to the one used in the proof of Lemma 6.8 with the following differences:

- The second-order derivations $\mathcal{B}_j^{\varepsilon, \gamma}$ can be either equal to some $\mathcal{A}_i^{\varepsilon, \gamma}$, $i \in \{1, \dots, N\}$, or to $\mathcal{R}_{N_1+1}^{\varepsilon, \gamma}$. In the latter case, this does not affect our previous reasoning because its coefficients have a growth at infinity that is at most polynomial of degree $r(q)$.
- In the definition (6.21) of $\mathcal{K}_i^{\varepsilon, \gamma}(t)$, the last term in the composition is $e^{s_1 \triangle^{\varepsilon, \gamma}}$.

Hence, the only main difference is in the last item: we have to explain how to deal with the composition by $e^{s_1 \triangle^{\varepsilon, \gamma}}$. There are two cases.

If $s_1 \geqslant \frac{t_0}{N}$, then the term $\chi_1 e^{s_1 \triangle^{\varepsilon, \gamma}}$ is expected to act as a smoothing operator. By the above reasoning, the operator $\mathcal{B}_1^{\varepsilon, \gamma} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{B}_i^{\varepsilon, \gamma} e^{s_{i+1} \widehat{\triangle}^q} \chi_2$ maps continuously $H^j(\mathbb{R}^n)$ to $H_{k_0 \alpha}^{-\alpha}(\mathbb{R}^n)$ for some $\alpha > 0$ and some $k_0 \in \mathbb{N}$. Now, we apply to $\chi_1 e^{s_1 \triangle^{\varepsilon, \gamma}}$ the property $(\mathbf{P}_6)$ stated in Section 5.3 (itself following from Corollary B.9 in Appendix B.2.2), which gives the result.

Otherwise, another term $e^{s_p \widehat{\triangle}^q}$ acts as a smoothing operator, and thus, by the above reasoning, the operator $\mathcal{B}_1^{\varepsilon, \gamma} e^{s_2 \widehat{\triangle}^q} \cdots \mathcal{B}_i^{\varepsilon, \gamma} e^{s_{i+1} \widehat{\triangle}^q} \chi_2$ maps continuously $H^j(\mathbb{R}^n)$ to $H_\beta^m(\mathbb{R}^n)$ for any $m \in \mathbb{N}$ and $\beta \geqslant 1$, and thus, using the global continuous embeddings (B.6), it maps $H^j(\mathbb{R}^n)$ to $\mathcal{D}_m^{\varepsilon, \gamma}$, for any $m \in \mathbb{N}$ arbitrarily large, uniformly with respect to $\varepsilon \in [-\varepsilon_0, \varepsilon_0]$. One then applies $e^{s_1 \triangle^{\varepsilon, \gamma}}$, which, by (6.15), maps continuously $\mathcal{D}_m^{\varepsilon, \gamma}$ to $\mathcal{D}_m^{\varepsilon, \gamma}$ (uniformly with respect to ε), and the final application of the localization operator χ_1 maps to $H^k(\mathbb{R}^n)$, by applying the uniform local subellipticity estimates (Theorem B.1), provided that m be large enough. $\qquad \square$

To conclude, it remains to establish the following lemma.

LEMMA 6.10. — We have
$$\mathcal{C}_i^{\varepsilon, \gamma}(t) = \mathcal{C}_i(t) + \mathrm{O}(|\varepsilon|^\infty)$$
as $\varepsilon \to 0$ in C^∞ topology, for every $i \in \{1, \ldots, N\}$, for every $t \in [t_0, t_1]$.

Proof. — It suffices to prove that $\chi_1(\mathcal{I}_i^{\varepsilon, \gamma}(t) - \mathcal{I}_i(t))\chi_2 = \mathrm{O}(|\varepsilon|^\infty)$, i.e., that
(6.22)
$$\int_{\Sigma_i(t)} \chi_1 e^{s_1 \widehat{\triangle}^q} \left(\mathcal{A}_{j_1}^{\varepsilon, \gamma} - \mathcal{A}_{j_1} \right) e^{s_2 \widehat{\triangle}^q} \cdots \left(\mathcal{A}_{j_i}^{\varepsilon, \gamma} - \mathcal{A}_{j_i} \right) e^{s_{i+1} \widehat{\triangle}^q} \chi_2 \, ds^{i+1} = \mathrm{O}(|\varepsilon|^\infty)$$
as $\varepsilon \to 0$, for every $i \in \mathbb{N}^*$. Recall that the second-order differential operators $\mathcal{A}_i$ and $\mathcal{A}_i^{\varepsilon, \gamma}$ are respectively defined by (6.12) and (6.14). It follows from their definition that there exists $R > 0$ such that
$$\mathrm{supp}\left((\mathcal{A}_i^{\varepsilon, \gamma} - \mathcal{A}_i) f \right) \subset \mathbb{R}^n \setminus \widehat{B}^q(0, R/\varepsilon^\gamma) \qquad \forall \varepsilon \in [-\varepsilon_0, \varepsilon_0] \qquad \forall f \in C^\infty(\mathbb{R}^n).$$
Besides, for $i = 1, 2$, let $R_i > 0$ be such that $\mathrm{supp}(\chi_i) \subset \widehat{B}^q(0, R_i)$. Here, $\widehat{B}^q(0, R)$ denotes the sR ball in $\mathbb{R}^n$ of center 0 and radius R, for the (nilpotent) sR structure associated with $\widehat{\triangle}^q$. We are going to use the exponential estimates (C.4) given in Appendix C.1.

The Sobolev regularity arguments are to those in the proof of Lemma 6.8, but we slightly modify those arguments with the following additional considerations.

Inside the integral (6.22), at least one of the real numbers s_p is such that $s_p \geqslant \frac{t_0}{N}$. We write $e^{s_p \widehat{\triangle}^q} = e^{\frac{s_p}{2} \widehat{\triangle}^q} e^{\frac{s_p}{2} \widehat{\triangle}^q}$. Let χ_3^ε be a function of compact support such that $\chi_3(x) = 1$ on $\widehat{B}^q(0, R/\varepsilon^{\gamma/2})$ and $\chi_3(x) = 0$ on $\mathbb{R}^n \setminus \widehat{B}^q(0, 2R/\varepsilon^{\gamma/2})$. In other words, one has $\chi_3 \simeq \mathbf{1}_{\widehat{B}^q(0, R/\varepsilon^{\gamma/2})}$. We write the operator inside the integral (6.22) as

$$\chi_1 \mathcal{D}_1^{\varepsilon, \gamma} e^{\frac{s_p}{2} \widehat{\triangle}^q} e^{\frac{s_p}{2} \widehat{\triangle}^q} \mathcal{D}_2^{\varepsilon, \gamma} \chi_2$$
$$= \chi_1 \mathcal{D}_1^{\varepsilon, \gamma} e^{\frac{s_p}{2} \widehat{\triangle}^q} \chi_3 e^{\frac{s_p}{2} \widehat{\triangle}^q} \mathcal{D}_2^{\varepsilon, \gamma} \chi_2 + \chi_1 \mathcal{D}_1^{\varepsilon, \gamma} e^{\frac{s_p}{2} \widehat{\triangle}^q} (1 - \chi_3) e^{\frac{s_p}{2} \widehat{\triangle}^q} \mathcal{D}_2^{\varepsilon, \gamma} \chi_2$$

where $\mathcal{D}_1^{\varepsilon,\gamma}$ and $\mathcal{D}_2^{\varepsilon,\gamma}$ are compositions of $e^{s_j\widehat{\triangle}^q}$ and of $\mathcal{A}_j^{\varepsilon,\gamma} - \mathcal{A}_j$. Using the exponential estimates (C.4) for heat kernels and their derivatives (see Appendix C.1), and combining with the previous reasonings, we infer that, not only the operator $e^{\frac{s_p}{2}\widehat{\triangle}^q}\mathcal{D}_2^{\varepsilon,\gamma}\chi_2$ maps $H^j(\mathbb{R}^n)$ to $H^\alpha_\beta(\mathbb{R}^n)$ for any α and β arbitrarily large, but also, its range is "essentially concentrated" in $\widehat{B}^q(0, R/\varepsilon^{\gamma/2})$, in the sense that

$$\left\| (1 - \chi_3) e^{\frac{s_p}{2}\widehat{\triangle}^q} \mathcal{D}_2^{\varepsilon,\gamma} \chi_2 \right\|_{L\left(H^j(\mathbb{R}^n), H^\alpha_\beta(\mathbb{R}^n)\right)} \leqslant C \, e^{-C/|\varepsilon|^\gamma} \qquad \forall \, \varepsilon \in [-\varepsilon_0, \varepsilon_0]$$

with $C = \mathrm{Cst}(j, k, t_0, t_1, N)$. The exponential term comes from the fact that the support of $1 - \chi_3$ is far from the support of χ_2, at a distance of the order of $1/|\varepsilon|^{\gamma/2}$. By a similar argument, we have

$$\left\| \left(\mathcal{A}_{j_p}^{\varepsilon,\gamma} - \mathcal{A}_{j_p} \right) e^{\frac{s_p}{2}\widehat{\triangle}^q} \chi_3 e^{\frac{s_p}{2}\widehat{\triangle}^q} \mathcal{D}_2^{\varepsilon,\gamma} \chi_2 \right\|_{L\left(H^j(\mathbb{R}^n), H^\alpha_\beta(\mathbb{R}^n)\right)} \leqslant C \, e^{-C/|\varepsilon|^\gamma} \quad \forall \, \varepsilon \in [-\varepsilon_0, \varepsilon_0].$$

The exponential term comes from the fact that the support of $\mathcal{A}_i^{\varepsilon,\gamma} - \mathcal{A}_i$ is far from the support of χ_3, at a distance of the order of $1/|\varepsilon|^\gamma$.

All in all, combining with the arguments already used in the proof of Lemma 6.8, we obtain the Lemma 6.10.   □

Proposition 6.7 is proved.   □

6.3. Asymptotic expansion in ε of the heat kernel

Taking the Schwartz kernels in the expansion (6.17), more precisely, considering their densities with respect to the measure $\widehat{\mu}^q$, and recalling that $\widehat{e}^q = e_{\widehat{\triangle}^q, \widehat{\mu}^q}$, we get that

$$(6.23) \qquad e_{\triangle^{\varepsilon,\gamma}, \widehat{\mu}^q}(t, x, x') = \widehat{e}^q(t, x, x') + \sum_{i=1}^{N} \varepsilon^i [\mathcal{C}_i(t)]_{\widehat{\mu}^q}(x, x') + \mathrm{o}\left(|\varepsilon|^N\right)$$

at any order N, as $\varepsilon \to 0$, in $C^\infty((0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$ topology. For instance, we have

$$[\mathcal{C}_1(t)]_{\widehat{\mu}^q}(x, x') = \int_0^t \int_{\mathbb{R}^n} \widehat{e}^q(t - s, x, z) \, ((\mathcal{A}_1)_x \widehat{e}^q) \, (s, z, x') \, dz \, ds$$

and all other (smooth) functions $[\mathcal{C}_i(t)]_{\widehat{\mu}^q}(x, x')$ can be expressed as well with convolutions. Now, using the formula (A.2) in Appendix A, we have

$$e^{\varepsilon,\gamma}(t, x, x') = e_{\triangle^{\varepsilon,\gamma}, \nu^{\varepsilon,\gamma}}(t, x, x') = \frac{1}{h^{\varepsilon,\gamma}(x')} e_{\triangle^{\varepsilon,\gamma}, \widehat{\mu}^q}(t, x, x')$$

where $h^{\varepsilon,\gamma} = \frac{d\nu^{\varepsilon,\gamma}}{d\widehat{\mu}^q}$ is the (smooth) density of $\nu^{\varepsilon,\gamma}$ with respect to $\widehat{\mu}^q$. By Lemma 6.2, $h^{\varepsilon,\gamma}$ converges to 1 uniformly on $\mathbb{R}^n$, and it depends smoothly on ε in C^∞ topology, hence $h^{\varepsilon,\gamma} = 1 + \varepsilon h^1 + \cdots + \varepsilon^N h^N + \mathrm{o}(|\varepsilon|^N)$ at any order N, in C^∞ topology. Using (6.23), we conclude that

$$(6.24) \qquad e^{\varepsilon,\gamma}(t, x, x') = \widehat{e}^q(t, x, x') + \sum_{i=1}^{N} \varepsilon^i f_i^q(t, x, x') + \mathrm{o}\left(|\varepsilon|^N\right)$$

at any order N, as $\varepsilon \to 0$, in $C^\infty((0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$ topology, where the functions f_i^q are smooth on $(0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Recall that $e^\varepsilon = e_{\triangle^\varepsilon, \nu^\varepsilon}$ where $\triangle^\varepsilon$ is defined by (4.4).

LEMMA 6.11. — *We have*
$$e^{\varepsilon, \gamma}(t, x, x') = e^\varepsilon(t, x, x') + \mathrm{O}(|\varepsilon|^\infty)$$
as $\varepsilon \to 0$ in $C^\infty((0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$ topology.

Proof. — Let $\beta > 0$ small. By (6.16), we have
$$e^{\varepsilon, \gamma}(t, x, x') = |\varepsilon|^{\beta \mathcal{Q}(q)} e^{\varepsilon^{1-\beta}, \gamma-\beta}\left(\varepsilon^{2\beta} t, \delta_{\varepsilon^\beta}(x), \delta_{\varepsilon^\beta}(x')\right).$$

Since $\triangle^{\varepsilon^{1-\beta}, \gamma-\beta}$ coincides with $\triangle^{\varepsilon^{1-\beta}}$ on every compact, we infer from the localization result, Theorem 3.2 (Section 3.2.1), that
$$|\varepsilon|^{\beta\mathcal{Q}(q)} e^{\varepsilon^{1-\beta}, \gamma-\beta}\left(\varepsilon^{2\beta} t, \delta_{\varepsilon^\beta}(x), \delta_{\varepsilon^\beta}(x')\right) = |\varepsilon|^{\beta\mathcal{Q}(q)} e^{\varepsilon^{1-\beta}}\left(\varepsilon^{2\beta} t, \delta_{\varepsilon^\beta}(x), \delta_{\varepsilon^\beta}(x')\right) + \mathrm{O}(|\varepsilon|^\infty)$$

as $\varepsilon \to 0$, on every compact subset of $(0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n$. Besides, it follows from the homogeneity property (4.5) that $e^\varepsilon(t, x, x') = \lambda^{\mathcal{Q}(q)} e^{\varepsilon/\lambda}(\lambda^2 t, \delta_\lambda(x), \delta_\lambda(x'))$ for $\lambda > 0$, and taking $\lambda = \varepsilon^{2\beta}$, we obtain
$$|\varepsilon|^{\beta\mathcal{Q}(q)} e^{\varepsilon^{1-\beta}}\left(\varepsilon^{2\beta} t, \delta_{\varepsilon^\beta}(x), \delta_{\varepsilon^\beta}(x')\right) = e^\varepsilon(t, x, x').$$

The Lemma 6.11 follows. $\qquad\square$

We note again, in the above proof, the crucial role of the parameter γ in our construction. Here, it has been instrumental to be able to apply the localization result, and thus show that the small-time asymptotics of $e^{\varepsilon, \gamma}$ and of e^ε coincide at the infinite order.

Using (6.24) and Lemma 6.11, we finally obtain that
$$e^\varepsilon(t, x, x') = \widehat{e}^q(t, x, x') + \sum_{i=1}^N \varepsilon^i f_i^q(t, x, x') + \mathrm{o}\left(|\varepsilon|^N\right)$$

at any order N, as $\varepsilon \to 0$, in $C^\infty((0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$ topology.

6.4. End of the proof

The end of the proof is now similar to the proof of Theorem B. Thanks to repeated applications of the localization result, Theorem 3.2 (hypoelliptic Kac's principle), we finally obtain
$$|\varepsilon|^{\mathcal{Q}(q)} e_{\triangle, \mu}\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) = \widehat{e}^q(t, x, x') + \sum_{i=1}^N \varepsilon^i f_i^q(t, x, x') + \mathrm{o}\left(|\varepsilon|^N\right)$$

at any order N, as $\varepsilon \to 0$, in C^∞ topology, which is exactly (1.3).

Let us next establish the homogeneity property for the (smooth) functions f_i^q.

LEMMA 6.12. — *For every $i \in \mathbb{N}^*$ and every $\varepsilon \neq 0$, we have*
$$f_i^q(t, x, x') = \varepsilon^{-i} |\varepsilon|^{\mathcal{Q}(q)} f_i^q\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right)$$
for all $(t, x, x') \in (0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Proof. — Given any fixed $\varepsilon \neq 0$, by (1.3), we have on the one part

$$|s\varepsilon|^{\mathcal{Q}(q)} e_{\triangle, \mu}\left(s^2\varepsilon^2 t, \delta_{s\varepsilon}(x), \delta_{s\varepsilon}(x')\right) = \widehat{e}^q(t, x, x') + \sum_{i=1}^N (s\varepsilon)^i f_i^q(t, x, x') + \mathrm{o}\left(s^N\right)$$

and on the other part

$$|s\varepsilon|^{\mathcal{Q}(q)} e_{\triangle, \mu}\left(s^2\varepsilon^2 t, \delta_{s\varepsilon}(x), \delta_{s\varepsilon}(x')\right)$$
$$= |\varepsilon|^{\mathcal{Q}(q)} \widehat{e}^q\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) + |\varepsilon|^{\mathcal{Q}(q)} \sum_{i=1}^N s^i f_i^q\left(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x')\right) + \mathrm{o}\left(s^N\right)$$

for every $s \in \mathbb{R}$ with $|s|$ sufficiently small and all $t > 0$ and $x, x' \in \mathbb{R}^n$. The result follows, since $\widehat{e}^q(t, x, x') = |\varepsilon|^{\mathcal{Q}(q)} \widehat{e}^q(\varepsilon^2 t, \delta_\varepsilon(x), \delta_\varepsilon(x'))$ (see (3.3)). $\square$

Remark 6.13. — Applying Lemma 6.12 with $\varepsilon = -1$, we obtain the "oddness" property

$$f_{2j-1}^q(t, x, x')$$
$$= -f_{2j-1}^q\left(t, \delta_{-1}(x), \delta_{-1}(x')\right) \forall\, j \in \mathbb{N}^* \quad \forall\, (t, x, x') \in (0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n.$$

In particular, we have $f_{2j-1}^q(t, 0, 0) = -f_{2j-1}^q(t, 0, 0)$ and thus $f_{2j-1}^q(t, 0, 0) = 0$ for every $j \in \mathbb{N}^*$ and for every $t > 0$. Note that, to obtain this property, it has been necessary to consider dilations δ_ε with $\varepsilon < 0$, whereas dilations are most often considered only with $\varepsilon > 0$ in the existing literature. We realized this fact in a discussion with Davide Barilari, whom we thank once again.

Case q regular. The case is treated exactly as in Section 4.4. We do not repeat the argument.

Part 3. Appendix
Appendix A. Schwartz kernels, heat kernels

Let M be a smooth manifold. We set $\mathcal{D}(M) = C_c^\infty(M)$ and we denote by $\mathcal{D}'(M)$ the space of distributions on M, i.e., the topological dual of $\mathcal{D}(M)$ endowed with the weak topology. Let μ be a smooth measure (density) on M.

Schwartz kernels. According to the Schwartz kernel theorem, there is a linear bijection between $\mathcal{D}'(M \times M)$ and the set of bilinear continuous functionals on $\mathcal{D}(M) \times \mathcal{D}(M)$. Given a linear continuous mapping $A : \mathcal{D}(M) \to \mathcal{D}'(M)$, the Schwartz kernel of A is the unique distribution $[A] \in \mathcal{D}'(M \times M)$ defined by $\langle Af, g \rangle_{\mathcal{D}'(M), \mathcal{D}(M)} = \langle [A], g \otimes f \rangle$ for all $f, g \in C_c^\infty(M)$, where $\langle \cdot, \cdot \rangle$ is the duality bracket.

When $[A] \in C^0(M \times M)$, identifying the distribution bracket by an integral with respect to the measure $\mu \otimes \mu$ and denoting by $[A]_\mu$ the density function, we have the familiar formula

$$Af(q) = \int_M f(q')\, [A]_\mu(q, q')\, d\mu(q') \qquad \forall\, q \in M \qquad \forall\, f \in \mathcal{D}(M).$$

We stress that, although the density function $[A]_\mu$ depends on μ, given any $q \in M$, the absolutely continuous measure $[A]_\mu(q, \cdot) \, d\mu(\cdot)$ depends only on A: it does not depend on the smooth measure μ, in the sense that $[A]_\mu(q, \cdot) \, d\mu(\cdot) = [A]_\nu(q, \cdot) \, d\nu(\cdot)$ for any other smooth measure ν on M.

Actually, in geometric terms, $[A]$ is a continuous section of the bundle $\pi_2^*(\Omega_M)$ on $M \times M$, where Ω_M is the line bundle of smooth measures (densities) on M and $\pi_2 : M \times M \to M$ is the projection defined by $\pi_2(q, q') = q'$.

Similarly, the diagonal part $[A]_\mu(q, q) \, d\mu(q)$ is an absolutely continuous measure, which does not depend on μ. Denoting by $\mathcal{M}_f$ the operator of multiplication by f, we have

$$\mathrm{Tr}(A\mathcal{M}_f) = \int_M f(q) \, [A]_\mu(q, q) \, d\mu(q) \qquad \forall \, f \in \mathcal{D}(M).$$

Hilbert–Schmidt norm. When $[A] \in L^2(M \times M, \mu \otimes \mu)$, the operator $A \in L(L^2(M, \mu))$ is Hilbert–Schmidt and the Hilbert–Schmidt norm of A is

$$\|A\|_{HS} = (\mathrm{Tr}(A^*A))^{1/2} = \|[A]_\mu\|_{L^2(M \times M, \mu \otimes \mu)}$$

and we recall that

$$\|A\|_{L(L^2(M, \mu))} \leqslant \|A\|_{HS}$$
$$\|AB\|_{HS} \leqslant \|A\|_{L(L^2(M, \mu))} \|B\|_{HS}$$
$$\|AB\|_{HS} \leqslant \|A\|_{HS} \|B\|_{L(L^2(M, \mu))}$$

where A and B are bounded operators on $L^2(M, \mu)$, with A or B Hilbert–Schmidt according to the inequality under consideration.

Action of pseudo-differential operators on Schwartz kernels. Given any pseudo-differential operators T_1 and T_2 on M, we have $[T_1 A T_2^*]_\mu = (T_1)_q (T_2)_{q'} [A]_\mu$ (where T_2^* is the transpose in $L^2(M, \mu)$), i.e.,

$$T_1 A T_2^* f(q) = \int_M f(q') \, (T_1)_q (T_2)_{q'} [A]_\mu(q, q') \, d\mu(q').$$

Heat kernels. Let $A : D(A) \to L^2(M, \mu)$ be a densely defined operator on $L^2(M, \mu)$, generating a strongly continuous semigroup $(e^{tA})_{t \geqslant 0}$. For every $t > 0$, the heat kernel $e_A(t)$ associated with A is the measure on M defined as the Schwartz kernel of e^{tA}, i.e., $e_A(t) = [e^{tA}]$. Of course, it does not depend on μ.

When this measure has a density $[e^{tA}]_\mu$ with respect to μ which is locally integrable, we define the heat kernel $e_{A,\mu}(t, \cdot, \cdot)$ associated with A and with the measure μ by $e_{A,\mu}(t, q, q') = [e^{tA}]_\mu(q, q')$. This means that

$$u(t, q) = \left(e^{tA} f\right)(q) = \int_M f(q') \, e_{A,\mu}(t, q, q') \, d\mu(q')$$

is the unique solution to $\partial_t u - Au = 0$ for $t > 0$, $u(0, \cdot) = f(\cdot)$, for every $f \in C_c^\infty(M)$. In other words, we have

$$e_A(t)(q, q') = \left[e^{tA}\right](q, q') = e_{A,\mu}(t, q, q') \, d\mu(q') \qquad \forall \, t > 0 \qquad \forall \, q, q' \in M.$$

As said above, this expression depends only on A, not on the smooth measure μ.

Extending $e_{A,\mu}$ by 0 for $t < 0$, for any fixed $q' \in M$ the mapping $(t,q) \mapsto e_{A,\mu}(t,q,q')$ is also solution of $(\partial_t - A)e_{A,\mu}(\cdot,\cdot,q') = \delta_{(0,q')}$ in the sense of distributions, where the distribution pairing is considered with respect to the measure $dt \times d\mu(q)$ on $\mathbb{R} \times M$.

We gather hereafter some useful facts.

- Let $\varphi : M \to M$ be a diffeomorphism, representing a change of variable in the manifold M. We have $\varphi^*\mu = |J_\mu(\varphi)|\mu$, where $J_\mu(\varphi)$ is the Jacobian of φ with respect to μ, and where $\varphi^*\mu$ is the pullback of μ under φ. Then

$$e_{\varphi^*A\varphi_*,\mu}(t,q,q') = |J_\mu(\varphi)(q')|\, e_{A,\mu}(t,\varphi(q),\varphi(q'))$$

(A.1) $$e_{A,\varphi^*\mu}(t,q,q') = \frac{1}{|J_\mu(\varphi)(q')|}\, e_{A,\mu}(t,q,q')$$

$$e_{\varphi^*A\varphi_*,\varphi^*\mu}(t,q,q') = e_{A,\mu}\left(t,\varphi(q),\varphi(q')\right)$$

for all $t > 0$ and $(q,q') \in M^2$. Note that the last one follows from the two first ones, in which we have replaced A with $\varphi^*A\varphi_*$ in the second one. The two first formulas in (A.1) are not symmetric, but there is no contradiction there: indeed if A is selfadjoint in $L^2(M,\mu)$ then $e_{A,\mu}$ is symmetric, but A need not be selfadjoint in $L^2(M,\varphi^*\mu)$ and thus $e_{A,\varphi^*\mu}$ need not be symmetric.

  Actually, we have

$$e_{\varphi^*A\varphi_*,\mu}(t,q,q')\,d(\varphi^*\mu)(q') = e_{\varphi^*A\varphi_*,\nu}(t,q,q')\,d\nu(q')$$

for any other smooth measure ν on M.
- As a particular case, given any $\lambda > 0$, we have $\lambda\, e_{A,\lambda\mu} = e_{A,\mu}$.
- Given any $\varepsilon > 0$, the kernel associated with $\varepsilon^2 A$ and with the measure ν is

$$e_{\varepsilon^2 A,\nu}(t,q,q') = e_{A,\nu}(\varepsilon^2 t,q,q')$$

for all $t > 0$ and $(q,q') \in M^2$.
- We assume that $\mu = h\nu$ with h a positive smooth function on M (density of μ with respect to ν). Then $h(q')e_{A,\mu}(t,q,q') = e_{A,\nu}(t,q,q')$ for all $(t,q,q') \in (0,+\infty) \times M \times M$, or equivalently,

(A.2) $$e_{A,\mu}(t,q,q')\,d\mu(q') = e_{A,\nu}(t,q,q')\,d\nu(q').$$

Appendix B. Subelliptic estimates and smoothing properties for hypoelliptic heat semigroups

This section can be read independently of the rest.

Let $d \in \mathbb{N}^*$. We set $\Lambda = (\mathrm{id} - \sum_{i=1}^d \partial_i^2)^{1/2}$. Denoting by $\|u\|_{L^2(\mathbb{R}^d)}$ the L^2 norm of a function $u \in L^2(\mathbb{R}^d)$ and by $\langle \cdot, \cdot \rangle_{L^2(\mathbb{R}^d)}$ the corresponding inner product, we recall that the Hilbert space $H^s(\mathbb{R}^d)$ is equipped with the norm $\|u\|_{H^s(\mathbb{R}^d)} = \|\Lambda^s u\|_{L^2(\mathbb{R}^d)}$ and with the inner product $\langle u, v \rangle_{H^s(\mathbb{R}^d)} = \langle u, \Lambda^{2s} v \rangle_{L^2(\mathbb{R}^d)}$.

Let $p \in \mathbb{N}^*$ and let $\mathcal{K}$ be a compact set (in our applications, we will take either $\mathcal{K} = [-\varepsilon_0, \varepsilon_0]$ or $\mathcal{K} = [-\varepsilon_0, \varepsilon_0] \times K$ for some $\varepsilon_0 > 0$ and for some compact subset K of M). For every $\tau \in \mathcal{K}$, let $Y_0^\tau, Y_1^\tau, \ldots, Y_p^\tau$ be smooth vector fields on $\mathbb{R}^d$ and

let $\mathbb{V}^\tau$ be a smooth function on $\mathbb{R}^d$, all of them depending continuously on τ in C^∞ topology. We set

(B.1) $$L^\tau = \sum_{i=1}^{p} (Y_i^\tau)^2 + Y_0^\tau - \mathbb{V}^\tau.$$

B.1. Local estimates

Let U be an arbitrary open subset of $\mathbb{R}^d$. Of course, all local estimates hereafter could be settled as well on a d-dimensional manifold.

B.1.1. Uniform local subelliptic estimates

THEOREM B.1. — *We assume that the Lie algebra* $\mathrm{Lie}(Y_0^\tau, Y_1^\tau, \ldots, Y_p^\tau)$ *generated by the vector fields is equal to $\mathbb{R}^d$ at any point of U, with a degree of nonholonomy r that is uniform with respect to $\tau \in \mathcal{K}$ (uniform weak Hörmander condition on U). Then there exists $\sigma > 0$ such that, for every $s \in \mathbb{R}$, for all smooth functions ζ and ζ' compactly supported in U with $\zeta' = 1$ on the support of ζ,*

$$\|\zeta u\|_{H^{s+\sigma}(\mathbb{R}^d)} \leqslant \mathrm{Cst}\,(s, \zeta, \zeta') \left(\|\zeta' L^\tau u\|_{H^s(\mathbb{R}^d)} + \|\zeta' u\|_{H^s(\mathbb{R}^d)} \right)$$

$$\forall\, u \in C^\infty\left(\mathbb{R}^d\right) \qquad \forall\, \tau \in \mathcal{K}.$$

This is a parameter-dependent version of the famous local subelliptic Hörmander estimate (see [Hör67]). It is straightforward to obtain by following the classical proofs, but for completeness, we provide the main steps (this is also useful in view of deriving global subelliptic estimates, further). The proof given in [Hör67] gives an optimal gain of regularity, which is $\sigma = 2/r$. Here, we rather follow the simpler proof given by Kohn in [Koh73] (see also [HN05]), which does not give the optimal gain of regularity and yields $\sigma = 1/2^{2r-1}$.

Note that, in the case where $\overline{U}$ is compact, Theorem B.1 implies that there exists $C > 0$ such that $\mathrm{Cst}\,\Lambda^\sigma \leqslant C\,\mathrm{id} - L^\tau - (L^\tau)^* \leqslant \mathrm{Cst}\,\Lambda^2$, for every $\tau \in \mathcal{K}$, where the inequalities are written for positive selfadjoint operators on $L^2(U)$.

Taking $s \in \{s, s+\sigma, \ldots, s+k\sigma\}$, we also infer from the theorem that, for every $k \in \mathbb{N}$, for all smooth functions ζ and ζ' compactly supported in U with $\zeta' = 1$ on the support of ζ, we have

(B.2) $\|\zeta u\|_{H^{k\sigma}(\mathbb{R}^d)}$

$$\leqslant \mathrm{Cst}\,(k, \zeta, \zeta') \left(\left\|\zeta'(L^\tau)^k u\right\|_{L^2(\mathbb{R}^d)} + \cdots + \|\zeta' L^\tau u\|_{L^2(\mathbb{R}^d)} + \|\zeta' u\|_{L^2(\mathbb{R}^d)} \right)$$

$$\forall\, u \in C^\infty\left(\mathbb{R}^d\right) \qquad \forall\, \tau \in \mathcal{K}.$$

Another useful consequence is the following. For every $k \in \mathbb{Z}$, we define the Hilbert space $\mathcal{D}_k^\tau$ as the completion for the norm $\|u\|_{\mathcal{D}_k^\tau} = \|(C\,\mathrm{id} - L^\tau - (L^\tau)^*)^k u\|_{L^2(\mathbb{R}^d)}$ of the set of smooth functions on $\mathbb{R}^d$ of compact support. For $k = 1$, the set $\mathcal{D}_1^\tau = \{u \in$

$L^2(\mathbb{R}^d) \mid L^\tau u \in L^2(\mathbb{R}^d)\}$ is the maximal domain of the operator L^τ on $L^2(\mathbb{R}^d)$. It follows from (B.2) that, for every $k \in \mathbb{N}$, for all smooth functions ζ and ζ' compactly supported in U with $\zeta' = 1$ on the support of ζ, we have

$$\mathrm{Cst}(k, \zeta, \zeta')\|\zeta u\|_{H^{k\sigma}(\mathbb{R}^d)} \leqslant \|\zeta' u\|_{\mathcal{D}_k^\tau} \leqslant \mathrm{Cst}(k, \zeta, \zeta')\|\zeta' u\|_{H^{2k}(\mathbb{R}^d)}$$

$$\forall\, u \in C^\infty\left(\mathbb{R}^d\right) \qquad \forall\, \tau \in \mathcal{K}.$$

By duality, we have the converse inequalities for k negative.

Remark B.2. — Note that, in the statement of Theorem B.1, we use cut-off functions ζ and ζ' (as in [Koh73]), in order to obtain the subelliptic estimates for any smooth function u on $\mathbb{R}^d$, and not only for any smooth function of compact support on U. This is because, in Appendix B.1.2, we are going to apply these estimates to solutions $u(t) = e^{tL^\tau}f$ of the heat equation $(\partial_t - L^\tau)u = 0$, which are not of compact support.

Proof of Theorem B.1. — We only recall the main steps of the proof, without providing all details. We modify the function $\mathbb{V}^\tau$ and the vector fields Y_j^τ (and thus L^τ) outside of U so that their coefficients and all their derivatives are uniformly bounded on $\mathbb{R}^d$, uniformly with respect to τ. Localization is performed by bracketting (see the localization lemma in [Koh73]).

First of all, we note that $(Y_j^\tau)^* = -Y_j^\tau - (\mathrm{div}\, Y_j^\tau)$, where the dual is taken in $L^2(\mathbb{R}^d)$. It follows that, given any smooth function g on $\mathbb{R}^d$, we have

$$\left|\left\langle Y_j^\tau u, gu\right\rangle_{L^2(\mathbb{R}^d)}\right| = \frac{1}{2}\left|\left\langle u, \left(g\, \mathrm{div}\, Y_j^\tau + Y_j^\tau g\right) u\right\rangle_{L^2(\mathbb{R}^d)}\right|$$

$$\leqslant \mathrm{Cst}(g)\|u\|_{L^2(\mathbb{R}^d)}^2 \qquad \forall\, u \in C_c^\infty\left(\mathbb{R}^d\right) \qquad \forall\, \tau \in \mathcal{K}.$$

We easily infer that

$$\sum_{i=1}^p \left\|Y_j^\tau u\right\|_{L^2(\mathbb{R}^d)}^2 \leqslant \mathrm{Cst}\left(\left|\langle L^\tau u, u\rangle_{L^2(\mathbb{R}^d)}\right| + \|u\|_{L^2(\mathbb{R}^d)}^2\right)$$

$$\leqslant \mathrm{Cst}\left(\|L^\tau u\|_{L^2(\mathbb{R}^d)}^2 + \|u\|_{L^2(\mathbb{R}^d)}^2\right) \quad \forall\, u \in C_c^\infty\left(\mathbb{R}^d\right) \quad \forall\, \tau \in \mathcal{K}.$$

For every $\sigma \in (0, 1]$, let $\mathcal{P}_\sigma$ be the set of pseudo-differential operators P of order 0 satisfying

$$\|Pu\|_{H^\sigma(\mathbb{R}^d)} \leqslant \mathrm{Cst}(\sigma)\left(\|L^\tau u\|_{L^2(\mathbb{R}^d)} + \|u\|_{L^2(\mathbb{R}^d)}\right) \qquad \forall\, u \in C_c^\infty\left(\mathbb{R}^d\right) \qquad \forall\, \tau \in \mathcal{K}.$$

for every smooth function u on $\mathbb{R}^d$ of compact support. Note that $0 \leqslant \sigma_1 \leqslant \sigma_2$ implies $\mathcal{P}_{\sigma_2} \subset \mathcal{P}_{\sigma_1}$. The proof consists of proving that:

    (1) $\mathcal{P}_\sigma$ is stable by taking the adjoint, for every $\sigma \in (0, 1/2]$;
    (2) $\mathcal{P}_\sigma$ is a left and right ideal in the ring of pseudo-differential operators of order 0;
    (3) $Y_j^\tau \Lambda^{-1} \in \mathcal{P}_\sigma$ for every $j \in \{0, 1, \ldots, p\}$, for every $\sigma \in (0, 1/2]$;
    (4) if $P \in \mathcal{P}_\sigma$ then $[Y_j^\tau, P] \in \mathcal{P}_{\sigma/4}$, for every $j \in \{0, 1, \ldots, p\}$ and for every $\sigma \in (0, 1/2]$.

Each of the steps is quite straightforward to establish, following [HN05, Koh73] and using elementary properties of brackets and of pseudo-differential operators. Then, noting that $[Y_j^\tau, Y_k^\tau]\Lambda^{-1} = [Y_j^\tau, Y_k^\tau\Lambda^{-1}] - Y_k^\tau\Lambda^{-1}\Lambda[Y_j^\tau, \Lambda^{-1}]$, using that $Y_k^\tau\Lambda^{-1} \in \mathcal{P}_{1/2}$ by the third property and thus that $[Y_j^\tau, Y_k^\tau\Lambda^{-1}] \in \mathcal{P}_{1/8}$, using that $\Lambda[Y_j^\tau, \Lambda^{-1}]$ is a pseudo-differential operator of order 0 and using the first and the third properties, we have $Y_k^\tau\Lambda^{-1}\Lambda[Y_j^\tau, \Lambda^{-1}] \in \mathcal{P}_{1/2}$, and we infer that $[Y_j^\tau, Y_k^\tau]\Lambda^{-1} \in \mathcal{P}_{1/8}$. By recurrence, we get that $Y_I^\tau\Lambda^{-1} \in \mathcal{P}_{1/2^{2k-1}}$, where Y_I^τ is any Lie bracket of the vector fields Y_j^τ of length k.

Using the uniform Hörmander condition, it follows that $\mathcal{P}_{1/2^{2r-1}}$ coincides with the ring of pseudo-differential operators of order 0. Indeed, we have proved that $\Lambda^{-1}\partial_{x_i} \in \mathcal{P}_{1/2^{2k-1}}$ for every $i \in \{1, \ldots, d\}$; using the properties of ideal and of stability by taking the adjoint, we have then $\Lambda^{-1}\partial_{x_i}^2\Lambda^{-1} \in \mathcal{P}_{1/2^{2k-1}}$ and thus $\Lambda^{-2}\partial_{x_i}^2 = \Lambda^{-1}\partial_{x_i}^2\Lambda^{-1} + \Lambda^{-1}[\Lambda^{-1}\partial_{x_i}^2\Lambda^{-1}, \Lambda] \in \mathcal{P}_{1/2^{2k-1}}$, and finally, $\mathrm{id} \in \mathcal{P}_{1/2^{2k-1}}$. The theorem is thus proved for $s = 0$.

We then prove the result for any $s \in \mathbb{R}$ by applying the inequality established for $s = 0$ to $\Lambda^s u$ and by managing the bracket $[L^\tau, \Lambda^s]$, using in particular the fact that $[(Y_i^\tau)^2, \Lambda^s] = 2[Y_i^\tau, \Lambda^s]Y_i^\tau + [Y_i^\tau, [Y_i^\tau, \Lambda^s]]$. □

B.1.2. Application: uniform local smoothing property for heat semigroups

Throughout the section, we assume that the operator L^τ defined by (B.1) on the domain $\mathcal{D}_1^\tau$ generates a strongly continuous semigroup $(e^{tL^\tau})_{t \geqslant 0}$ on $L^2(\mathbb{R}^d)$, for every $\tau \in \mathcal{K}$, satisfying the following uniform estimate: for all positive real numbers $t_0 < t_1$ we have

$$\text{(B.3)} \qquad \left\| e^{tL^\tau} \right\|_{L(L^2(\mathbb{R}^d))} \leqslant \mathrm{Cst}(t_0, t_1) \qquad \forall\, t \in [t_0, t_1] \qquad \forall\, \tau \in \mathcal{K}.$$

This is the case if $\mathbb{V}^\tau$ is uniformly bounded below on $\mathbb{R}^d$ and if either L^τ is selfadjoint for any τ or the vector fields $Y_0^\tau, Y_1^\tau, \ldots, Y_p^\tau$ are bounded on $\mathbb{R}^d$ as well as their first derivatives uniformly with respect to τ. Indeed, under these assumptions, reasoning as in the proof of Lemma 3.1, there exists $C \geqslant 0$ such that the operator $L^\tau - C\,\mathrm{id}$ on $L^2(\mathbb{R}^d)$ is closed and dissipative, as well as its adjoint, and thus the strongly continuous semigroup $(e^{t(L^\tau - C\,\mathrm{id})})_{t \geqslant 0}$ that it generates is a contraction semigroup, and then $(e^{tL^\tau})_{t \geqslant 0}$ is a strongly continuous semigroup, with uniform norm estimates.

We denote by e^τ the heat kernel associated with L^τ for the Lebesgue measure on $\mathbb{R}^d$. The function e^τ is defined on $(0, +\infty) \times \mathbb{R}^d \times \mathbb{R}^d$ and depends on three variables (t, x, y). In what follows, the notation ∂_1 (resp., ∂_2, ∂_3) denotes the partial derivative with respect to t (resp., to x, to y). The indices x in χ_1 and y in χ_2 below mean that, when e^τ is taken at (t, x, y), χ_1 is taken at x and χ_2 is taken at y.

COROLLARY B.3. — *We assume that the Lie algebra* $\mathrm{Lie}(Y_1^\tau, \ldots, Y_p^\tau)$ *generated by the vector fields* Y_i^τ, $i = 1, \ldots, p$, *is equal to* $\mathbb{R}^d$ *at any point of* U, *with a degree of nonholonomy* r *that is uniform with respect to* $\tau \in \mathcal{K}$ *(uniform strong*

Hörmander condition on U). Note that this assumption is more restrictive[9] than in Appendix B.1.1, because we have excluded Y_0^τ.

For all $\chi_1, \chi_2 \in C_c^\infty(U)$, for all positive real numbers $0 < t_0 < t_1$, for all $s, s' \in \mathbb{R}$, for all $(k, \alpha, \beta) \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$, we have

$$(B.4) \quad \left\| (\chi_1)_x (\chi_2)_y \partial_1^k \partial_2^\alpha \partial_3^\beta e^\tau(\cdot, \cdot, \cdot) \right\|_{L^\infty\left((t_0, t_1) \times \mathbb{R}^d \times \mathbb{R}^d\right)}$$
$$\leqslant \mathrm{Cst}\,(\chi_1, \chi_2, t_0, t_1, k, \alpha, \beta) \quad \forall\, \tau \in \mathcal{K}$$

and

$$\left\| \chi_1 e^{tL^\tau} \chi_2 f \right\|_{H^s\left(\mathbb{R}^d\right)}$$
$$\leqslant \mathrm{Cst}\,(\chi_1, \chi_2, t_0, t_1, s, s') \, \|f\|_{H^{s'}\left(\mathbb{R}^d\right)} \quad \forall\, t \in [t_0, t_1] \quad \forall\, f \in C_c^\infty\left(\mathbb{R}^d\right) \quad \forall\, \tau \in \mathcal{K}.$$

The corollary says that the family $(e^\tau)_{\tau \in \mathcal{K}}$ is bounded in $C^\infty((0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n)$, uniformly with respect to τ (for the Fréchet topology defined by seminorms on an exhaustive sequence of compact subsets of $(0, +\infty) \times \mathbb{R}^n \times \mathbb{R}^n$). It also states that, with the above notations, $\chi_1 e^{tL^\tau} \chi_2 \in L(H^{s'}(\mathbb{R}^d), H^s(\mathbb{R}^d))$ for every $t > 0$, where the norm of this operator can be bounded by a positive constant depending on χ_1 and χ_2 but not depending on τ.

Proof. — Defining on $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ the vector fields

$$\widetilde{Y}_0^\tau = \begin{pmatrix} -2 \\ Y_0^\tau \\ (Y_0^\tau)^* \end{pmatrix}, \quad \widetilde{Y}_i^\tau = \begin{pmatrix} 0 \\ Y_i^\tau \\ (Y_i^\tau)^* \end{pmatrix}, \quad i = 1, \ldots, p,$$

we consider the operator

$$P^\tau = (L^\tau)_x + (L^\tau)_{x'}^* - 2\partial_t = \sum_{i=1}^p (\widetilde{Y}_i^\tau)^2 + \widetilde{Y}_0^\tau - \mathbb{V}^\tau$$

on $L^2(\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d)$, where $(L^\tau)_x = L^\tau \otimes \mathrm{id}$ and $(L^\tau)_{x'}^* = \mathrm{id} \otimes (L^\tau)^*$ are differential operators acting on functions $g = g(x, x')$. The uniform weak Hörmander condition is satisfied on U for the $p + 1$ vector fields $(\widetilde{Y}_0^\tau, \widetilde{Y}_1^\tau, \ldots, \widetilde{Y}_p^\tau)$ on $\mathbb{R}^{1+2d}$. Applying Theorem B.1 to the heat kernel e^τ, which satisfies $P^\tau e^\tau = 0$ on $(0, +\infty) \times U \times U$, we obtain that $\|\zeta e^\tau\|_{H^{s+\sigma}(\mathbb{R}^{1+2d})} \leqslant \mathrm{Cst}(s, \zeta, \zeta') \|\zeta' e^\tau\|_{H^s(\mathbb{R}^{1+2d})}$, and by recurrence,

$$\|\zeta e^\tau\|_{H^{s+k\sigma}\left(\mathbb{R}^{1+2d}\right)} \leqslant \mathrm{Cst}(s, k, \zeta, \zeta') \|\zeta' e^\tau\|_{H^s\left(\mathbb{R}^{1+2d}\right)} \quad \forall\, \tau \in \mathcal{K} \quad \forall\, k \in \mathbb{N}^* \quad \forall\, s \in \mathbb{R}$$

for all smooth functions ζ and ζ' on $(0, +\infty) \times U \times U$ with compact support, satisfying $\zeta' = 1$ on the support of ζ. In what follows, we take $\zeta = \chi_0 \otimes \chi_1 \otimes \chi_2$ and $\zeta' = \chi_0' \otimes \chi_1' \otimes \chi_2'$. In order to obtain (B.4), we need an initialization, which is given in the following lemma.

LEMMA B.4. — *There exists $s \in \mathbb{R}$ such that*

$$\|(\chi_0')_t (\chi_1')_x (\chi_2')_{x'} e^\tau\|_{H^s\left(\mathbb{R}^{1+2d}\right)} < +\infty.$$

---

[9] This more restrictive assumption ensures that $L^\tau - \partial_t$ is hypoelliptic. Otherwise hypoellipticity may fail: take $Y_0 = \partial_{x_1}$ and all Y_i, $i \geqslant 1$, not depending on x_1: then $(Y_0, Y_1, \ldots, Y_n)$ satisfies the Hörmander condition, but not $(\partial_t + Y_0, Y_1, \ldots, Y_n)$.

Proof of Lemma B.4. Recall that $\Lambda = (\mathrm{id} - \sum_{i=1}^{d} \partial_i^2)^{1/2}$. Given any $j \in \mathbb{Z}$, we define on $\mathbb{R}^d$ the elliptic pseudo-differential operator $\Lambda'_{j,j}$ of order j by $\Lambda'_{j,j} u(x) = \langle x \rangle^j \Lambda^j u(x)$ for every $u \in C^\infty(\mathbb{R}^d)$, where $\langle x \rangle = (1 + \|x\|_2^2)^{1/2}$ (Japanese bracket), with $\|\cdot\|_2$ the Euclidean norm in $\mathbb{R}^d$. There exists $m \in \mathbb{N}^*$ such that $\Lambda'_{-m,-m}$ is Hilbert–Schmidt (as an operator on $L^2(\mathbb{R}^d)$), i.e., $\|\Lambda'_{-m,-m}\|_{HS} < +\infty$. Indeed, the Schwartz kernel of $\Lambda'_{-m,-m}$ is

$$\left[\Lambda'_{-m,-m} \right](x,y) = \frac{1}{(2\pi)^n} \frac{1}{(1 + \|x\|_2^2)^{m/2}} \int \frac{e^{i(x-y)\cdot\xi}}{(1 + \|\xi\|_2^2)^{m/2}} \, d\xi$$

which is the product of $(1 + \|x\|_2^2)^{-m/2}$ and of the Fourier transform taken at $y - x$ of the function $\xi \mapsto (1 + \|\xi\|_2^2)^{-m/2}$, and it is therefore square-integrable when m is chosen large enough. Using the general facts (concerning Schwartz kernels) recalled in Appendix A with $A = e^{tL^\tau}$, $T_1 = \Lambda'_{-m,-m}\chi'_1$ and $T_2 = \chi'_2$, we obtain that

$$\left\| (\chi_0)_t (\Lambda'_{-m,-m})_x (\chi'_1)_x (\chi'_2)_{x'} e^\tau \right\|_{L^2(\mathbb{R}^{1+2d})}$$
$$= \left\| (\chi_0)_t \Lambda'_{-m,-m} \chi'_1 e^{tL^\tau} \chi'_2 \right\|_{HS} \leqslant \left\| \Lambda'_{-m,-m} \right\|_{HS} \left\| (\chi_0)_t e^{tL^\tau} \right\|_{L(L^2(\mathbb{R}^d))} \leqslant \mathrm{Cst}(t_0, t_1)$$

when $\mathrm{supp}(\chi_0) \subset [t_0, t_1]$, where we have used the uniform estimate (B.3) to derive the latter inequality. $\qquad \square$

We have therefore proved (B.4). To get the estimate on the semigroup, it suffices to write that, for all $j, k \in \mathbb{Z}$,

$$\left\| \Lambda^j \chi_1 e^{tL^\tau} \chi_2 \Lambda^k \right\|_{L(L^2(\mathbb{R}^d))} \leqslant \left\| \Lambda^j \chi_1 e^{tL^\tau} \chi_2 \Lambda^k \right\|_{HS}$$
$$= \left\| \Lambda_x^j (\chi_1)_x \Lambda_{x'}^k (\chi_2)_{x'} e^\tau(t, \cdot, \cdot) \right\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)}$$

and then to apply (B.4). $\qquad \square$

B.2. Global estimates

In the sequel, we use the notation $\langle x \rangle = (1 + \|x\|_2^2)^{1/2}$ for $x \in \mathbb{R}^d$ (Japanese bracket), where $\|x\|_2$ is the Euclidean norm of x. Throughout the section, we assume that there exists $n \in \mathbb{N}^*$ such that the function $\mathbb{V}^\tau$ and all coefficients of the vector fields $Y_0^\tau, Y_1^\tau, \ldots, Y_p^\tau$, as well as all their derivatives, are bounded by multiples of $\langle x \rangle^n$. In other words, we assume that the growth at infinity of the vector fields Y_i^τ and of $\mathbb{V}^\tau$, as well as their derivatives, is at most polynomial.

The objective of this section is to establish global subelliptic estimates and global smoothing properties of the corresponding heat semigroup and heat kernel. We follow [EH03], where alternative global subelliptic Hörmander estimates can be found (see [EH03, Theorems 3.1 and 4.1]) but are not sufficient for our needs.

Given any $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}$, we define on $L^2(\mathbb{R}^d)$ the elliptic selfadjoint pseudo-differential operator $\Lambda_{\alpha,\beta} = \frac{1}{2}(\Lambda'_{\alpha,\beta} + (\Lambda'_{\alpha,\beta})^*)$ of order α, with

$$\Lambda'_{\alpha,\beta} u(x) = \langle x \rangle^\beta \left(\mathrm{id} - \sum_{i=1}^{d} \partial_{x_i}^2 \right)^{\alpha/2} u(x) = \langle x \rangle^\beta \Lambda^\alpha u(x).$$

and we define the weighted Hilbert space $H_\beta^\alpha(\mathbb{R}^d)$ as the completion of $C_c^\infty(\mathbb{R}^d)$ for the norm

$$\|u\|_{H_\beta^\alpha(\mathbb{R}^d)} = \|\Lambda_{\alpha,\beta} u\|_{L^2(\mathbb{R}^n)}.$$

The inner product is $\langle u, v \rangle_{H_\beta^\alpha(\mathbb{R}^d)} = \langle \Lambda_{\alpha,\beta} u, \Lambda_{\alpha,\beta} v \rangle_{L^2(\mathbb{R}^n)}$. We have the following properties (see also [EH03]):

- $\Lambda_{\alpha,\beta} \in L(H_{\beta'}^{\alpha'}(\mathbb{R}^d), H_{\beta'-\beta}^{\alpha'-\alpha}(\mathbb{R}^d))$.
- We have the continuous embedding $H_{\beta'}^{\alpha'}(\mathbb{R}^d) \hookrightarrow H_\beta^\alpha(\mathbb{R}^d)$ whenever $\alpha' \geqslant \alpha$ and $\beta' \geqslant \beta$, with a compact embedding if both inequalities are strict.
- The dual of $H_\beta^\alpha(\mathbb{R}^d)$ with respect to the pivot space $L^2(\mathbb{R}^d) = H_0^0(\mathbb{R}^d)$ is $H_{-\beta}^{-\alpha}(\mathbb{R}^d)$.

We denote by $(Y_i^\tau)_{i \geqslant 0}$ the family of vector fields consisting of the vector fields $Y_0^\tau, Y_1^\tau, \ldots, Y_p^\tau$ completed with all their successive Lie brackets.

B.2.1. Uniform global subelliptic estimates

DEFINITION B.5. — *Following* [EH03], *we say that the* uniform polynomial weak Hörmander condition *is satisfied if there exist nonnegative integers N_0 and N_1 such that*

$$(\text{B.5}) \qquad \|y\|_2^2 \leqslant \text{Cst} \langle x \rangle^{2N_0} \sum_{i=0}^{N_1} \langle Y_i^\tau(x), y \rangle^2 \qquad \forall\, (x,y) \in \mathbb{R}^d \times \mathbb{R}^d \qquad \forall\, \tau \in \mathcal{K}.$$

THEOREM B.6. — *Under the uniform polynomial weak Hörmander condition, there exist $\sigma > 0$ and $N \in \mathbb{N}$ such that*

$$\|u\|_{H_\beta^{s+\sigma}(\mathbb{R}^d)} \leqslant \text{Cst}(s,\beta) \left(\|L^\tau u\|_{H_{\beta+N}^s(\mathbb{R}^d)} + \|u\|_{H_{\beta+N}^s(\mathbb{R}^d)} \right)$$
$$\forall\, s, \beta \in \mathbb{R} \qquad \forall\, u \in C_c^\infty\left(\mathbb{R}^d\right).$$

For every $k \in \mathbb{Z}$, for every $\beta \in \mathbb{R}$, we define the Hilbert space $\mathcal{D}_{k,\beta}^\tau$ as the completion of $C_c^\infty(\mathbb{R}^d)$ for the norm

$$\|u\|_{\mathcal{D}_{k,\beta}^\tau} = \left\| (C \,\text{id} - L^\tau - (L^\tau)^*)^k \, u \right\|_{H_\beta^0(\mathbb{R}^d)}$$

where $C > 0$ is a fixed constant such that $(C \,\text{id} - L^\tau - (L^\tau)^*)^k$ is a positive selfadjoint operator on $L^2(\mathbb{R}^d)$ (of maximal domain $\mathcal{D}_{k,0}^\tau$). When $\beta = 0$, we denote simply $\mathcal{D}_k^\tau = \mathcal{D}_{k,\beta}^\tau$. For $k = 1$, the set $\mathcal{D}_1^\tau = \{u \in L^2(\mathbb{R}^d) \mid L^\tau u \in L^2(\mathbb{R}^d)\}$ is the maximal domain of the operator L^τ on $L^2(\mathbb{R}^d)$.

Note that the weight β in the iterated domains will be used only in the proof of Corollary B.9 further.

By iteration, Theorem B.6 implies that

$$\|u\|_{H_{\beta-kN}^{s+k\sigma}(\mathbb{R}^d)} \leqslant \text{Cst}(s,k,\beta) \left(\|(L^\tau)^k u\|_{H_\beta^s(\mathbb{R}^d)} + \cdots + \|L^\tau u\|_{H_\beta^s(\mathbb{R}^d)} + \|u\|_{H_\beta^s(\mathbb{R}^d)} \right)$$
$$\forall\, s, \beta \in \mathbb{R} \qquad \forall\, k \in \mathbb{N}^* \qquad \forall\, u \in C^\infty\left(\mathbb{R}^d\right) \qquad \forall\, \tau \in \mathcal{K}$$

and therefore we have the continuous embeddings

$$(B.6) \qquad H^{2k}_{\beta+2kN}\left(\mathbb{R}^d\right) \hookrightarrow \mathcal{D}^\tau_{k,\beta} \hookrightarrow H^{k\sigma}_{\beta-kN}\left(\mathbb{R}^d\right) \quad \forall\, k \in \mathbb{N} \quad \forall\, \beta \in \mathbb{R} \quad \forall\, \tau \in \mathcal{K}$$

with constants depending on k but not on τ. By duality, we have

$$(B.7) \qquad H^{-k\sigma}_{\beta+kN}\left(\mathbb{R}^d\right) \hookrightarrow \mathcal{D}^\tau_{-k,\beta} \hookrightarrow H^{-2k}_{\beta-2kN}\left(\mathbb{R}^d\right) \quad \forall\, k \in \mathbb{N} \quad \forall\, \beta \in \mathbb{R} \quad \forall\, \tau \in \mathcal{K}.$$

Remark B.7. — Actually, by bracketting (see the sketch of proof of the theorem hereafter), we obtain the following slight variant of Theorem B.6: for all smooth functions ζ and ζ' on $\mathbb{R}^n$, which are globally bounded as well as all their derivatives, satisfying $\zeta' = 1$ on the support of ζ, we have

$$\|\zeta u\|_{H^{s+\sigma}_{\beta-N}\left(\mathbb{R}^d\right)} \leqslant \mathrm{Cst}(s,\beta,\zeta,\zeta')\left(\|\zeta' L^\tau u\|_{H^s_\beta\left(\mathbb{R}^d\right)} + \|\zeta' u\|_{H^s_\beta\left(\mathbb{R}^d\right)}\right)$$

$$\forall\, s,\beta \in \mathbb{R} \quad \forall\, u \in C^\infty\left(\mathbb{R}^d\right) \quad \forall\, \tau \in \mathcal{K}.$$

Note that the functions ζ and ζ' are not assumed here to be of compact support. This fact will be useful in Appendix B.2.2 to localize in time only (but not in space).

Proof of Theorem B.6. For every $m \in \mathbb{R}$ and every $\delta \in \mathbb{R}$, we define the following class of global symbols $\mathcal{S}^{m,\delta}$, by considering their growth in x: a symbol $a : \mathbb{R}^d \times \mathbb{R}^d \setminus \{0\} \to \mathbb{C}$ (of order m) is a smooth function having an asymptotic expansion $a \sim \sum_{j=0}^{+\infty} a_{m-j}$, where a_{m-j} is smooth and such that, for any $k \in \mathbb{N}$, the function $a_k = a - \sum_{j=0}^{k-1} a_{m-j}$ satisfies

$$\left|\partial_x^\alpha \partial_\xi^\beta a_k(x,\xi)\right| \leqslant \mathrm{Cst}(\alpha,\beta)\langle x\rangle^{\delta-|\alpha|}(1+\|\xi\|_2)^{m-k-|\beta|} \qquad \forall\, \alpha,\beta \in \mathbb{N}^d.$$

We denote by $\Psi^{m,\delta}$ the set of pseudo-differential operators on $\mathbb{R}^d$ whose symbol belongs to $\mathcal{S}^{m,\delta}$. We have the following useful property: if $a \in \mathcal{S}^{m,\delta}$ then $\mathrm{Op}(a) \in L(H^\alpha_\beta(\mathbb{R}^d), H^{\alpha-m}_{\beta-\delta}(\mathbb{R}^d))$. Here, Op is the usual left (for instance) quantization operator.

Now, if we consider in $\mathcal{S}^{m,\delta}$ a family of symbols $(a^\tau)_{\tau\in\mathcal{K}}$, with the above constants $\mathrm{Cst}(\alpha,\beta)$ being uniform with respect to τ, then we have the uniform estimate

$$\|\mathrm{Op}(a^\tau)\|_{L\left(H^\alpha_\beta\left(\mathbb{R}^d\right), H^{\alpha-m}_{\beta-\delta}\left(\mathbb{R}^d\right)\right)} \leqslant \mathrm{Cst}(\alpha,\beta) \quad \forall\, \tau \in \mathcal{K}.$$

Thanks to this general remark, we obtain that $Y_i^\tau \in L(H^\alpha_\beta(\mathbb{R}^d), H^{\alpha-1}_{\beta-N}(\mathbb{R}^d))$, that $\Lambda^{-1}Y_i^\tau \in L(H^\alpha_\beta(\mathbb{R}^d), H^\alpha_{\beta-N}(\mathbb{R}^d))$ and that $[Y_i^\tau, \Lambda^\gamma] \in L(H^\alpha_\beta(\mathbb{R}^d), H^{\alpha-\gamma}_{\beta-N}(\mathbb{R}^d))$, with uniform constants in the estimates (see also [EH03, Lemma 3.2] for other useful properties of the spaces $H^\alpha_\beta(\mathbb{R}^d)$).

These preliminaries being done, we follow the steps of the proof of Theorem B.1, keeping track of constants. It is then straightforward to establish that

$$\sum_{j=1}^p \left\|Y_j^\tau u\right\|^2_{L^2\left(\mathbb{R}^d\right)} \leqslant \mathrm{Cst}\left(\left|\langle L^\tau u, u\rangle_{L^2\left(\mathbb{R}^d\right)}\right| + \langle x\rangle^{2N}\|u\|^2_{L^2\left(\mathbb{R}^d\right)}\right)$$

$$\|u\|_{H^\sigma\left(\mathbb{R}^d\right)} \leqslant \mathrm{Cst}\langle x\rangle^N\left(\|L^\tau u\|_{L^2\left(\mathbb{R}^d\right)} + \|u\|_{L^2\left(\mathbb{R}^d\right)}\right)$$

$$\forall\, u \in C_c^\infty(B(x,1)) \qquad \forall\, x \in \mathbb{R}^d \qquad \forall\, \tau \in \mathcal{K}$$

where $N \in \mathbb{N}$ is a sufficiently large integer. Then, noting that, roughly, $\|u\|_{H^\alpha_\beta(\mathbb{R}^d)} \simeq \langle x \rangle^\beta \|u\|_{H^\alpha(\mathbb{R}^d)}$ for every $u \in C_c^\infty(B(x,1))$ and for all $\alpha, \beta \in \mathbb{R}$, using a partition of unity (see [EH03, Proof of Theorem 4.1] for details), we obtain

$$\|u\|_{H^\sigma(\mathbb{R}^d)} \leqslant \mathrm{Cst}\left(\|L^\tau u\|_{H^0_N(\mathbb{R}^d)} + \|u\|_{H^0_N(\mathbb{R}^d)}\right) \quad \forall\, u \in C_c^\infty\left(\mathbb{R}^d\right) \quad \forall\, \tau \in \mathcal{K}.$$

Using the brackets $[\Lambda_{s,\beta}, Y_i^\tau]$, $[\Lambda_{s,\beta}, L^\tau]$, increasing the integer N if necessary (but N remains uniform with respect to s and β), we then obtain

$$\sum_{j=1}^p \left\|Y_j^\tau u\right\|^2_{H^s_\beta(\mathbb{R}^d)} \leqslant \mathrm{Cst}(s,\beta)\left(\left|\langle L^\tau u, u\rangle_{H^s_\beta(\mathbb{R}^d)}\right| + \|u\|^2_{H^s_{\beta+N}(\mathbb{R}^d)}\right)$$

$$\forall\, u \in C_c^\infty\left(\mathbb{R}^d\right) \quad \forall\, \tau \in \mathcal{K}$$

and we then establish the global estimate of Theorem B.6. We do not provide any details. $\qquad\square$

Remark B.8. — It is interesting to note that the uniform polynomial weak Hörmander condition (B.5) is satisfied for any m-tuple $(Y_0, Y_1, \ldots, Y_p)$ of polynomial vector fields that satisfy the weak Hörmander condition at every point with a uniform degree of nonholonomy.

Indeed, defining as above the family $(Y_i)_{i \geqslant 0}$ by completing the $(p+1)$-tuple $(Y_0, \ldots, Y_p)$ with their iterated Lie brackets, by the assumption, there exists a large enough integer N_1 such that the span of the finite family $(Y_i)_{0 \leqslant i \leqslant N_1}$ is equal to $\mathbb{R}^n$ at every point. Defining the matrix

$$P(x) = \begin{pmatrix} Y_0(x) & Y_1(x) & \cdots & Y_{N_1}(x) \end{pmatrix} \quad \forall\, x \in \mathbb{R}^n,$$

we have to prove that there exists $N_0 \in \mathbb{N}$ such that

(B.8) $\qquad y^\top P(x) P(x)^\top y \geqslant \mathrm{Cst}\langle x \rangle^{-2N_0} \|y\|_2^2 \qquad \forall\, x, y \in \mathbb{R}^n.$

By [Hör58, Lemma 2], it is known that, for every polynomial function Q on $\mathbb{R}^n$ such that $Q(x) > 0$ for every $x \in \mathbb{R}^n$, there exists an integer $N \in \mathbb{N}$ such that $Q(x) \geqslant \frac{\mathrm{Cst}}{\langle x \rangle^{2N}}$. In other words, the decay at infinity of any polynomial is polynomial, it cannot be arbitrarily small. This non-obvious fact follows from a Łojasiewicz inequality combined with an inversion argument. The existence of N_0 such that (B.8) is satisfied follows.

B.2.2. Application: uniform global smoothing property of heat semigroups

We say that the *uniform polynomial strong Hörmander assumption* is satisfied if there exist nonnegative integers N_0 and N_1 such that

(B.9) $\qquad \|y\|_2^2 \leqslant \mathrm{Cst}\, \langle x \rangle^{2N_0} \sum_{i=1}^{N_1} \langle Y_i^\tau(x), y \rangle^2 \quad \forall\, (x,y) \in \mathbb{R}^d \times \mathbb{R}^d \quad \forall\, \tau \in \mathcal{K}.$

Compared with (B.5), in (B.9) we have excluded the vector field Y_0^τ: the sum starts at $i = 1$. By a similar argument as the one elaborated in Remark B.8, note that (B.9) is satisfied for polynomial vector fields satisfying the strong H"ormander condition at every point, with a uniform degree of nonholonomy.

Throughout this section, we assume that the operator $L^\tau : D(L^\tau) = \mathcal{D}_1^\tau \to L^2(\mathbb{R}^d)$ generates a strongly continuous semigroup $(e^{tL^\tau})_{t \geqslant 0}$, for every $\tau \in \mathcal{K}$, satisfying the uniform estimate: for all positive real numbers $t_0 < t_1$, we have

(B.10) $$\left\| e^{tL^\tau} \right\|_{L\left(L^2(\mathbb{R}^d)\right)} \leqslant \mathrm{Cst}(t_0, t_1) \quad \forall \, t \in [t_0, t_1].$$

We denote by e^τ the heat kernel associated with L^τ for the Lebesgue measure on $\mathbb{R}^d$. The function e^τ is defined on $(0, +\infty) \times \mathbb{R}^d \times \mathbb{R}^d$ and depends on three variables (t, x, y). In what follows, the notation ∂_1 (resp., ∂_2, ∂_3) denotes the partial derivative with respect to t (resp., to x, to y).

Corollary B.9. — *Under the uniform polynomial strong Hörmander assumption, there exist $s_0 > 0$ and $k_0 \in \mathbb{N}^*$ such that, for all $s \geqslant s_0$ and $s' \geqslant s_0$, for all positive real numbers $0 < t_0 < t_1$,*

$$\left\| (\Lambda_{s,-k_0 s})_x \, (\Lambda_{s',-k_0 s'})_y \, e^\tau(\cdot, \cdot, \cdot) \right\|_{L^2\left((t_0, t_1) \times \mathbb{R}^d \times \mathbb{R}^d\right)} \leqslant \mathrm{Cst}(s, s', t_0, t_1) \quad \forall \, \tau \in \mathcal{K}$$

and

$$\left\| e^{tL^\tau} \right\|_{L\left(H^{-s'}_{k_0 s'}(\mathbb{R}^d), H^s_{-k_0 s}(\mathbb{R}^d)\right)} \leqslant \mathrm{Cst}(s, s', t_0, t_1) \quad \forall \, t \in [t_0, t_1] \quad \forall \, \tau \in \mathcal{K}$$

This result is a global smoothing property for hypoelliptic heat kernels in Sobolev spaces with polynomial weight.

Proof. — We follows the same lines as in the proof of Corollary B.3, by considering the operator P^τ. It follows from Remark B.7 that, for all $s, \beta \in \mathbb{R}$, for every $k \in \mathbb{N}$, we have $\|\zeta e^\tau\|_{H^{s+k\sigma}_{\beta - kN}(\mathbb{R}^{1+2d})} \leqslant \mathrm{Cst}(s, k, \beta, \zeta, \zeta') \|\zeta' e^\tau\|_{H^s_\beta(\mathbb{R}^{1+2d})}$, for all smooth functions ζ and ζ' on $(0, +\infty) \times \mathbb{R}^d \times \mathbb{R}^d$ satisfying $\zeta' = 1$ on the support of ζ, and such that ζ and ζ' and all their derivatives are bounded by constants. We choose $\zeta = (\chi_0)_t$ and $\zeta' = (\chi_0')_t$ only depending on t, in order to localize in time over $t \in [t_0, t_1]$.

To initialize the bootstrap argument, we first observe that, like in the proof of Corollary B.3, using (B.10), there exists $m \in \mathbb{N}^*$ such that

$$\left\| (\chi_0')_t \, (\Lambda_{-m,-m})_x \, e^\tau \right\|_{L^2(\mathbb{R}^{1+2d})} \leqslant \mathrm{Cst}(t_0, t_1)$$

(assuming that $\mathrm{supp}(\chi_0') \subset [t_0, t_1]$), hence

$$\left\| (\chi_0')_t e^\tau \right\|_{H^{-m}_{-m}(\mathbb{R}^{1+2d})} \leqslant \mathrm{Cst}(t_0, t_1).$$

Therefore, taking $s = \beta = -m$, we get that $\|(\chi_0)_t e^\tau\|_{H^{-m+k\sigma}_{-m-kN}(\mathbb{R}^{1+2d})} \leqslant \mathrm{Cst}(k, t_0, t_1)$ for every $k \in \mathbb{N}^*$. The result follows. $\qquad\square$

Appendix C. Global smoothing properties of sR heat semigroups

This section can be read independently of the rest.

C.1. Global smoothing properties in iterated domains

Throughout this section, we consider a selfadjoint sR Laplacian

$$\triangle_{\mathrm{sR}} = -\sum_{i=1}^{m} X_i^* X_i$$

(see Section 2.2) on $L^2(\mathbb{R}^n)$, where $X_1, \ldots, X_m$ are smooth vector fields on $\mathbb{R}^n$, satisfying the Hörmander condition at every point of $\mathbb{R}^n$. We denote by $(e^{t\triangle_{\mathrm{sR}}})_{t \geqslant 0}$ the associated sR heat semigroup, and by e the sR heat kernel.

We show how to use the Kannai transform to establish global smoothing properties of the sR heat semigroup in the scale of Sobolev spaces associated with $\triangle_{\mathrm{sR}}$ (iterated domains $D((\mathrm{id} - \triangle_{\mathrm{sR}})^j)$, for $j \in \mathbb{Z}$). This technique is well known (see [CGT82]) and, conveniently combined with complex analysis (e.g., Phragmen-Lindelöf principle, see [CS08]), leads to exponential estimates of heat kernels. However, we recall some of the precise arguments and we give statements that we have not been able to find in this form in the existing literature.

Since the Kannai transform is based on the spectral theorem, it requires to consider a selfadjoint operator. This is why, in this section, we only consider selfadjoint sR Laplacians.

C.1.1. Rough global smoothing properties in iterated domains

Starting from the inequality

$$\left(1 + \lambda^2\right)^k e^{-t\lambda^2} \leqslant \frac{\mathrm{Cst}(k, t_1)}{t^k} \quad \forall \lambda \geqslant 0 \quad \forall t_1 > 0 \quad \forall t \in (0, t_1] \quad \forall k \in \mathbb{Z},$$

we infer from the spectral theorem applied to the selfadjoint operator $\triangle_{\mathrm{sR}}$ that $\|(\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}}\|_{L(L^2(\mathbb{R}^n))} \leqslant \frac{\mathrm{Cst}(k, t_1)}{t^k}$. Hence, given any $t > 0$ and any $j, k \in \mathbb{Z}$, the operator $e^{t\triangle_{\mathrm{sR}}}$ is bounded as an operator from $\mathcal{D}_j = D((\mathrm{id} - \triangle_{\mathrm{sR}})^j)$ to $\mathcal{D}_k = D((\mathrm{id} - \triangle_{\mathrm{sR}})^k)$, and we obtain the following result.

LEMMA C.1. — *Given any $t > 0$ and any $j, k \in \mathbb{Z}$, the operator $e^{t\triangle_{\mathrm{sR}}}$ is bounded as an operator from $\mathcal{D}_j = D((\mathrm{id} - \triangle_{\mathrm{sR}})^j)$ to $\mathcal{D}_k = D((\mathrm{id} - \triangle_{\mathrm{sR}})^k)$, and*

$$\left\| e^{t\triangle_{\mathrm{sR}}} \right\|_{L(\mathcal{D}_j, \mathcal{D}_k)} \leqslant \begin{cases} \mathrm{Cst}(k - j, t_1)\frac{1}{t^{2(k-j)}} & \text{if } k \geqslant j, \\ \mathrm{Cst}(k) & \text{if } k \leqslant j, \end{cases} \quad \forall t_1 > 0 \quad \forall t \in (0, t_1].$$

This property is completely general and is even true without any Hörmander condition on the vector fields X_i: it is satisfied for any selfadjoint nonpositive operator.

We establish in Section C.1 much finer global smoothing properties, thanks to the Kannai transform that we recall in the next section.

C.1.2. Kannai transform

Hereafter, we recall a classical argument developed in [CGT82], consisting of obtaining estimates of the heat semigroup (in the Riemannian case), via the Kannai transform, Here, we are in the subelliptic case and we consider a sR Laplacian $\triangle_{\mathrm{sR}}$ on $\mathbb{R}^n$: this is a nonpositive selfadjoint operator on $L^2(\mathbb{R}^n)$, of domain $D(\triangle_{\mathrm{sR}})$. We define the sR wave operator $\cos(t\sqrt{-\triangle_{\mathrm{sR}}})$ on $L^2(\mathbb{R}^n)$ as follows: $u(t, \cdot) = \cos\left(t\sqrt{-\triangle_{\mathrm{sR}}}\right)\delta_y$ is the solution to

$$(\partial_{tt} - \triangle_{\mathrm{sR}})u = 0,$$
$$u(0, \cdot) = \delta_y,$$
$$\partial_t u(0, \cdot) = 0.$$

For every $t \in \mathbb{R}$ we have

(C.1)
$$\left\| \cos\left(t\sqrt{-\triangle_{\mathrm{sR}}}\right) \right\|_{L(L^2(\mathbb{R}^n))} \leqslant 1.$$

Given any smooth even real-valued function $F \in L^2(\mathbb{R})$, its Fourier transform is defined by $\widehat{F}(s) = \int_{\mathbb{R}} F(\lambda)\cos(\lambda s)\,ds$ and we have

$$F(\lambda) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{F}(s)\cos(\lambda s)\,ds = \frac{1}{\pi}\int_0^{+\infty} \widehat{F}(s)\cos(\lambda s)\,ds.$$

As a consequence of the spectral theorem, we have

$$F\left(\sqrt{-\triangle_{\mathrm{sR}}}\right) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{F}(s)\cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right)\,ds = \frac{1}{\pi}\int_0^{+\infty} \widehat{F}(s)\cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right)\,ds.$$

In particular, fixing some arbitrary $t \geqslant 0$ and taking $F(\lambda) = e^{-t\lambda^2}$, we have $\widehat{F}(s) = \sqrt{\frac{\pi}{t}}e^{-s^2/4t}$ and thus

$$e^{-t\lambda^2} = \frac{1}{2\sqrt{\pi t}}\int_{\mathbb{R}} e^{-s^2/4t}\cos(s\lambda)\,ds = \frac{1}{\sqrt{\pi t}}\int_0^{+\infty} e^{-s^2/4t}\cos(s\lambda)\,ds$$

for every $\lambda \in \mathbb{R}$, and therefore

$$e^{t\triangle_{\mathrm{sR}}} = \frac{1}{2\sqrt{\pi t}}\int_{\mathbb{R}} e^{-s^2/4t}\cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right)\,ds = \frac{1}{\sqrt{\pi t}}\int_0^{+\infty} e^{-s^2/4t}\cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right)\,ds$$

for every $t > 0$. This formula is usually called transmutation, or Kannai transform (see [Kan77]).

We also have

$$e^{-t\left(1+\lambda^2\right)} = \frac{1}{2\sqrt{\pi t}}\int_{\mathbb{R}} e^{-t-s^2/4t}\cos(s\lambda)\,ds = \frac{1}{\sqrt{\pi t}}\int_0^{+\infty} e^{-t-s^2/4t}\cos(s\lambda)\,ds.$$

Differentiating k times with respect to t, for some $k \in \mathbb{N}$, one gets

$$\left(1+\lambda^2\right)^k e^{-t(1+\lambda^2)} = \int_0^{+\infty} \frac{d^k}{dt^k}\left(\frac{1}{\sqrt{\pi t}}e^{-t-s^2/4t}\right)\cos(s\lambda)\,ds$$

i.e.,

$$\left(1+\lambda^2\right)^k e^{-t\lambda^2} = \int_0^{+\infty} \frac{P_{2k}(s,t)}{t^{2k+1/2}}e^{-s^2/4t}\cos(s\lambda)\,ds$$

where P_{2k} is a polynomial of the two variables s and t, of degree $2k$ (actually, depending on s^2 and of t), and thus

$$(\text{C.2}) \quad (\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}} = \int_0^{+\infty} \frac{P_{2k}(s,t)}{t^{2k+1/2}} e^{-s^2/4t} \cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right) ds$$

$$\forall\, t > 0 \quad \forall\, k \in \mathbb{N} \quad \forall\, j \in \mathbb{Z}.$$

C.1.3. Exponential estimates, using the finite speed propagation property

Let us establish an additional estimate, by using the finite speed propagation property for the sR wave equation. The finite speed propagation property for sR wave equations has been established in [Mel85]; actually, it follows from the finite speed propagation property for usual Riemannian wave equations, by considering a Riemannian ε-regularization of the sR Laplacian (see [Ge93]) and then passing to the limit, using the fact that the ε-regularized Riemannian distance converges to the sR distance as ε tends to 0 (see also [Gro96, Section 1.4.D] and [Var90, Appendix A4]).

Let $x \in \mathbb{R}^n$ and let $0 < r_1 < r_2$ be arbitrary. In what follows, we denote by $\mathrm{B}_{\mathrm{sR}}(x, r_i)$ the sR ball in $\mathbb{R}^n$ of center x and of radius r_i, for the sR distance d_{sR} associated with $\triangle_{\mathrm{sR}}$ (we assume that the Hörmander condition is satisfied).

Given any smooth function f on $\mathbb{R}^n$ such that $\mathrm{supp}(f) \subset \mathrm{B}_{\mathrm{sR}}(x, r_1)$, let us estimate $\|(\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}} f\|_{L^2(\mathbb{R}^n \setminus \mathrm{B}_{\mathrm{sR}}(x, r_2))}$. Since $\mathrm{supp}(f) \subset \mathrm{B}_{\mathrm{sR}}(x, r_1)$, we have, by the finite speed propagation property,

$$\mathrm{supp}\left(\cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right) f\right) \subset \mathrm{B}_{\mathrm{sR}}(x, r_1 + s) \qquad \forall\, s \geqslant 0.$$

In other words, a minimal time $s \geqslant r_2 - r_1$ is required in order to transport information from $\mathrm{B}_{\mathrm{sR}}(x, r_1)$ to $\mathbb{R}^n \setminus \mathrm{B}_{\mathrm{sR}}(x, r_2)$. Therefore, using (C.1) and (C.2) we get that

$$(\text{C.3}) \quad \left\|(\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}} f\right\|_{L^2(\mathbb{R}^n \setminus \mathrm{B}_{\mathrm{sR}}(x, r_2))}$$

$$= \left\|\int_{r_2 - r_1}^{+\infty} \frac{P_{2k}(s,t)}{t^{2k+1/2}} e^{-s^2/4t} \cos\left(s\sqrt{-\triangle_{\mathrm{sR}}}\right) f\, ds\right\|_{L^2(\mathbb{R}^n)}$$

$$\leqslant \int_{r_2 - r_1}^{+\infty} \frac{|P_{2k}(s,t)|}{t^{2k+1/2}} e^{-s^2/4t}\, ds\, \|f\|_{L^2(\mathbb{R}^n)}$$

$$\leqslant \mathrm{Cst}(k) \frac{1 + t^{2k}}{t^{2k}} \left(1 + (r_2 - r_1)^{2k}\right) e^{-(r_2 - r_1)^2/4t} \|f\|_{L^2(\mathbb{R}^n)}$$

$$\forall\, t > 0 \quad \forall\, k \in \mathbb{N} \quad \forall\, f \in C_c^\infty(\mathbb{R}^n) \text{ s.t. } \mathrm{supp}(f) \subset \mathrm{B}_{\mathrm{sR}}(x, r_1).$$

Note that $r_2 - r_1 = \mathrm{d}_{\mathrm{sR}}(\mathrm{B}_{\mathrm{sR}}(x, r_1), \mathbb{R}^n \setminus \mathrm{B}_{\mathrm{sR}}(x, r_2))$. To obtain the latter inequality, we have used the inequalities (proved by recurrence)

$$\int_a^{+\infty} e^{-s^2/4t}\, ds \leqslant \mathrm{Cst}\ \sqrt{t}\, e^{-a^2/4t},$$

$$\int_a^{+\infty} s^j e^{-s^2/4t}\, ds \leqslant \mathrm{Cst}(j) t \left(1 + t^{(j-1)/2}\right) \left(1 + a^{j-1}\right) e^{-a^2/4t}$$

$$\forall\, j \in \mathbb{N}^* \quad \forall\, a > 0 \quad \forall\, t > 0.$$

A consequence of this analysis is the following.

PROPOSITION C.2. — *Let χ and χ' be smooth functions on compact support on $\mathbb{R}^n$, such that $\chi' = 1$ on $\mathrm{supp}(\chi)$. Then, for every $k \in \mathbb{N}$ and every $t > 0$, the operator $(1 - \chi')(\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}}\chi$ is bounded on $L^2(\mathbb{R}^n)$, and*

$$\left\|(1 - \chi')(\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}}\chi\right\|_{L(L^2(\mathbb{R}^n))} \leqslant \mathrm{Cst}(k, \chi, \chi')\frac{1 + t^{2k}}{t^{2k}} e^{-\,\mathrm{Cst}(\chi, \chi')/t} \qquad \forall\, t > 0.$$

The same property holds for the operator $(\mathrm{id} - \triangle_{\mathrm{sR}})^k(1 - \chi')e^{t\triangle_{\mathrm{sR}}}\chi$ and (taking the adjoint and by selfadjointness) for the operators $\chi(\mathrm{id} - \triangle_{\mathrm{sR}})^k e^{t\triangle_{\mathrm{sR}}}(1 - \chi')$ and $(\mathrm{id} - \triangle_{\mathrm{sR}})^k\chi e^{t\triangle_{\mathrm{sR}}}(1 - \chi')$.

Proof. — Using a partition of unity, we reduce the proof to the case where $\mathrm{supp}(\chi) \subset \mathrm{B}_{\mathrm{sR}}(\bar{x}, r)$ and $\chi' = 1$ on $\mathrm{B}_{\mathrm{sR}}(\bar{x}, 2r)$ for some $\bar{x} \in \mathbb{R}^n$ and some $r > 0$, so that $\mathrm{supp}(\chi') \subset \mathbb{R}^n \setminus \mathrm{B}_{\mathrm{sR}}(\bar{x}, 2r)$. Then (C.3) gives the conclusion. $\qquad\square$

Note that all estimates that we have obtained are uniform with respect to $t > 0$.

C.2. Global smoothing properties for nilpotent sR Laplacians

In this section, we consider the nilpotent sR Laplacian $\widehat{\triangle}^q$, associated with the nilpotentization $(\widehat{M}^q, \widehat{D}^q, \widehat{g}^q)$ of the sR structure (M, D, g) at q. Recall that $\widehat{M}^q$ is a homogeneous space of a Carnot group (see Section 2.5.4), but is not a Carnot group whenever q is singular.

We will establish global smoothing properties for the semi-group $(e^{t\widehat{\triangle}^q})_{t \geqslant 0}$ in iterated domains and Sobolev spaces with polynomial weight.

Note that, since we deal in this section with polynomial vector fields, we could apply the smoothing properties established in Appendix B.2 (more precisely, see Corollary B.9) by deriving global subelliptic estimates in Sobolev spaces with polynomial weight. But here, in this specific context, the global smoothing properties that we are going to establish are much stronger.

C.2.1. Reminders on exponential estimates

There are various ways to establish exponential estimates for heat kernels. Such estimates have been proved in [Var90] by means of large deviation theory, using the probabilistic interpretation of the sR heat kernel (see also [KS88] for long time estimates). We also refer the reader to [SC84] and [JSC86] for estimates of the sR heat kernel and of its derivatives, established, however, on a compact manifold.

In [CGT82] (see also [Mel85]) the authors show how to combine the Kannai transform (i.e., the finite propagation speed of waves) with the Harnack principle and with a classical Moser iteration argument. In [SC92], the author combines the Kannai transform with the Harnack parabolic principle in the hypoelliptic case. In [CS08], the authors use an elegant Phragmen-Lindelöf argument to obtain off-diagonal exponential estimates.

Finally, we quote the paper [Mah98], in which it is shown that all known exponential estimates remain valid as well when considering a sR Laplacian on a homogeneous space (thus, extending results of [Var90] that were established on Lie groups).

According to these references, denoting by $\widehat{e}^q$ the heat kernel of the sR Laplacian $\widehat{\triangle}^q = \sum_{i=1}^m (\widehat{X}_i^q)^2$, for every $\varepsilon > 0$, for all $i_1, \ldots, i_s \in \{1, \ldots, m\}$ with $s \in \mathbb{N}^*$, we have

$$(\text{C.4}) \quad \left| \partial_t^m X_q^I \widehat{e}^q(t, x, y) \right| \leqslant \frac{\text{Cst}(\varepsilon, I)}{t^{m+|I|/2}} \frac{1}{\text{Vol}\left(\widehat{B}^q\left(y, \sqrt{t}\right)\right)} \exp\left(\frac{-\widehat{\text{d}}^q(x, y)^2}{4(1+\varepsilon)t}\right)$$

$$\forall\, t > 0 \quad \forall\, x, y \in \widehat{M}^q \simeq \mathbb{R}^n$$

where $X^I = \widehat{X}_{i_1} \cdots \widehat{X}_{i_s}$, $I = (i_1, \ldots, i_s)$ and $|I| = s$. In other words, the heat kernel $\widehat{e}^q$ and its derivatives are exponentially decreasing off the diagonal. Note that $\text{Vol}(\widehat{B}^q(y, \sqrt{t})) = t^{\mathcal{Q}(y)/2}$. By symmetry of the heat kernel, one can obtain the inequality (C.4) with $t^{\mathcal{Q}(y)/2}$ replaced by $t^{\mathcal{Q}(x)/2}$, or even by $t^{\mathcal{Q}(x)/4} t^{\mathcal{Q}(y)/4}$. Note anyway that all these inequalities are different, because $\mathcal{Q}(x) \neq \mathcal{Q}(y)$ in general, unless $\widehat{M}^q$ is a Carnot group (which is the case when q is regular).

C.2.2. Global smoothing properties

As in Appendix B.2.1, given any $j \in \mathbb{Z}$ and any $\alpha \in \mathbb{R}$, we define $\widehat{\mathcal{D}}_{j,\alpha}^q$ as the completion of $C_c^\infty(\mathbb{R}^n)$ for the norm

$$\|u\|_{\widehat{\mathcal{D}}_{j,\alpha}^q} = \left\| \left(\text{id} - \widehat{\triangle}^q\right)^j u \right\|_{H_\alpha^0(\mathbb{R}^n)} = \left\| \langle x \rangle^\alpha \left(\text{id} - \widehat{\triangle}^q\right)^j u \right\|_{L^2(\mathbb{R}^n)}.$$

Here, we recall that the Japanese bracket is defined by $\langle x \rangle = (1 + \|x\|_2^2)^{1/2}$ where $\|x\|_2^2 = x_1^2 + \cdots + x_n^2$ is the Euclidean norm.

It is useful to define the *sR Japanese bracket* by $\langle x \rangle_{\text{sR}} = (1 + \|x\|_{\text{sR}}^2)^{1/2}$, where the sR pseudo-norm is defined by $\|x\|_{\text{sR}} = \sum_{i=1}^n |x_i|^{1/w_i}$ for every $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ in privileged coordinates around 0 (here, $w_i = w_i(q)$). Accordingly, we define $\widehat{\mathcal{D}}_{j,\alpha}^{q,\text{sR}}$ as the completion of $C_c^\infty(\mathbb{R}^n)$ for the norm

$$\|u\|_{\widehat{\mathcal{D}}_{j,\alpha}^{q,\text{sR}}} = \left\| \langle x \rangle_{\text{sR}}^\alpha \left(\text{id} - \widehat{\triangle}^q\right)^j u \right\|_{L^2(\mathbb{R}^n)}.$$

Setting $r = w_n$ (degree of nonholonomy at 0), we have the inequality

$$\text{Cst}\, \langle x \rangle^{1/r} \leqslant \langle x \rangle_{\text{sR}} \leqslant \text{Cst}\, \langle x \rangle \qquad \forall\, x \in \mathbb{R}^n$$

from which it follows that

(C.5) $\quad \mathrm{Cst}\, \|u\|_{\widehat{\mathcal{D}}^q_{j,\,\alpha/r}} \leqslant \|u\|_{\widehat{\mathcal{D}}^{q,\,\mathrm{sR}}_{j,\,\alpha}} \leqslant \mathrm{Cst}\, \|u\|_{\widehat{\mathcal{D}}^q_{j,\,\alpha}}$

$$\forall\, j \in \mathbb{Z} \quad \forall\, \alpha > 0 \quad \forall\, u \in C^\infty_c(\mathbb{R}^n).$$

PROPOSITION C.3. — *Given any $t > 0$, any $j, k \in \mathbb{Z}$ and any real number $\beta \geqslant 1$, the operator $e^{t\widehat{\triangle}^q}$ is bounded as an operator from $\widehat{\mathcal{D}}^{q,\,\mathrm{sR}}_{j,\,\beta}$ to $\widehat{\mathcal{D}}^{q,\,\mathrm{sR}}_{k,\,\beta}$, and*
(C.6)

$$\|e^{t\widehat{\triangle}^q}\|_{L\left(\widehat{\mathcal{D}}^{q,\,\mathrm{sR}}_{j,\,\beta},\,\widehat{\mathcal{D}}^{q,\,\mathrm{sR}}_{k,\,\beta}\right)} \leqslant \begin{cases} \frac{\mathrm{Cst}(j,k,\beta,t_1)}{t^{k-j}} & \text{if } k \geqslant j, \\ \mathrm{Cst}(k,\beta,t_1) & \text{if } k \leqslant j, \end{cases} \quad \forall\, t \in (0, t_1] \quad \forall\, t_1 > 0.$$

This result states a global smoothing property in the iterated domains with polynomial weight, by keeping the same weight $\beta \geqslant 1$. In particular, for $j = k$, we have

(C.7) $\quad \left\| e^{t\widehat{\triangle}^q} f \right\|_{\mathcal{D}^{\mathrm{sR}}_{k,\,\beta}} \leqslant \mathrm{Cst}(k,\beta)\, \|f\|_{\mathcal{D}^{\mathrm{sR}}_{k,\,\beta}}$

$$\forall\, t \in [0, 1] \quad \forall\, k \in \mathbb{Z} \quad \forall\, \beta \geqslant 1 \quad \forall\, f \in C^\infty_c(\mathbb{R}^n).$$

This implies that $(e^{t\widehat{\triangle}^q})_{t \geqslant 0}$ is a C_0 semigroup in the Hilbert space $\mathcal{D}^{\mathrm{sR}}_{k,\,\beta}$, for every $k \in \mathbb{Z}$ and every $\beta \geqslant 1$, which may not be a semigroup of contractions. Note that, for $k = 0$, $\mathcal{D}^{\mathrm{sR}}_{k,\,\beta}$ is the space $L^2(\mathbb{R}^n)$ with polynomial weight $\langle x \rangle^\beta_{\mathrm{sR}}$ and indeed, for $\beta \geqslant 1$, $\widehat{\triangle}^q$ may not be dissipative in this Hilbert space.

Proof. — Since $(\mathrm{id} - \widehat{\triangle}^q)^j e^{t\widehat{\triangle}^q} = e^{t\widehat{\triangle}^q}(\mathrm{id} - \widehat{\triangle}^q)^j$, it suffices to prove (C.6) for $j = 0$. We proceed in two steps: we first establish (C.6) for $t = 1$, and then we show how to obtain the estimate for every $t \in (0, t_1]$ by homogeneity.

Let $\beta \geqslant 1$ and $k \in \mathbb{N}$ be arbitrary. We follow and adapt a classical argument (Schur test): by the Cauchy–Schwarz inequality, we have

(C.8) $\quad \left(\langle x \rangle^\beta_{\mathrm{sR}} \left(\mathrm{id} - \widehat{\triangle}^q \right)^k e^{\widehat{\triangle}^q} f(x) \right)^2$

$$= \left(\int_{\mathbb{R}^n} \frac{\langle x \rangle^\beta_{\mathrm{sR}}}{\langle y \rangle^\beta_{\mathrm{sR}}} \left(\mathrm{id} - \widehat{\triangle}^q \right)^k_x \widehat{e}^q(1, x, y) \langle y \rangle^\beta_{\mathrm{sR}} f(y)\, dy \right)^2$$

$$\leqslant \int_{\mathbb{R}^n} \left(\frac{\langle x \rangle_{\mathrm{sR}}}{\langle y \rangle_{\mathrm{sR}}} \right)^{2\beta} \left| \left(\mathrm{id} - \widehat{\triangle}^q \right)^k_x \widehat{e}^q(1, x, y) \right| dy \int_{\mathbb{R}^n} \left| \left(\mathrm{id} - \widehat{\triangle}^q \right)^k_x \widehat{e}^q(1, x, y) \right| \langle y \rangle^{2\beta}_{\mathrm{sR}} f(y)^2\, dy$$

for every $f \in C^\infty_c(\mathbb{R}^n)$.

We claim that $\mathrm{Cst}\, \widehat{d}^q(0, x) \leqslant \|x\|_{\mathrm{sR}} \leqslant \mathrm{Cst}\, \widehat{d}^q(0, x)$ for every $x \in \mathbb{R}^n$ (see [Jea14, Lemma 2.1 p. 28]). Indeed, the distance $\widehat{d}^q(0, \cdot)$ is continuous (by the Hörmander condition) and thus reaches its bounds on the compact set $\{x \in \mathbb{R}^n \mid \|x\|_{\mathrm{sR}} = 1\}$; the inequality follows by homogeneity. We infer that

(C.9) $\quad \dfrac{\langle x \rangle_{\mathrm{sR}}}{\langle y \rangle_{\mathrm{sR}}} \leqslant \mathrm{Cst}\, \dfrac{1 + \widehat{d}^q(0, x)}{1 + \widehat{d}^q(0, y)} \leqslant \mathrm{Cst}\, \left(1 + \widehat{d}^q(x, y) \right) \qquad \forall\, x, y \in \mathbb{R}^n.$

It follows from (C.4) and (C.9) that

$$\int_{\mathbb{R}^n} \left(\frac{\langle x\rangle_{\mathrm{sR}}}{\langle y\rangle_{\mathrm{sR}}}\right)^{2\beta} \left|\left(\mathrm{id}-\widehat{\triangle}^q\right)_x^k \widehat{e}^q(1,x,y)\right| dy$$

$$\leqslant \mathrm{Cst}(k) \int_{\mathbb{R}^n} \left(1 + \widehat{d}^q(x,y)\right)^{2\beta} e^{-\widehat{d}^q(x,y)^2/8}\, dy$$

$$\leqslant \mathrm{Cst}(k) \int_{\mathbb{R}^n} \left(1 + \widehat{d}^q(x,y)\right)^{2\beta} e^{-\widehat{d}^q(x,y)^2/8}\, dy$$

$$\leqslant \mathrm{Cst}(k,\beta) \int_{\mathbb{R}^n} e^{-\widehat{d}^q(x,y)^2/16}\, dy$$

$$= \mathrm{Cst}(k,\beta)$$

and hence, plugging into (C.8), integrating with respect to x and using (C.4) again, we obtain

$$(\mathrm{C.10}) \quad \left\|e^{\widehat{\triangle}^q} f\right\|_{\widehat{\mathcal{D}}_{k,\beta}^{q,\,\mathrm{sR}}}^2 = \int_{\mathbb{R}^n} \left(\langle x\rangle_{\mathrm{sR}}^{\beta} \left(\mathrm{id}-\widehat{\triangle}^q\right)^k e^{\widehat{\triangle}^q} f(x)\right)^2 dx$$

$$\leqslant \mathrm{Cst}(k,\beta) \int_{\mathbb{R}^n} e^{-\widehat{d}^q(x,y)^2/8}\, dx \int_{\mathbb{R}^n} \langle y\rangle_{\mathrm{sR}}^{2\beta} f(y)^2\, dy = \mathrm{Cst}(k,\beta) \|f\|_{\widehat{\mathcal{D}}_{0,\beta}^{q,\,\mathrm{sR}}}^2 .$$

We have obtained (C.6) for $t = 1$.

To obtain (C.6) for every $t \in (0,t_1]$, we make the changes of variable $x = \varphi(z)$ and $y = \varphi(z)$ in (C.10), with $\varphi = \delta_{1/\sqrt{t}}$. Noting that $\varphi^*\widehat{\triangle}^q\varphi_* = t\widehat{\triangle}^q$, $\varphi^* e^{\widehat{\triangle}^q}\varphi_* = e^{t\widehat{\triangle}^q}$, $\varphi^* dx = |\det(\varphi)|\, dx = \frac{1}{t^{\mathcal{Q}(q)/2}}\, dx$, and $\|\varphi(x)\|_{\mathrm{sR}} = \frac{1}{\sqrt{t}}\|x\|_{\mathrm{sR}}$, we get from (C.10) that

$$\frac{1}{t^{\mathcal{Q}(q)/2}} \int_{\mathbb{R}^n} \left(\left(1 + \frac{1}{t}\|z\|_{\mathrm{sR}}^2\right)^{2\beta} \left(\mathrm{id}-t\widehat{\triangle}^q\right)^k e^{t\widehat{\triangle}^q}(\varphi^* f)(z)\right)^2 dz$$

$$\leqslant \mathrm{Cst}(k,\beta) \frac{1}{t^{\mathcal{Q}(q)/2}} \int_{\mathbb{R}^n} \left(1 + \frac{1}{t}\|z\|_{\mathrm{sR}}^2\right)^{2\beta} (\varphi^* f)(z)^2\, dz$$

from which (C.6) is easily inferred. $\square$

Using (C.5) and the embeddings (B.6) and (B.7), we infer the following result.

COROLLARY C.4. — *Given any $t > 0$, any $j,k \in \mathbb{Z}$ and any real number $\beta \geqslant 1$, the operator $e^{t\widehat{\triangle}^q}$ is bounded as an operator from $H_{\beta-jN}^{j\sigma}(\mathbb{R}^n)$ when $j \leqslant 0$ (resp., $H_{\beta+2jN}^{2j}(\mathbb{R}^n)$ when $j \geqslant 0$) to $H_{\beta+2kN}^{2k}(\mathbb{R}^n)$ when $k \leqslant 0$ (resp., $H_{\beta-kN}^{k\sigma}(\mathbb{R}^n)$ when $k \geqslant 0$). Moreover, for all $t_0, t_1 \in (0, +\infty)$, their norms are uniformly bounded with respect to $t \in [t_0, t_1]$.*

Remark C.5. — Given any $t > 0$, it is not true that $e^{t\widehat{\triangle}^q}$ maps continuously any $H_\beta^\alpha(\mathbb{R}^n)$ to any $H_{\beta'}^{\alpha'}(\mathbb{R}^n)$. Indeed, otherwise, using Hilbert–Schmidt norms we would infer that its heat kernel $e(t,\cdot,\cdot)$ is bounded in the Schwartz space $\mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$, which is not true around the diagonal.

BIBLIOGRAPHY

[ABB12]   Andrei Agrachev, Davide Barilari, and Ugo Boscain, *On the Hausdorff volume in sub-Riemannian geometry*, Calc. Var. Partial Differ. Equ. **43** (2012), no. 3-4, 355–388. ↑915

[ABB20]   _____, *A Comprehensive Introduction to Sub-Riemannian geometry. From the Hamiltonian viewpoint.*, Cambridge Studies in Advanced Mathematics, vol. 181, Cambridge University Press, 2020, with an appendix by Igor Zelenko. ↑908

[ABGR09]   Andrei Agrachev, Ugo Boscain, Jean-Paul Gauthier, and Francesco Rossi, *The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups*, J. Funct. Anal. **256** (2009), no. 8, 2621–2655. ↑914

[AM05]   Andrei Agrachev and Alessia Marigo, *Rigid Carnot algebras: a classification*, J. Dyn. Control Syst. **11** (2005), no. 4, 449–494. ↑914

[BA89]   Gérard Ben Arous, *Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale*, Ann. Inst. Fourier **39** (1989), no. 1, 73–99. ↑900, 903, 905, 920

[BAL91a]   Gérard Ben Arous and Rémi Léandre, *Décroissance exponentielle du noyau de la chaleur sur la diagonale. I*, Probab. Theory Relat. Fields **90** (1991), no. 2, 175–202. ↑903

[BAL91b]   _____, *Décroissance exponentielle du noyau de la chaleur sur la diagonale. II*, Probab. Theory Relat. Fields **90** (1991), no. 3, 377–402. ↑903, 904

[Bar13]   Davide Barilari, *Trace heat kernel asymptotics in 3D contact sub-Riemannian geometry*, J. Math. Sci., New York **195** (2013), no. 3, 391–411, Translation of Sovrem. Mat. Prilozh. No. 82 (2012). ↑913, 928

[Bel96]   André Bellaïche, *The tangent space in sub-Riemannian geometry*, Sub-Riemannian geometry. Proceedings of the satellite meeting of the 1st European congress of mathematics "Journées nonholonomes : géométrie sous-riemannienne, théorie du contrôle, robotique", Paris, France, June 30–July 1, 1992, Progress in Mathematics, vol. 144, Birkhäuser, 1996, pp. 1–78. ↑908, 909, 911, 913

[Bon69]   Jean-Michel Bony, *Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés*, Ann. Inst. Fourier **19** (1969), no. fasc. 1, 277–304. ↑919

[CdVHT]   Yves Colin de Verdière, Luc Hillairet, and Emmanuel Trélat, *Spectral asymptotics for sub-Riemannian Laplacians*, Work in progress. ↑900, 904, 905, 906

[CdVHT16] _____, *Quantum ergodicity and quantum limits for sub-Riemannian Laplacians*, Sémin. Laurent Schwartz, EDP Appl. **2014-2015** (2016), article no. 20. ↑915

[CdVHT18] _____, *Spectral asymptotics for sub-Riemannian Laplacians, I: Quantum ergodicity and quantum limits in the 3-dimensional contact case*, Duke Math. J. **167** (2018), no. 1, 109–174. ↑905

[CGT82]   Jeff Cheeger, Mikhail Gromov, and Michael Taylor, *Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds*, J. Differ. Geom. **17** (1982), no. 1, 15–53. ↑962, 963, 966

[CS08]   Thierry Coulhon and Adam Sikora, *Gaussian heat kernel upper bounds via the Phragmén–Lindelöf theorem*, Proc. Lond. Math. Soc. **96** (2008), no. 2, 507–544. ↑962, 966

[EH03]   Jean-Pierre Eckmann and Martin Hairer, *Spectral properties of hypoelliptic operators*, Commun. Math. Phys. **235** (2003), no. 2, 233–253. ↑926, 957, 958, 959, 960

[EN00]   Klaus-Jochen Engel and Rainer Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics, vol. 194, Springer, 2000. ↑916, 919, 923, 926, 943

[Fuk87]   Kenji Fukaya, *Collapsing of Riemannian manifolds and eigenvalues of Laplace operator*, Invent. Math. **87** (1987), no. 3, 517–547. ↑923

[Ge93]   Zhong Ge, *Collapsing Riemannian metrics to Carnot–Carathéodory metrics and Laplacians to sub-Laplacians*, Can. J. Math. **45** (1993), no. 3, 537–553. ↑923, 964

[Gro96]   Mikhael Gromov, *Carnot–Carathéodory spaces seen from within*, Sub-Riemannian geometry, Progress in Mathematics, vol. 144, Birkhäuser, 1996, pp. 79–323. ↑908, 910, 911, 964

[HN05]   Bernard Helffer and Francis Nier, *Hypoelliptic estimates and spectral theory for Fokker–Planck operators and Witten Laplacians*, Lecture Notes in Mathematics, vol. 1862, Springer, 2005. ↑953, 955

[Hör58]   Lars Hörmander, *On the division of distributions by polynomials*, Ark. Mat. **3** (1958), 555–568. ↑940, 960

[Hör67]   ———, *Hypoelliptic second order differential equations*, Acta Math. **119** (1967), 147–171. ↑900, 911, 953

[Hsu95]   Elton P. Hsu, *On the principle of not feeling the boundary for diffusion processes*, J. Lond. Math. Soc. **51** (1995), no. 2, 373–382. ↑922

[Hue60]   Denise Huet, *Phénomènes de perturbation singulière dans les problèmes aux limites*, Ann. Inst. Fourier **10** (1960), 61–150. ↑923

[Jea14]   Frédéric Jean, *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*, SpringerBriefs in Mathematics, Springer, 2014. ↑908, 911, 912, 915, 967

[JSC86]   David S. Jerison and Antonio Sánchez-Calle, *Estimates for the heat kernel for a sum of squares of vector fields*, Indiana Univ. Math. J. **35** (1986), no. 4, 835–854. ↑921, 965

[Kan77]   Yakar Kannai, *Off diagonal short time asymptotics for fundamental solutions of diffusion equations*, Commun. Partial Differ. Equations **2** (1977), no. 8, 781–830. ↑963

[Koh73]   Joseph J. Kohn, *Pseudo-differential operators and hypoellipticity*, Partial differential equations. Univ. California, Berkeley, Calif., 1971), Proceedings of Symposia in Applied Mathematics, vol. 23, American Mathematical Society, 1973, pp. 61–69. ↑953, 954, 955

[KS88]   Shigeo Kusuoka and Daniel W. Stroock, *Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator*, Ann. Math. **127** (1988), no. 1, 165–189. ↑965

[LD10]   Enrico Le Donne, *Lecture Notes on sub-Riemannian geometry*, 2010, Ongoing work, preprint. ↑908

[Lio73]   Jacques-Louis Lions, *Perturbations singulières dans les problèmes aux limites et en contrôle optimal*, Lecture Notes in Mathematics, vol. 323, Springer, 1973. ↑923

[Mah98]   Patrick Maheux, *Estimations du noyau de la chaleur sur les espaces homogènes*, J. Geom. Anal. **8** (1998), no. 1, 65–96. ↑966

[Mar07]   Alessia Marigo, *Classification of Carnot algebras: the semi-rigid cases*, J. Dyn. Control Syst. **13** (2007), no. 1, 95–119. ↑914

[Mel85]   Richard Melrose, *The wave equation for a hypoelliptic operator with symplectic characteristics of codimension two*, J. Anal. Math. **44** (1984/85), 134–182. ↑964, 966

[Mét76]   Guy Métivier, *Fonction spectrale et valeurs propres d'une classe d'opérateurs non elliptiques*, Commun. Partial Differ. Equations **1** (1976), no. 5, 467–519. ↑900, 903, 905

[Mon02]   Richard Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, 2002. ↑908, 911, 913

[MS78]   Arthur Menikoff and Johannes Sjöstrand, *On the eigenvalues of a class of hypoelliptic operators*, Math. Ann. **235** (1978), no. 1, 55–85. ↑906

[Paz83]   Amnon Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer, 1983. ↑916, 919, 923

[Rif14]   Ludovic Rifford, *Sub-Riemannian geometry and optimal transport*, SpringerBriefs in Mathematics, Springer, 2014. ↑908, 909

[Rum00]   Michel Rumin, *Sub-Riemannian limit of the differential form spectrum of contact manifolds*, Geom. Funct. Anal. **10** (2000), no. 2, 407–452. ↑923

[SC84]   Antonio Sánchez-Calle, *Fundamental solutions and geometry of the sum of squares of vector fields*, Invent. Math. **78** (1984), no. 1, 143–160. ↑965

[SC92]   Laurent Saloff-Coste, *A note on Poincaré, Sobolev, and Harnack inequalities*, Int. Math. Res. Not. (1992), no. 2, 27–38. ↑966

[Str86]   Robert S. Strichartz, *Sub-Riemannian geometry*, J. Differ. Geom. **24** (1986), no. 2, 221–263. ↑910

[Var90]   Nicolas Th. Varopoulos, *Small time Gaussian estimates of heat diffusion kernels. II. The theory of large deviations*, J. Funct. Anal. **93** (1990), no. 1, 1–33. ↑964, 965, 966

Yves COLIN DE VERDIÈRE
Institut Fourier,
Université Grenoble Alpes,
100 rue des Mathématiques,
38610 Gières, (France)

yves.colin-de-verdiere@univ-grenoble-alpes.fr

Luc HILLAIRET
Institut Denis Poisson,
Université d'Orléans,
route de Chartres,
45067 Orléans Cedex 2, (France)

luc.hillairet@univ-orleans.fr

Emmanuel TRÉLAT
Sorbonne Université, CNRS,
Université de Paris, Inria,
Laboratoire Jacques-Louis Lions (LJLL),
F-75005 Paris, (France)

emmanuel.trelat@sorbonne-universite.fr