Scuola Internazionale Superiore di Studi Avanzati

Area of Mathematics

VOLUME VARIATION AND HEAT KERNEL FOR AFFINE CONTROL PROBLEMS

Ph.D. Thesis

Supervisor
Prof. Andrei A. Agrachev

Candidate
Elisa Paoli

Academic Year 2014/2015

Il presente lavoro costituisce la tesi presentata da Elisa Paoli, sotto la direzione del Prof. Andrei A. Agrachev, al fine di ottenere l'attestato di ricerca post-universitaria *Doctor Philosophiæ* presso la SISSA, Curriculum in Matematica Applicata, Area di Matematica. Ai sensi dell'art. 1, comma 4, dello Statuto della Sissa pubblicato sulla G.U. no. 36 del 13.02.2012, il predetto attestato è equipollente al titolo di *Dottore di Ricerca in Matematica*.

Trieste, Anno Accademico 2014-2015.

Contents

Chapter 1

Introduction

Optimal control problems have attracted an increasing attention in the last decades and have been proved to be wide-ranging enough to cover many classical, but also new fields of mathematics. We are thinking in particular to the powerful approach that they have given to the study of Riemannian and sub-Riemannian geometry ([4, 40, 5]), which can be described through the properties of an optimal control problem, that is linear in the controls and with a quadratic cost. In this thesis we are interested in the contribution that this theory can offer to the study of the heat kernel of hypoelliptic operators and to the analysis of the variation of a volume form under the projection of the Hamiltonian flow.

The results obtained on the fundamental solution of the heat equation, the so called *heat kernel*, on a Riemannian manifold have inspired new interest in the study of the heat kernel for hypoelliptic second order operators. It is worth observing at this point that already from the daily experience, one could guess that there exists a relation between heat and geometry. At a more accurate level, it has indeed been observed a deep interaction between the small time asymptotics of the heat kernel with geometric quantities such as distance [48], cut and conjugate locus [38, 41], and curvature invariants [19].

However, the extension of these results to non-Riemannian situations (from the geometric viewpoint) or to non-elliptic operators (from the viewpoint of PDE) is non trivial: some results have been obtained in the sub-Riemannian context, relating the hypoelliptic heat kernel with its associated Carnot-Carathéodory distance [33, 34] and its cut locus [7, 8], but much less is known concerning the relation with curvature or the generalization to non sub-Riemannian situations.

One of the most celebrated results in the Riemannian setting, which is by now classical, reads as follows. Denote by $p(t, x, y)$ the heat kernel associated with the Laplace-Beltrami operator Δ_g on an n-dimensional Riemannian manifold (M, g). Then the coefficients appearing in the small time heat kernel expansion on the diagonal

$$p(t, x_0, x_0) = \frac{1}{t^{n/2}} \left(\sum_{i=0}^{m} a_i(x_0) t^i + O(t^m) \right) \tag{1.1}$$

contain information about the curvature of the manifold at the point x_0, namely all $a_i(x_0)$ can be written as universal polynomials of the Riemann tensor and its covariant derivatives computed at the point x_0 (see for instance [19, 44]).

On the other hand, let us consider a smooth second order elliptic operator L, possibly with drift, on the n-dimensional Euclidean space. Then one can extract from the principal symbol of the operator L a good Riemannian metric g, such that the associated heat kernel contains geometric information. Namely it is possible to choose g so that the first coefficients of the on-the-diagonal heat kernel expansion depend on the curvature associated to the metric g and the local structure of the drift at the point. For instance, in [21], Bismut proved that for the operator $L = \Delta_g + X_0$ associated to a Riemannian manifold (M, g) the corresponding heat kernel satisfies

$$p(t, x_0, x_0) = \frac{1}{(4\pi t)^{n/2}} \left(1 - \left(\frac{\operatorname{div}(X_0)}{2} + \frac{\|X_0(x_0)\|^2}{2} - \frac{S(x_0)}{6}\right) t + O(t^2)\right), \tag{1.2}$$

for $t \to 0$, where S is the scalar curvature of the Riemannian metric g.

However, as soon as the ellipticity assumption on the operator is removed, even the structure of the asymptotic expansion of the fundamental solution is much less understood, and the drift field plays a central role in the velocity of decay of the asymptotics.

Let us consider for instance the second order Hörmander-type operator on a closed submanifold M of $\mathbb{R}^n$

$$L = \frac{1}{2} \sum_{i=1}^{k} f_i + f_0, \tag{1.3}$$

where $f_0, f_1, \ldots, f_k$ are smooth bounded vector fields, with bounded derivatives of any order. We assume that the vector fields in (1.3) satisfy the *Hörmander condition*

$$\operatorname{Lie}\{(\operatorname{ad} f_0)^j f_i \,|\, i = 1, \ldots, k,\ j \in \mathbb{N}\}\big|_x = T_x M, \qquad \forall\, x \in M. \tag{wHC}$$

Here $(\operatorname{ad} X)Y = [X, Y]$ and $\operatorname{Lie} \mathcal{F}$ denotes the smallest Lie algebra containing a family of vector fields $\mathcal{F}$. As proved by Hörmander [29], this hypothesis implies the hypoellipticity of the operator L and the existence of a smooth fundamental solution $p(t, x, y)$.

A first step in the study of the asymptotic expansion of the heat kernel in the hypoelliptic setting has been done by Ben Arous and Léandre ([33, 34, 13, 15, 14]), and by Barilari, Boscain and Neel ([8]). These results concern hypoelliptic operators without drift field or such that the fields $f_1, \ldots, f_k$ satisfy the *strong Hörmander condition*

$$\operatorname{Lie}\{f_i \,|\, i = 1, \ldots, k\}\big|_x = T_x M, \qquad \forall\, x \in M. \tag{sHC}$$

In this condition, it is not necessary to include f_0 in the bracket generating process. This is the reason why, in contrast, condition (wHC) is also referred to as *weak Hörmander condition.*

Under hypothesis (sHC) it is possible to endow M with a structure of sub-Riemannian manifold. This determines a metric on the distribution $\mathcal{D}$, spanned by the fields $f_1, \ldots, f_k$. Then it is well defined on M the *sub-Riemannian* distance function $d(x, y)$, also called *CC-distance*, determined by the length of curves whose tangent vector lies in $\mathcal{D}$. Using probabilistic techniques, Ben Arous and Léandre were able to relate the small time behavior of the heat kernel to the sub-Riemannian distance function ([33, 34]), generalizing the result by Varadhan [48] in the elliptic case; they also proved that for points $x \neq y$, such that y is not in the cut locus of x, the asymptotics of $p(t, x, y)$ has still a polynomial decay analogous to the elliptic case. This has been then extended in [8] to points x and y that are not conjugate along any minimal geodesic.

However, considering the expansion on the diagonal, they experienced rather new behaviors, that make clear why even the structure of the hypoelliptic heat kernel is not yet well-understood. When $x = y$ the asymptotic expansion depends strongly on the interaction between the drift field and the diffusion generated by the second order term of L. Let

$$\mathcal{D}_x = \text{span}\{f_1, \dots, f_k\}_x, \qquad \text{and} \qquad \mathcal{D}_x^i = \mathcal{D}_x^{i-1} + \text{span}\{[\mathcal{D}, \mathcal{D}^{i-1}]\}_x \qquad \forall i > 1,$$

i.e., $\mathcal{D}_x^i$ is the subspace of T_xM generated by all the Lie brackets of $f_1, \dots, f_k$ up to length i. Ben Arous showed in [15] (see also [14]) that if the drift is a smooth section of $\mathcal{D}^2$, the heat kernel on the diagonal has still a polynomial decay, but of different degree, and precisely

$$p(t, x, x) = \frac{C + O(\sqrt{t})}{t^{\mathcal{Q}/2}},$$

where $\mathcal{Q}$ is the Hausdorff dimension of the manifold and $C > 0$ is a constant depending on x.

On the other hand, if $f_0(x) \notin \mathcal{D}_x^2$, then Ben Arous and Léandre showed in [17, 18] that $p(t, x, x)$ decays to zero exponentially fast, as $\exp\left(-\frac{C}{t^\alpha}\right)$, for a positive α depending on x and bounded above by 1.

A heuristic interpretation of this behavior is that when f_0 points outside $\mathcal{D}^2$, its action is so strong that it cannot be compensated by the diffusion of the fields $f_1, \dots, f_k$ and it moves the concentration of heat far from the initial concentration point x. The break in the behavior is given by the second order Lie brackets. This can be explained since the second order part of the operator reproduces a Brownian motion, which moves as $\sqrt{t}$, while the drift field has velocity t. Then if f_0 points outside $\mathcal{D}^2$, the diffusion generated by the second order part is too slow.

Concerning the generalization of (1.1) about the geometric meaning of the coefficients of the asymptotic expansion, few results are available and again only under the stronger Hörmander condition, in particular when the drift field is either zero or horizontal. In [6] it is computed the first term of the asymptotics for 3D contact structures, where an invariant κ of the sub-Riemannian structure playing the role of the curvature appears. Concerning higher dimensional structures, to our best knowledge, the only known results are [46] for the case of a Sasakian manifold (where the trace of the Tanaka Webster curvature appears) and the case of the two higher dimensional model spaces: CR spheres [11] and quaternionic Hopf fibrations [12].

The research presented in the first part of this thesis wants to be an initial step to understand the behavior of the heat kernel when the operator fails to satisfy the strong Hörmander condition, and only its weak version (wHC) holds.

In Chapter 3, we study the order of the asymptotics at the diagonal. Let x_0 be a point where the drift field lies in $\mathcal{D}_{x_0}^2$. Then we prove that the asymptotic expansion depends on the structure of the Lie algebra generated by the fields and on the controllability of an approximating system. More precisely, we say that the control problem

$$\dot{x} = f_0(x) + \sum_{i=1}^{k} u_i(t) f_i(x) \qquad \text{for } u = (u_1, \dots, u_k) \in L^\infty(\mathbb{R}; \mathbb{R}^k) \tag{1.4}$$

is small time locally controllable at x_0, if for every time t it is possible to reach every point of an open neighborhood of x_0 with curves described by (1.4) starting from x_0 in time not greater than t. Then we show that only the following cases can appear.

If (1.4) is *not* small time locally controllable at x_0, then $p(t, x_0, x_0) = 0$ for every $t > 0$.

If (1.4) is controllable, then we consider an appropriate nilpotent approximation of the fields. This nilpotent approximation is chosen so that it keeps just the necessary information on $f_0, f_1, \dots, f_k$, in order to generate still an hypoelliptic operator. If the associated control problem is small time locally controllable at x_0, then the original fundamental solution $p(t, x_0, x_0)$ has again a polynomial decay (as in the sub-Riemannian setting) and we find the exact order $\mathcal{N}$ of this polynomial decay, generalizing the already known results for the case when (sHC) holds. The number $\mathcal{N}$ is a generalization of the Hausdorff dimension of sub-Riemannian manifolds and depends on the Lie algebra generated by $f_0, f_1, \dots, f_k$.

If instead the approximating system is *not* controllable, then we show that the decay is faster then $t^{-\mathcal{N}}$. Indeed in this case, already Ben Arous and Léandre have shown examples of exponentially fast decay.

In Chapter 4 we perform the first step in the characterization from a geometric viewpoint of the coefficients of the asymptotics on the diagonal. In particular we focus on the model class of linear Hörmander operators in $\mathbb{R}^n$ with constant second order part. These operators are the simplest class of hypoelliptic, but not elliptic, operators satisfying (wHC) and are classical in the literature, starting from the pioneering work of Hörmander [29] (see also [32] for a detailed discussion on this class of operators).

For any point in the kernel of the drift field, we show that $p(t, x, x)$ has a polynomial decay, and we characterize all the coefficients in the asymptotitc expansion through the trace of the drift field and some geometric-like operators defined in [2], and related to the minimal cost of geodesics of the associated optimal control problem. This is a result in the spirit of (1.1).

For points that are not in the kernel of the drift field, we show that the decay depends on the value of the drift. More precisely, if the drift is in the space generated by the constant fields of the second order term, $f_1, \dots, f_k$, then the asymptotics is still polynomial and we find an expansion like (1.2). If the drift points outside this space, then the decay is exponentially fast, even faster than what was found in the sub-Riemannian case. This difference reflects the heuristic opinion that if the drift points outside the space generated by the second order part, then it drifts apart the heat from the initial point, and the fields $f_1, \dots, f_k$ cannot compensate this strong effect.

The generalization of this geometric result to general hypoelliptic operators has been proved to be much more complicated. As an example, in Appendix A we consider a slightly more general operator, the Kolmogorov operator in dimension 2. Here the second order part is constant, while the drift field has no restrictions. Then we compute the first terms of the asymptotic expansion of the fundamental solution at the equilibrium points, x_0, of the drift. We determine also the curvature operator $\mathcal{R}_0$ of an associated geodesic fixed in x_0. By a comparison between the asymptotic expansion and the curvature, we conclude that the first coefficients in the asymptotics do not depend on $\mathcal{R}_0$. For this reason, we suppose that they could be determined by some other invariants depending, in this case, only by the drift field. Even if we have not found explicitly what is the correct geometric interpretation, we hope that this further example can cast a new light on the characterization of the coefficients.

The second part of the thesis concerns the variation of a smooth volume on a manifold, under the projection of the Hamiltonian flow, for a quadratic Hamiltonian. In particular, this class of dynamics contains the sub-Riemannian manifolds.

This study was inspired by the fact that one of the possible ways of introducing curvature in Riemannian geometry is by looking for the variation of a smooth volume under the geodesic flow. Indeed, given a point x on a Riemannian manifold (M, g) and a tangent unit vector $v \in T_xM$, it is well-known that the asymptotic expansion of the Riemannian volume vol_g in the direction of v depends on the Ricci curvature at x. More precisely, the volume element, that is written in coordinates centered in x as $\mathrm{vol}_g = \sqrt{\det g_{ij}} dx_1 \dots dx_n$, satisfies the expansion for $t \to 0$

$$\sqrt{\det g_{ij}(\exp(tv))} = 1 - \frac{1}{6}\mathrm{Ric}(v, v)t^2 + O(t^3),$$

where $\exp(tv)$ denotes the exponential map in the direction v and Ric is the Ricci curvature tensor.

In the sub-Riemannian setting this asymptotic cannot be directly generalized. Indeed, the exponential map is not a diffeomorphism and to compute the volume of small balls one should have a precise knowledge of the structure of the cut locus, which is not easy. Nevertheless the geodesic flow on the Riemannian manifold can be seen as a Hamiltonian flow on the cotangent bundle, associated to a non-degenerate quadratic Hamiltonian. On a sub-Riemannian manifold, and more in general even for structures deriving from an affine control, the Hamiltonian flow is defined in a similar way. In particular, if the structure is sub-Riemannian, the restriction of the Hamiltonian to any fiber is a degenerate non-negative quadratic form. The projection on the manifold, M, of its integral curves are geodesics, but, contrary to the Riemannian case, in general not all the geodesics can be obtained in this way. These projected geodesics are then parametrized by the initial covector in the cotangent bundle and if they are sufficiently regular (ample and equiregular geodesics), it is possible to compute the variation of the volume in a "smooth" way by looking at the measure as an n-form in the cotangent space T^*M, which has dimension $2n$, restricted to the fiber T_x^*M.

To give some insight on this procedure, let us come back to a Riemannian manifold (M, g), endowed with a smooth volume $\mu = e^{\psi}\mathrm{vol}_g$. In the Hamiltonian language, the exponential map on M can be seen as the projection of its Hamiltonian flow in the cotangent bundle. Indeed let $\exp_x(t, v)$ denote the point reached by a curve at time t starting from x with velocity v, i.e., $\exp_x(t, v) = \exp_x(tv)$. The metric g induces a canonical identification between T_xM and the cotangent space T_x^*M. So the exponential map can be seen as a Hamiltonian flow

$$\exp_x(tv) = \exp_x(t, v) = \pi\left(e^{t\vec{H}}\lambda\right),$$

where in the last expression λ denotes the element in T_x^*M corresponding to v. Then the variation of μ is obtained as its pull-back through the map $\pi \circ e^{t\vec{H}} : T^*M \to M$. Observe that the pull-back $(\pi \circ e^{t\vec{H}})^*\mu$ defines an n-form on the cotangent bundle T^*M, that has dimension $2n$. The quantity that we compute is the restriction of this form to the n-dimensional fiber T_x^*M. Moreover, the volume μ_x defines naturally a volume μ_λ on the fiber T_x^*M. With this Hamiltonian interpretation, the classical Riemannian asymptotics can be read as the variation of $(\pi \circ e^{t\vec{H}})^*\mu$ restricted to the fiber T_x^*M, with respect to the volume μ_λ, i.e.,

$$(\pi \circ e^{t\vec{H}})^*\mu\Big|_{T_x^*M} = t^n \, e^{\int_0^t \psi'(\gamma(\tau))d\tau} \left(1 - \frac{1}{6}\mathrm{Ric}_g(v,v)t^2 + O(t^3)\right) \mu_\lambda. \tag{1.5}$$

Figure 5.3 illustrates this variation from the metric measure view point. Indeed let $\Omega \subset T_x^*M$ be a small neighborhood of λ and let $\Omega_{x,t} := \pi \circ e^{t\vec{H}}(\Omega)$ be its image on M with respect to the Hamiltonian flow. For every t it is a neighborhood of $\gamma(t)$. Then

$$\mu(\Omega_{x,t}) = \int_\Omega (\pi \circ e^{t\vec{H}})^*\mu,$$

and (1.5) represents the variation of the volume element along γ.

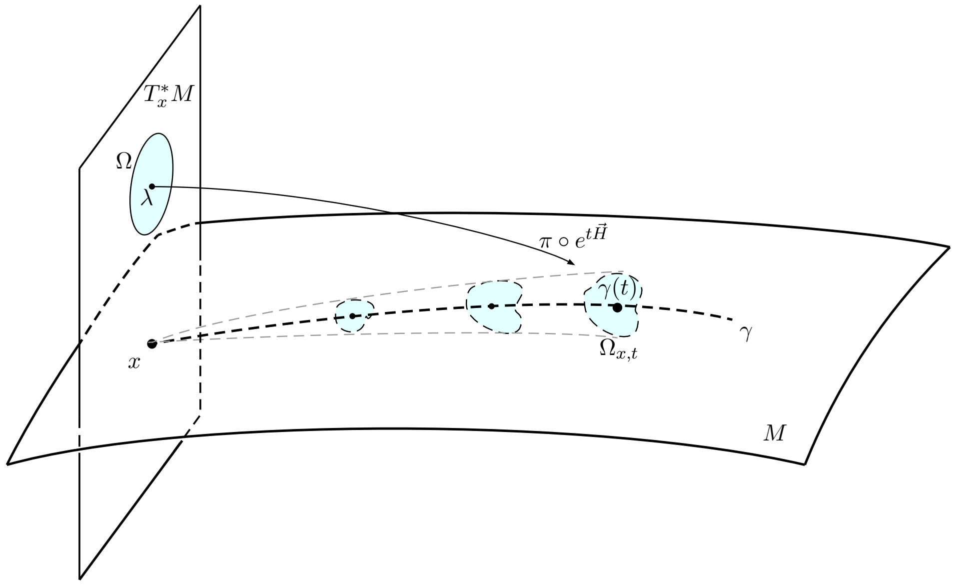


Figure 1.1: Variation of volume

Eq. (1.5) underlines geometric properties of the variation of the volume, as well as its measure properties, separated in distinct parts. Indeed, we see that the order term t^n depends only on the dimension of the manifold. The asymptotics in the brackets contains only geometric information, that depend on the metric g on M. The constant term $e^{\psi(x)}$ depending on μ at the initial point is contained in the associated volume μ_λ. Finally, the measure information is encoded in the exponential term. Indeed it represents the variation of μ along the geodesic and is equal to the exponential of $\int_0^t \psi'(\tau)d\tau = \int_0^t \langle \text{grad}\psi, \dot\gamma(\tau)\rangle d\tau$. In particular, it defines a measure invariant function ρ at every initial cotangent vector λ:

$$\rho(\lambda)\mu_\lambda := \frac{d}{dt}\bigg|_{t=0} \left(t^{-n} \; (\pi \circ e^{t\vec{H}})^* \mu|_\lambda \Big|_{T_x^*M} \right), \qquad \lambda \in T_x^*M. \tag{1.6}$$

In Chapter 5 we generalize the asymptotics (1.5) to a sub-Riemannian structure, and more in general to any structure arising from a non-negative quadratic Hamiltonian. Let M be a smooth manifold and let $\vec{H}$ denote a quadratic, possibly degenerate, Hamiltonian. A special class of dynamics is given by the Hamiltonian, whose restriction to a fiber T_x^*M is a degenerate homogeneous quadratic form (i.e., without linear or constant terms). Then this case recovers the sub-Riemannian structures on the manifold M.

Fix $\lambda \in T_x^*M$ and let $\gamma(t) = \pi(e^{t\vec{H}}\lambda)$ be the associated geodesic on M. The asymptotics that we obtain is expressed as in (1.5) and we interpret every component as a generalization of the corresponding Riemannian element. In particular, the Hamiltonian at λ generates a constant leading term c_0 and influences the order of the asymptotic. Indeed, we observe that the order of the asymptotics is not constant, but depends on the particular geodesic. Moreover, the asymptotics depends on two geometric invariants, that are rational functions in the initial covector λ. The first one is a modification of the Ricci tensor, that is substituted now by the trace of a curvature operator in the direction of λ. This curvature operator, $\mathcal{R}_\lambda$, is a generalization of the sectional curvature and is defined in [3] for the wide class of geometric structures arising from affine control systems. The second invariant is the generalization of $\rho(\lambda)$, introduced in (1.6). It is a measure metric invariant and represents how the volume changes along the curve with respect to a reference n-dimensional form given by the Hamiltonian.

The structure of the thesis is as follows. In Chapter 2 we give a brief introduction on classical results in stochastic differential equations and on the Hörmander theorem, since they play an important role in the understanding of the heat kernel. Chapter 3 is based on [43] and shows the proofs about the order of the heat kernel asymptotic expansion on the diagonal. Chapter 4 contains the results on the model class of the Hörmander operators with linear drift and constant second order term, which can be found in [9]. Chapter 5 concerns the asymptotic expansion of the volume under the Hamiltonian flow, which is the content of [1]. Finally in Appendix A we give a further example of the on-the-diagonal asymptotics of the heat kernel for Kolmogorov hypoelliptic operators in dimension 2.

Aknowledgements. I would like to thank my supervisor Andrei Agrachev for having suggested these problems and for having introduced me to the research. His dedication and enthusiasm for mathematics have fascinated me since the beginning. I'm also grateful to Davide Barilari, that has followed this project with contagious interest and shared with me many parts of this work.

Chapter 2

On stochastic differential equations and Hörmander's theorem

In this chapter we give a brief survey on stochastic differential equations and on the hypoellipticity of operators, proved by Hörmander's theorem. In particular we exploit the relation between the probability density of a stochastic process with the fundamental solution of certain PDEs, which is based on Kolmogorov backward equation. These are classical theorems that we want to recall since their relation is useful for the understanding of the results of the next chapters. We refer to [30] and [42] for a complete presentation of stochastic processes and differential equations and to [29, 35, 27, 28] for what concerns Hörmander's theorem and the existence of the fundamental solution.

In this chapter we give a presentation in the Euclidean setting $\mathbb{R}^n$, but everything is valid also on a closed submanifold M of $\mathbb{R}^n$. Details can be found in [30], Chapter 5.

2.1 Stochastic differential equations

Let $\mathbb{R}^n$ be the n-dimensional Euclidean space and $f_0, f_1, \ldots, f_k$ be smooth vector fields on $\mathbb{R}^n$. In coordinates $(x_1, \ldots, x_n)$, we denote by f_j^i the i-th component of the vector field f_j, i.e., $f_j = \sum_{i=1}^n f_j^i \frac{\partial}{\partial x_i}$.

Consider the time independent stochastic differential equation in Stratonovich form for an n-dimensional stochastic process ξ_t

$$d\xi_t^i = f_0^i(\xi_t) + \sum_{j=1}^k f_j^i(\xi_t) \circ dw_j(t), \tag{2.1}$$

where $w(t) = (w_1(t), \ldots, w_k(t))$ is a Brownian motion. This equation can be written equivalently in Itô differential form, by a modification of the field f_0. Indeed (2.1) is equivalent to

$$d\xi_t^i = \bar{f}_0^i(\xi_t) + \sum_{j=1}^k f_j^i(\xi_t) dw_j(t), \tag{2.2}$$

where

$$\bar{f}_0^i = f_0^i + \frac{1}{2}\sum_{j=1}^{k}\sum_{h=1}^{n} f_j^h \frac{\partial f_j^i}{\partial x_h}.$$

We recall the definition of solution of the stochastic differential equation.

Definition 2.1. *By a solution of equation* (2.2) *we mean an n-dimensional continuous stochastic process ξ_t, $t \geq 0$, defined on a probability space $(\Omega, \mathcal{F}, P)$ such that*

(i) there exists an increasing, right-continuous family $(\mathcal{F}_t)_{t\geq 0}$ of sub-σ-algebras of $\mathcal{F}$;

(ii) there exists a k-dimensional $(\mathcal{F}_t)$-Brownian motion $w(t)$, with $w(0) = 0$ a.s.;

(iii) $(\xi_t)_{t\geq 0}$ is an n-dimensional continuous process adapted to $(\mathcal{F}_t)_{t\geq 0}$, i.e., $\xi_t(\omega)$ is continuous in t for every $\omega \in \Omega$, and it is $\mathcal{F}_t$ measurable for every t;

(iv) the processes $f_j(\xi_t(\omega))$ for $j = 1, \ldots, k$ and $\bar{f}_0(\xi_t(\omega))$ are a.s. $L^2_{loc}([0,T])$ and $L^1_{loc}([0,T])$ integrable respectively, for every $T > 0$;

(v) with probability one, ξ_t and w_t satisfy

$$\xi_t^i - \xi_0^i = \int_0^t \bar{f}_0^i(\xi_s)ds + \sum_{j=1}^{k}\int_0^t f_j^i(\xi_s)dw_j(s) \qquad \forall 1 \leq i \leq n.$$

Solutions to equation (2.2) with Lipschitz continuous fields and starting condition $\xi_0 = x$ are also called *(time-homogeneous) Itô diffusions*. In particular, Itô diffusions satisfy the Markov property, i.e., for any bounded Borel function φ from $\mathbb{R}^n$ to $\mathbb{R}$ and any $t, s \geq 0$ it holds

$$\mathrm{E}[\varphi(\xi_{t+s})|\mathcal{F}_t] = \mathrm{E}[\varphi(\xi_{t+s})|\xi_t],$$

where E denotes the expectation value. In other words, the probability of the process at future steps depends only on the process at the moment ξ_t and not on the previous times.

We recall classical existence and uniqueness theorems for the solution to (2.2), with the sufficient assumptions for this thesis. For less restrictive conditions and the proofs we refer to [30]. In particular, by admitting that solutions can explode, it is possible to prove an existence theorem also for only continuous fields.

Theorem 2.2 (Existence theorem). *If the fields $f_0, f_1, \ldots, f_k$ are bounded with bounded first order derivatives, then for any given probability measure μ on $\mathbb{R}^n$ with compact support, there exists a solution (ξ, w) of* (2.2) *such that the law of ξ_0 coincides with μ, i.e., $P[\xi_0 \in A] = \mu[A]$ for any Borel set $A \in \mathbb{R}^n$.*

Definition 2.3 (Pathwise uniqueness). *We say that the solution of Eq.* (2.2) *is pathwise unique if whenever ξ and ξ' are two solutions defined on the same probability space $(\Omega, \mathcal{F}, P)$ with the same reference family $(\mathcal{F}_t)_{t\geq 0}$ and the same k-dimensional Brownian motion $w(t)$, such that $\xi_0 = \xi'_0$ a.s., then $\xi_t = \xi'_t$ for all $t \geq 0$ a.s.*

Theorem 2.4 (Uniqueness theorem). *Suppose $\bar{f}_0, f_1, \ldots, f_k$ are locally Lipschitz continuous, then equation* (2.2) *has a pathwise unique solution.*

We end this section by recalling Itô's differential formula for stochastic processes. For a proof we refer to [42] Chapter 4.

Theorem 2.5 (Itô formula). *Let ξ_t be a stochastic process solution of* (2.2) *and let $\psi(t,x)$ be a function of class $C^2(\mathbb{R}^+ \times \mathbb{R}^n; \mathbb{R})$ (by $\mathbb{R}^+$ we mean the strictly positive real line). Then it holds*

$$d\psi(t,\xi_t) = \frac{\partial \psi(t,\xi_t)}{\partial t} dt + \sum_{i=1}^{n} \frac{\partial \psi(t,\xi_t)}{\partial x_i} d\xi_t^i + \frac{1}{2} \sum_{i,h=1}^{n} \sum_{j=1}^{k} f_j^i(\xi_t) f_j^h(\xi_t) \frac{\partial^2 \psi(t,\xi_t)}{\partial x_i \partial x_h} dt.$$

2.2 Fundamental solution and Kolmogorov backward equation

In this section we recall the relation between second order operators and solutions to equation (2.1). Let us consider the operator

$$\frac{\partial \varphi}{\partial t} - f_0(\varphi) - \frac{1}{2} \sum_{i=1}^{k} f_i^2(\varphi) \qquad \forall \varphi \in C^\infty(\mathbb{R} \times \mathbb{R}^n) \tag{2.3}$$

and denote by L the operator $f_0 + \frac{1}{2}\sum_{i=1}^k f_i^2$.

Definition 2.6. *The* fundamental solution *of an operator $\frac{\partial}{\partial t} - L$ over $\mathbb{R} \times \mathbb{R}^n$ is a function $p(t,x,y) \in C^\infty(\mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^n)$ such that*

- *for every fixed $y \in \mathbb{R}^n$, it holds $\frac{\partial}{\partial t} p(t,x,y) = L_x p(t,x,y)$, where the operator L acts on the x variable;*
- *for any $\varphi_0 \in L^2(\mathbb{R}^n)$, we have*

$$\lim_{t \searrow 0} \int_{\mathbb{R}^n} p(t,x,y) \varphi_0(y) dy = \varphi_0(x).$$

In other words, if we want to solve the partial differential equation $\frac{\partial \varphi}{\partial t} = L\varphi$ with initial condition $\varphi(0,x) = \varphi_0(x)$, the fundamental solution allows to reconstruct φ by convolution of φ_0 with $p(t,x,y)$.

The Kolmogorov backward equation relates the fundamental solution of (2.3) to the stochastic process solution of (2.2) (see also Corollary 2.8 below).

Theorem 2.7 (Kolmogorov backward equation). *Let $\varphi \in C_c^\infty(\mathbb{R}^n)$ and ξ_t be the stochastic process solution of* (2.2)*. Then the function*

$$u(t,x) := \mathrm{E}[\varphi(\xi_T) | \xi_t = x] \qquad T > t \geq 0, x \in \mathbb{R}^n$$

satisfies

$$\begin{cases} \frac{\partial u(t,x)}{\partial t} + Lu(t,x) = 0 \\ \lim_{t \nearrow T} u(t,x) = \varphi(x). \end{cases} \tag{2.4}$$

Proof. Let us fix the function $\varphi \in C_c^\infty(\mathbb{R}^n)$. The limit condition is easily true.

For the differential equation, notice that the function $u(t,x)$ has the following properties: $\mathrm{E}[u(t,\xi_t)|\xi_t = x] = u(t,x)$ and $\mathrm{E}[u(T,\xi_T)|\xi_t = x] = u(t,x)$. Then by Itô formula and since the Brownian motion has zero expectation value, we have the following series of identities

$$\begin{aligned} 0 =& \mathrm{E}[u(T,\xi_T) - u(t,\xi_t)|\xi_t = x] = \mathrm{E}\left[\int_t^T du(s,\xi_s)\bigg| \xi_t = x\right] \\ =& \mathrm{E}\left[\int_t^T \frac{\partial u(s,\xi_s)}{\partial s} + \sum_{i=1}^n \bar{f}_0^i(\xi_s)\frac{\partial u(s,\xi_s)}{\partial x_i}\right. \\ &\left. + \frac{1}{2}\sum_{i,h=1}^n \sum_{j=1}^k f_j^i(\xi_s) f_j^h(\xi_s)\frac{\partial^2 u(s,\xi_s)}{\partial x_i \partial x_h} ds \bigg| \xi_t = x\right] \\ =& \mathrm{E}\left[\int_t^T \frac{\partial u(s,\xi_s)}{\partial t} + Lu(s,\xi_s)\bigg| \xi_t = x\right]. \end{aligned}$$

Since this equation holds for every $t < T$ and $x \in \mathbb{R}^n$, also the punctual equation (2.4) is satisfied. □

Let us assume for a moment that the process ξ_t admits a smooth density function, that means that there exists a smooth function $p(T,y;t,x)$ for $0 < t \leq T$ and $x, y \in \mathbb{R}^n$ such that for all Borel sets $A \subset \mathbb{R}^n$

$$P[\xi_T \in A|\xi_t = x] = \int_A p(T,y;t,x)dy,$$

where P denotes the probability. Then the function u defined in Theorem 2.7 is

$$u(t,x) = \int_{\mathbb{R}^n} \varphi(T,y)p(T,y;t,x)dy.$$

Since the theorem holds for any function φ with compact support, the probability density itself satisfies (2.4). More precisely

$$\begin{cases} \frac{\partial p(T,y;t,x)}{\partial t} + L_x p(T,y;t,x) = 0 \\ \lim_{t\nearrow T} p(T,y;t,x) = \delta_y(x), \end{cases}$$

where δ_y is the Dirac delta centered in y and L_x denotes that the operator is acting on the x variable. Let $p(t,x,y)$ denote the density of ξ_t with initial condition $\xi_0 = x$. By the Markov property of ξ_t we know that $p(T,y;t,x) = p(T-t,x,y)$. With a change of the time variable we have then proved the following corollary.

Corollary 2.8. *Let ξ_t be the solution of* (2.1)*, with initial condition $\xi_0 = x$ and assume that it admits a smooth density function*

$$p(t,x,y)dy := P[\xi_t \in dy|\xi_0 = x].$$

Then $p(t,x,y)$ is the fundamental solution of the operator in (2.3)*, i.e.,*

$$\begin{cases} \frac{\partial p(t,x,y)}{\partial t} = L_x p(t,x,y) \\ \lim_{t\searrow 0} p(t,x,y) = \delta_y(x). \end{cases}$$

2.2.1 Fundamental solution of the adjoint operator

Let $p(t,x,y)$ be the fundamental solution of the operator $\frac{\partial}{\partial t} - L$ and define the heat operator

$$e^{tL}\varphi(x) = \int_{\mathbb{R}^n} p(t,x,y)\varphi(y)dy \qquad \forall \varphi \in L^2(\mathbb{R}^n).$$

Let L^* be the adjoint operator to L.

Lemma 2.9. *The fundamental solution, $p^*(t,x,y)$, associated to the adjoint operator L^* can be obtain from the fundamental solution of L as*

$$p^*(t,x,y) = p(t,y,x). \tag{2.5}$$

Proof. Let $\varphi, \psi \in L^2(\mathbb{R}^n)$ then by definition of adjoint operator we have

$$\begin{aligned} \int_{\mathbb{R}^n} \varphi(y) e^{tL^*}\psi(y)dy &= \int_{\mathbb{R}^n} \psi(x)e^{tL}\varphi(x)dx \\ &= \int_{\mathbb{R}^n} \psi(x) \int_{\mathbb{R}^n} p(t,x,y)\varphi(y)dydx \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \psi(x)p(t,x,y)dx \right) \varphi(y)dy. \end{aligned}$$

Since this identity holds for any $\varphi, \psi \in L^2(\mathbb{R}^n)$, Eq. (2.5) follows. □

2.3 Hörmander Theorem

In the previous section we have seen how to find the fundamental solution for the operator (2.3), provided that (2.1) admits a smooth probability density. This condition is satisfied if the operator is elliptic, namely if the vector fields $f_1, \dots, f_k$ span $\mathbb{R}^n$. However many interesting operators do not satisfy this property and one would like to have a criterion to admit the existence of a fundamental solution also for only hypoelliptic operators, i.e., for operators that have the following smoothing property.

Definition 2.10. *The operator $(\partial_t - L)$ is said to be* hypoelliptic *if for every function u and every open set $A \in \mathbb{R} \times \mathbb{R}^n$ such that $(\partial_t - L)u \in C^\infty(A)$, then $u \in C^\infty(A)$.*

Here we use the shorthand ∂_t to denote the partial derivative in the time variable.

We say that the operator (2.3) satisfies *Hörmander condition* if the Lie algebra generated by the vector fields $\partial_t - f_0, f_1, \dots, f_k$ is equal to $\mathbb{R} \times \mathbb{R}^n$ at every point, i.e., if

$$\mathrm{Lie}_{(t,x)}\{\partial_t - f_0, f_1, \dots, f_k\} = \mathbb{R} \times \mathbb{R}^n \qquad \forall x \in \mathbb{R}^n. \tag{2.6}$$

By the Lie algebra spanned by the vector fields $\{X_i\}_{i=0}^k$ we mean the space generated by all the vector fields obtained with the Lie brackets of any order of $X_0, \dots, X_k$. Since the vector fields ∂_t and f_i are completely independent condition (2.6) is equivalent to require that

$$\mathrm{span}_x\{f_i, [f_i, f_j], [f_0, f_i], [f_i, [f_j, f_h]], \dots : 1 \le i, j, h \le k\} = \mathbb{R}^n \qquad \forall x \in \mathbb{R}^n,$$

where in the generating set we take the fields $f_1, \dots, f_k$ (not f_0) and all the Lie brackets of any order of all the vector fields $f_0, \dots, f_k$. In other words, f_0 is not in the set, it is there only inside a Lie bracket with other vector fields.

Hörmander proved in [29] that this hypothesis is sufficient to imply that the operator (2.3) is hypoelliptic. Moreover if we assume also the boundedness of the vector fields $f_0, \dots, f_k$ and of their derivatives of any order, it is possible to prove that (2.3) admits a fundamental solution. This conclusion was first proved by Hörmander in [29] with analytical tecnics and subsequently by Malliavin [35] with a probabilistic approach, using Malliavin calculus. Here we refer to Hairer's version in [28] and [27].

Theorem 2.11. *Assume that all vector fields in* (2.1) *are bounded with bounded derivatives of any order. If moreover they satisfy Hörmander condition* (2.6)*, then the solution of* (2.1) *admits a smooth density with respect to Lebesgue measure.*

Proof. The proof of this deep theorem can be found in [28] and [27], where the Malliavin calculus is used.

Here we want to stress that under the hypothesis of the theorem, there exists a fundamental solution of (2.3) and it is determined in Corollary 2.8. □

Remark 2.12. *The hypothesis on the boundedness of the fields in Theorem 2.11 can be partially weaken. For the interested reader we refer to [28] Section 4. Here we write the main ideas.*

Let $\Phi_t(x)$ be the solution map of (2.1) *with initial condition x. For a given initial condition x_0 we denote by $J_{0,t}$ the derivative of Φ_t evaluated at x_0. Moreover let $\xi_t = \Phi_t(x_0)$. A differentiation of* (2.1)*, yields that $J_{0,t}$ is a solution of*

$$dJ_{0,t} = Df_0(\xi_t)\, J_{0,t}\, dt + \sum_{i=1}^{k} Df_i(\xi_t)\, J_{0,t} \circ dw_i(t), \qquad J_{0,0} = \mathrm{Id}_n$$

where Id_n is the $n \times n$ identity matrix. Higher order derivatives $J_{0,t}^{(k)}$ with respect to the initial condition can be defined in a similar way. Moreover let $J_{0,t}^{-1}$ be the inverse of the Jacobian matrix, which also solves an analogous differential equation.

Then in Theorem 2.11 it is enough to require that the fields $f_0, f_1, \dots, f_k$ are C^∞, with all their derivatives that grow at most polynomially at infinity. Furthermore, we assume that they are such that the processes ξ_t, $J_{0,t}^{(k)}$ and $J_{0,t}^{-1}$ satisfy

$$\mathrm{E}[\sup_{t \le T} |\xi_t|^p] < \infty, \qquad \mathrm{E}[\sup_{t \le T} |J_{0,t}^{(k)}|^p] < \infty \qquad \mathrm{E}[\sup_{t \le T} |J_{0,t}^{-1}|^p] < \infty$$

for every initial condition $x_0 \in \mathbb{R}^n$, every terminal time $T > 0$, every $k > 0$ and every $p > 0$.

Notice that the boundedness of the vector fields and of all their derivatives, in particular implies this weaker hypothesis.

2.4 Example: linear operators

The simplest example of hypoelliptic operator is the linear partial differential operator with constant second order coefficients and affine drift field, that satisfy Hörmander condition.

Since this is the subject of chapter 4, for the reader's convenience we derive here explicitly its fundamental solution.

Let us consider the partial differential equation

$$\frac{\partial \varphi}{\partial t} - \sum_{j=1}^{n} (\alpha + Ax)_j \frac{\partial \varphi}{\partial x_j} - \frac{1}{2} \sum_{j,h=1}^{n} (BB^*)_{jh} \frac{\partial^2 \varphi}{\partial x_j \partial x_h} = 0, \qquad \forall \varphi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^n), \tag{2.7}$$

where α is a constant column vector of dimension n, A is an $n \times n$ real matrix, which represents the linear part of a *drift* field, and B is an $n \times k$ real matrix, with $1 \leq k \leq n$, that generates the diffusion coefficients.

We will further assume Hörmander condition (2.6), which in this context becomes a condition on the matrices A and B, called *Kalman condition*, namely

$$\mathrm{rk}[B, AB, A^2B, \ldots, A^{n-1}B] = n. \tag{2.8}$$

We have seen in Corollary 2.8 that the fundamental solution, $p(t, x, y)$, of (2.7) is given by the probability density of the solution ξ_t of the associated stochastic differential equation

$$\begin{cases} d\xi_t = (\alpha + A\xi_t)dt + Bdw(t) \\ \xi_0 = x \end{cases} \tag{2.9}$$

where $w(t) = (w_1(t), \ldots, w_k(t))$ is a k-dimensional Brownian motion. (Notice that since B is constant, the Itô and the Stratonovich equations coincide.) In other words, $p(t, x, y)$ is the $C^\infty(\mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^n)$ function such that for every Borel set $U \in \mathbb{R}^n$ the probability of ξ_t to be in U at time t is given by

$$P\left[\xi_t \in U | \xi_0 = x\right] = \int_U p(t, x, y) dy.$$

By $\mathbb{R}^+$ we denote, here and in the following, the strictly positive real line. The stochastic process solution of (2.9) is equal to

$$\xi_t = e^{tA}\xi_0 + e^{tA} \int_0^t e^{-sA} ds\alpha + e^{tA} \int_0^t e^{-sA} Bdw(s),$$

as it can be readily verified with a derivation. In particular, if the initial value ξ_0 is Gaussian distributed, then ξ_t is a Gaussian process, since by definition of solution ξ_0 and $w(t)$ are independent. The initial condition is a Dirac delta centered in x, so it is a degenerate Gaussian with mean x and vanishing covariance matrix, therefore to find the distribution of ξ_t it is enough to determine its mean value and covariance matrix. To this end we use Itô's differential formula of Theorem 2.5 applied to some auxiliary functions ψ.

To find the mean value of ξ_t let us fix $\psi := x_j$ for $1 \leq j \leq n$, then

$$\xi_j(t) - \xi_j(0) = \int_0^t (\alpha + A\xi(\tau))_j d\tau + \int_0^t \sum_{i=1}^{k} B_{ji} dw_i(\tau).$$

Let $m_j(t)$ be the j-th component of the mean value of ξ_t. By the previous formula, since the integral of the Brownian motion has zero mean value, we find that $m(t)$ satisfies the differential equation $\dot{m}_j(t) = (\alpha + Am(t))_j$ with initial condition $m(0) = x$. Therefore

$$\mathrm{E}[\,\xi_t \,|\, \xi_0\,] = e^{tA}\left(x + \int_0^t e^{-sA} ds \alpha\right). \tag{2.10}$$

To compute the covariance matrix let us choose $\psi = (x_j - m_j(t))(x_h - m_h(t))$, then by Itô differential formula it holds

$$\begin{aligned}\psi(t,\xi(t)) - \psi(0,\xi(0)) &= \int_0^t -\dot{m}_j(\tau)(\xi_h(\tau) - m_h(\tau)) - \dot{m}_h(\tau)(\xi_j(\tau) - m_j(\tau)) d\tau \\ &+ \int_0^t (\alpha + A\xi)_j(\xi_h(\tau) - m_h(\tau)) + (\alpha + A\xi)_h(\xi_j(\tau) - m_j(\tau)) + (BB^*)_{jh} d\tau \\ &+ \int_0^t \sum_{i=1}^k B_{ji}(\xi_h(\tau) - m_h(\tau)) + B_{hi}(\xi_j(\tau) - m_j(\tau)) dw_i(\tau).\end{aligned}$$

Let us take the expectation value and denote by $\rho_{jh}(t)$ the jh-component of the covariance matrix. Recall that $\mathrm{E}[(\xi_j(t) - m_j(t))|\xi(0)] = 0$. To evaluate

$$\mathrm{E}\left[(A\xi(\tau))_j(\xi_h - m_h(\tau))|\xi(0)\right]$$

we rewrite it as

$$\mathrm{E}\left[(A(\xi(\tau) - m(\tau)))_j(\xi_h - m_h(\tau)) + (Am(\tau))_j(\xi_h - m_h(\tau))|\xi(0)\right] = (A\rho(\tau))_{jh} + 0.$$

Then $\rho(t)$ satisfies the differential equation $\dot{\rho}_{jh}(t) = (A\rho(t))_{jh} + (A\rho(t))_{hj} + (BB^*)_{jh}$ with vanishing initial value, whose solution is

$$D_t := \mathrm{E}[(\xi_j - m_j(t))(\xi_h - m_h(t))|\xi(0)] = e^{tA}\int_0^t e^{-\tau A} BB^* e^{-\tau A^*} d\tau e^{tA^*}.$$

By Kalman's condition (2.8), the matrix D_t is invertible for every $t > 0$. Therefore we can conclude that the C^∞ fundamental solution of equation (2.7) is given by the non-degenerate Gaussian

$$p(t,x,y) = \frac{e^{\varphi(t,x,y)}}{(2\pi)^{n/2}\sqrt{\det D_t}},$$

where

$$\varphi(t,x,y) = -\frac{1}{2}\left(y - e^{tA}\left(x + \int_0^t e^{-sA} ds\alpha\right)\right)^* D_t^{-1}\left(y - e^{tA}\left(x + \int_0^t e^{-sA} ds\alpha\right)\right)$$

and it is determined by the mean value (2.10) and the covariance matrix D_t. In the case when $\alpha = 0$ the formula for $p(t,x,y)$ reduces to

$$p(t,x,y) = \frac{e^{-\frac{1}{2}(y - e^{tA}x)^* D_t^{-1}(y - e^{tA}x)}}{(2\pi)^{n/2}\sqrt{\det D_t}} \qquad \text{for } t > 0,\ x, y \in \mathbb{R}^n. \tag{2.11}$$

Chapter 3

Order of the asymptotic expansion of the heat kernel on the diagonal

This chapter is based on the results of [43] and contains the proofs about the order of the heat kernel asymptotic expansion on the diagonal.

3.1 Overview of the Chapter

Let M be a closed n-dimensional submanifold of the Euclidean space and let μ be a volume form on M. Given $f_0, f_1, \ldots, f_k$ smooth vector fields on M we consider the following partial differential operator:

$$\frac{\partial \varphi}{\partial t} - f_0(\varphi) - \frac{1}{2}\sum_{i=1}^{k} f_i^2(\varphi) \qquad \forall \varphi \in C^\infty(\mathbb{R} \times M). \tag{3.1}$$

We assume that the fields $f_0, f_1, \ldots, f_k$ are bounded with bounded derivatives of any order and that they satisfy the *Hörmander condition*

$$\mathrm{Lie}_{(t,x)}\left\{\frac{\partial}{\partial t} - f_0, f_1, f_k\right\} = \mathbb{R} \times T_x M \qquad \forall x \in M, t > 0, \tag{3.2}$$

where Lie denotes the Lie algebra generated by the fields. As explained in Chapter 2, under these conditions the operator is hypoelliptic and admits a fundamental solution, $p(t, x, y)$.

As soon as the ellipticity assumptions on the operator is removed, even the structure of the asymptotic expansion of the fundamental solution is not well understood, and the drift field plays a central role in the velocity of decay of the asymptotics. Already the order of the small time asymptotic expansion of p on the diagonal is not completely known and some results exist only under the assumption of the strong Hörmander condition (sHC). Let

$$\mathcal{D}_x = \mathrm{span}\{f_1, \ldots, f_k\}_x, \qquad \text{and} \qquad \mathcal{D}_x^i = \mathcal{D}_x^{i-1} + \mathrm{span}\{[\mathcal{D}, \mathcal{D}^{i-1}]\}_x \qquad \forall i > 1,$$

i.e., $\mathcal{D}_x^i$ is the subspace of $T_x M$ generated by all the Lie brackets of $f_1, \ldots, f_k$ up to length i. Ben Arous showed in [15] (see also [14]) that if the drift is a smooth section of $\mathcal{D}^2$, the heat kernel on the diagonal has a polynomial decay, but the degree in different from the elliptic case, and precisely

$$p(t,x,x) = \frac{C + O(\sqrt{t})}{t^{\mathcal{Q}/2}}, \tag{3.3}$$

where $\mathcal{Q}$ is the Hausdorff dimension of the manifold and $C > 0$ is a constant depending on x.

Conversely, if $f_0(x) \notin \mathcal{D}^2$, then Ben Arous and Léandre showed in [17, 18] that $p(t,x,x)$ decays to zero exponentially fast, as $\exp\left(-\frac{C}{t^\alpha}\right)$, for a positive α depending on x and bounded above by 1.

In this chapter we study the order of the asymptotics at the diagonal, when only the weak Hörmander condition (3.2) holds. Our results apply to any point, x_0, where the *drift field* f_0 lies in $\mathcal{D}^2_{x_0}$. We underline that this is a property at the point x_0, it is not necessary that the drift is a section of $\mathcal{D}^2$.

However, if f_0 is even a section of $\mathcal{D}^2$, then the weak Hörmander condition implies also the strong one, and indeed we recover the small time asymptotics presented in (3.3), which depends on the Hausdorff dimension $\mathcal{Q}$ of the sub-Riemannian manifold.

In the general case, when $f_0(x_0) \in \mathcal{D}^2_{x_0}$ only at the point x_0, then the decay can be either polynomially fast, as in the sub-Riemannian case, or exponential, depending on the principal part of the fields $f_0, f_1 \dots, f_k$, that we will soon introduce. Moreover, in the polynomial case we show the exact order of decay. This is given by a number $\mathcal{N}$ that depends on the Lie algebra generated by the fields. This number generalizes the Hausdorff dimension, which in this case could actually even not be defined, since the strong Hörmander condition is not guaranteed. However, if $\mathcal{Q}$ exists, then we show that $\mathcal{N} \le \mathcal{Q}$, that means that the drift field produces a slower decay of the heat at its equilibrium points.

To present the details of these results, let us introduce some notation. The proof relies on a homogeneity property of the operator in (3.1) under dilation and it is suggested by the following observation. The fundamental solution, $p(t,x,y)$, is characterized as the probability density of the stochastic process ξ_t, starting from the point x, and solution of the stochastic differential equation written in Stratonovich form

$$d\xi_t = f_0(\xi_t)dt + \sum_{i=1}^{k} f_i(\xi_t) \circ dw_i(t), \tag{3.4}$$

where $w = (w_1, \dots, w_k)$ denotes a k-dimensional Brownian motion. Heuristically the flow of the drift field has order t, while a Brownian motion moves as $\sqrt{t}$. Therefore the idea is to introduce weights of the fields and to assign to f_0 a double weight with respect to the other fields. Accordingly, we then define a new filtration of the tangent space at x_0, $\mathcal{G} = \{G_i(x_0)\}$, which involves f_0 as follows: every layer G_i is spanned by the Lie brackets of $f_0, \dots, f_k$ up to length i, where f_0 is counted twice. For example the first 4 layers are

$$\begin{aligned} G_0 &= \{0\}, \\ G_1 &= \operatorname{span}\{f_1, \dots, f_k\}, \\ G_2 &= \operatorname{span}\{f_i, [f_i, f_j], f_0 : i,j = 1, \dots, k\}, \\ G_3 &= \operatorname{span}\{f_i, [f_i, f_j], f_0, [f_i, [f_j, f_h]], [f_0, f_i] : i,j,h = 1, \dots, k\}. \end{aligned} \tag{3.5}$$

Let $k_i := \dim G_i(x_0)$. By condition (wHC) there exists a smallest integer m such that $G_m(x_0) = T_{x_0}M$. The integer $\mathcal{N}$ that determines the order of the polynomial decay is defined as

$$\mathcal{N} := \sum_{i=1}^{m} i \cdot (k_i - k_{i-1}) = \sum_{i=1}^{m} i(\dim G_i(x_0) - \dim G_{i-1}(x_0)),$$

i.e. it is the dimension of the manifold, where each coordinate in the $i-th$ layer, that is not in the $(i-1)-th$ layer, is counted i times. Notice that, if the drift field is identically zero, we recover the sub-Riemannian case, where every layer G_i of the filtration coincides with $\mathcal{D}^i$ and $\mathcal{N} = \mathcal{Q}$.

Analogously to the procedure usually done in the sub-Rimannian contest, we take nilpotent approximations of the fields $f_0, f_1, \ldots, f_k$ with respect to the filtration $\mathcal{G}_{x_0}$. Namely, let $(x_1, \ldots, x_n)$ be coordinates, in a neighborhood of x_0, adapted to the filtration (see Definition 3.9) and for $\epsilon > 0$ define a dilation, which multiplies every coordinate of the i-th layer of a factor ϵ^i, for $1 \leq i \leq m$. Then let $\hat{f}_1, \ldots, \hat{f}_k$ be the principal parts of the fields f_i homogeneous of order 1, with respect to the dilations, and let $\hat{f}_0$ be the principal part of the drift field, homogeneous of order 2.

This construction produces a split in the operator L as follows: let L_0 be the operator defined by the principal parts of the fields

$$L_0 := \hat{f}_0 + \frac{1}{2} \sum_{i=1}^{k} \hat{f}_i^2. \tag{3.6}$$

Then the operator L can be seen as a sum of two parts, L_0 and $L - L_0$. The careful choice of the filtration $\mathcal{G}$, guaranties that the principal operator L_0 preserves the weak Hörmander condition in the equilibrium points of the drift field, therefore it is still hypoelliptic and it admits a smooth fundamental solution $q_0(t, x, y)$ on $\mathbb{R}^n$. Moreover, q_0 behaves well under a rescaling, indeed for every $t > 0$ we have the homogeneity property

$$q_0(t, x, y) = \frac{1}{t^{\mathcal{N}/2}} q_0(1, \delta_{1/\sqrt{t}} x, \delta_{1/\sqrt{t}} y) \qquad \forall x, y \in \mathbb{R}^n.$$

Finally we can state our main results. We prove that the asymptotics of the fundamental solution depends on the controllability of some control problems, associated to the stochastic equation (3.4). Consider the control problem

$$\dot{x}(t) = f_0(x(t)) + \sum_{i=1}^{k} u_i(t) f_i(x(t)) \tag{3.7}$$

where $u = (u_1, \ldots, u_k) \in L^\infty(\mathbb{R}; \mathbb{R}^k)$ are the controls. If the control problem (3.7) is not small time locally controllable around the point x_0, i.e. from x_0 we can not reach a neighborhood of x_0 using curves described by the control problem, then

$$p(t, x_0, x_0) = 0 \qquad \forall t > 0.$$

We consider also the control problem induced by the approximating system

$$\dot{x}(t) = \hat{f}_0(x(t)) + \sum_{i=1}^{k} u_i(t) \hat{f}_i(x(t)). \tag{3.8}$$

If this control problem is small time locally controllable, then so is also the original one (3.7) and we can prove that the small time asymptotics of the fundamental solution is polynomial, precisely

$$p(t, x_0, x_0) = \frac{q_0(1, x_0, x_0) + O(t)}{t^{\mathcal{N}/2}}. \tag{3.9}$$

In particular, if the operator arises from a sub-Rimemannian manifold with no drift field, the approximating control problem is always controllable. Moreover, the integer $\mathcal{N}$ is equal to the homogeneous dimension $\mathcal{Q}$ of the manifold and we recover Ben Arous result (3.3). On the other hand, if the drift field is not identically zero, but still the fields $f_1, \ldots, f_k$ satisfy the strong Hörmander condition, then we prove the inequality $\mathcal{N} \leq \mathcal{Q}$.

In the intermediate case in which the approximating control problem (3.8) is not small time locally controllable in x_0, but the original control problem (3.7) is still small time locally controllable, then the behavior of the asymptotics can be more general. It blows up faster than $t^{-\mathcal{N}/2}$, and possibly even exponentially fast, as it was already pointed out in some examples by Ben Arous and Léandre in [18].

These conclusions are obtained by a careful generalization of Stroock and Varadhan's support theorem for diffusion operators, that allows to characterize when the leading term $q_0(1, x_0, x_0)$ in (3.9) is strictly positive.

The structure of the chapter is as follows. We begin by describing in details in Section 3.2 the homogeneity properties of the operator (3.1). In particular we derive the conditions that the dilations have to satisfy in order to produce the right split of the operator, into a hypoelliptic principal part plus a perturbation. We also give a brief introduction into Duhamel's formula in Subsection 3.2.1, since it is an important tool to study the perturbed operator. In Section 3.3 we introduce the coordinates that give the right dilations of the space. These coordinates are defined from a filtration of the tangent space to x_0 determined by the fields $f_0, f_1, \ldots, f_k$ and give rise to a graded structure around x_0, which defines an anisotropic dilation. In Section 3.4 we define the nilpotent approximation, that determines the principal operator (3.6). We compute also the integer $\mathcal{N}$ appearing in the asymptotics (3.9), that comes from the change of the volume form under the dilations. In Section 3.5 we prove the asymptotics (3.9), by using the tools introduced in the previous sections. In the following Section 3.6 we focus our study on the operator derived from the nilpotent approximating system and its associated control problem. By proving a modification of Stroock and Varadhan's support theorem, we give a necessary and sufficient condition for the positivity of the fundamental solution of the principal operator, that is based on the controllability of the approximating control system. Finally we end the chapter with Section 3.7, where we show a series of examples, illustrating how this formula recovers in particular the known results recalled in the introduction.

3.2 The fundamental solution and its behavior under the action of a dilation

Let us consider the operator in (3.1). We will call f_0 the *drift* field and we will denote by L the operator $f_0 + \frac{1}{2}\sum_{i=1}^{k} f_i^2$. Let us recall the definition of fundamental solution.

Definition 3.1. *The* fundamental solution *of an operator* $\frac{\partial}{\partial t} - L$ *over* $\mathbb{R} \times M$ *with respect to the volume* μ *is a function* $p(t, x, y) \in C^\infty(\mathbb{R}^+ \times M \times M)$ *such that*

- *for every fixed $y \in M$, it holds $\frac{\partial}{\partial t}p(t,x,y) = L_x p(t,x,y)$, where the operator L acts on the x variable;*
- *for any $\varphi_0 \in L^2(M)$, we have*

$$\lim_{t \searrow 0} \int_M p(t,x,y)\varphi_0(y)\mu(y) = \varphi_0(x).$$

In other words, if we want to solve the partial differential equation $\frac{\partial \varphi}{\partial t} = L\varphi$ with initial condition $\varphi(0,x) = \varphi_0(x)$, the fundamental solution allows to reconstruct φ by convolution of φ_0 with $p(t,x,y)$.

Remark 3.2. *The choice of the volume form μ, that defines the fundamental solution, is not relevant in our study. Indeed the order of the asymptotics of $p(t,x,y)$ on the diagonal does not depend on the fixed smooth volume form. This can be proved by noting that the fundamental solution changes with respect to the given volume μ in the following way: let μ and ν be two volume forms on M, and let g be a smooth function such that $\nu = e^g\mu$. Let p_μ and p_ν denote the fundamental solutions of $\frac{\partial}{\partial t} - L$ with respect to μ and ν respectively. Then for every initial condition $\varphi_0 \in C_0^\infty(M)$, the solution $\varphi(t,x)$ of*

$$\begin{cases} \frac{\partial \varphi}{\partial t} = f_0(\varphi) + \frac{1}{2}\sum_{i=1}^k f_i^2(\varphi) \\ \varphi(0,x) = \varphi_0(x) \end{cases}$$

is given by

$$\begin{aligned} \varphi(t,x) &= \int_M p_\mu(t,x,y)\varphi_0(y)\mu(y) \\ &= \int_M p_\nu(t,x,y)\varphi_0(y)\nu(y) = \int_M p_\nu(t,x,y)\varphi_0(y)e^{g(y)}\mu(y), \end{aligned}$$

where the equalities follow since the solution is unique for smooth vector fields. Since φ_0 is arbitrary, we have

$$p_\mu(t,x,y) = e^{g(y)}p_\nu(t,x,y) \qquad \forall t > 0, \forall x,y \in M.$$

From the point of view of the asymptotics of the fundamental solution on the diagonal, it follows that the two asymptotics are the same for both volume forms up to a multiplicative constant $e^{g(x_0)} \neq 0$ depending on the relation between the two volumes and on the point where we compute the asymptotics.

For the study of the small time asymptotics on the diagonal, we will then suppose without loss of generality that $\mu = dx_1 \wedge \ldots \wedge dx_n$ near the point x_0.

Let x_0 be a point where the drift field lies in $\mathcal{D}^2_{x_0}$. In the rest of this section we explain the perturbative method, that we use for the proof of our results on the order of the asymptotic of $p(t,x_0,x_0)$.

Definition 3.3. *Let (U,x) be a coordinate neighborhood of x_0, i.e. $U \subset M$ is an open set, $x_0 \in U$ and $x = (x_1,\ldots,x_n) : U \to \mathbb{R}^n$ is such that $x(x_0) = 0$. Let $(w_1,\ldots,w_n)$ be positive integers, called* weights *of the coordinates $(x_1,\ldots,x_n)$.*

For $0 < \epsilon \leq 1$ *we define the* dilation, δ_ϵ, *of order* ϵ *and weights* $(w_1, \ldots, w_n)$ *around* x_0*, as the function* $\delta_\epsilon : U \to U$*, such that*

$$\delta_\epsilon(x_1, \ldots, x_n) := (\epsilon^{w_1} x_1, \epsilon^{w_2} x_2, \ldots, \epsilon^{w_n} x_n) \qquad \forall x = (x_1, \ldots, x_n) \in U.$$

For $\epsilon > 1$ *we define the dilation in the same way, but only from the smaller domain* $\delta_{\frac{1}{\epsilon}} U$*.*

Under the action of the dilations δ_ϵ, the coordinate functions and the coordinate vector fields behave as

$$x_i \circ \delta_\epsilon = \epsilon^{w_i} x_i, \qquad \delta_{\frac{1}{\epsilon}*} \frac{\partial}{\partial x_i} = \frac{1}{\epsilon^{w_i}} \frac{\partial}{\partial x_i} \qquad \forall 1 \leq i \leq n. \tag{3.10}$$

Here $\delta_{1/\epsilon *} X$ denotes the pushforward of a vector field X under the action of δ_ϵ. Let the volume μ be represented in the coordinate neighborhood (U, x) by $\mu = dx_1 \wedge \ldots \wedge dx_n$. By Remark 3.2 this assumption is not restrictive for the study of the asymptotics along the diagonal. Then the volume μ changes under the action of the dilation δ_ϵ as

$$\delta_\epsilon^*(dx_1 \wedge \ldots \wedge dx_n) = \epsilon^{\sum_{i=1}^n w_i} dx_1 \wedge \ldots \wedge dx_n, \tag{3.11}$$

where δ_ϵ^* denotes the pull-back.

When we apply a dilation to the space around x_0 and we rescale the time variable, also the fundamental solution is changed accordingly, as it is proved in the following proposition.

Proposition 3.4. *Let* (U, x) *be a coordinate neighborhood around the point* x_0 *and let* μ *be a volume form on* M *such that* $\mu = dx_1 \wedge \ldots \wedge dx_n$ *in* U*. For weights* $(w_1, \ldots, w_n)$ *and* $0 < \epsilon < 1$ *consider the dilation* $\delta_\epsilon : U \longrightarrow U$*. Let* α *be any real positive number. Let* $p(t, x, y)$ *be the fundamental solution of the operator in* (3.1) *with respect to the volume* μ*. Then the fundamental solution on* U *of the operator*

$$\frac{\partial}{\partial t} - \epsilon^\alpha \left(\delta_{1/\epsilon *} f_0 + \frac{1}{2} \sum_{i=1}^k \left(\delta_{1/\epsilon *} f_i \right)^2 \right) \tag{3.12}$$

is the function

$$q_\epsilon(t, x, y) := \epsilon^{\sum_{i=1}^n w_i} \, p(\epsilon^\alpha t, \delta_\epsilon(x), \delta_\epsilon(y)) \qquad \forall x, y \in U.$$

Remark 3.5. *The coefficient of normalization* $\epsilon^{\sum_{i=1}^n w_i}$*, that we have used to define* q_ϵ*, is necessary in order to reconstruct all the solutions of the differential operator, by convolution with* q_ϵ*. This coefficient appears as soon as we make a change of coordinates in the integral of the convolution. Moreover, we will see that this coefficient defines the order of the asymptotics of the fundamental solution for small time.*

Proof. First of all notice that the dilation $\delta_{1/\epsilon} : \delta_\epsilon(U) \to U$ can be defined only on the smaller neighborhood $\delta_\epsilon(U)$ of U. Then the fields $\delta_{1/\epsilon *} f_i$ are vector fields just on the coordinate neighborhood U.

Next let us prove the first property of the fundamental solution, i.e. that the function q_ϵ is a solution of the operator in (3.12). For convenience, we call ψ the dilation from

$\mathbb{R}^+ \times \delta_{1/\epsilon}U$ to $\mathbb{R}^+ \times U$ defined by $\psi(t,x) := (\epsilon^\alpha t, \delta_\epsilon(x))$. Then the function q_ϵ can be written as $q_\epsilon(t,x,y) = p(\psi(t,x), \delta_\epsilon(y))$ and the operator in (3.12) is

$$\psi_*^{-1}\frac{\partial}{\partial t} - \psi_*^{-1}f_0 - \frac{1}{2}\sum_{i=1}^k \left(\psi_*^{-1}f_i\right)^2,$$

where we are using a little abuse of notation, by considering $\frac{\partial}{\partial t}, f_0, \ldots, f_k$ as vector fields defined on the product space $\mathbb{R}^+ \times U$. Recall the definition of the pushforward of a vector field X under the action of a diffeomorphism ψ: for every function g we have $X_x(\left.(g\circ\psi)\right|_x) = \psi_{x*}(X)(g)|_{\psi(x)}$. Then we compute for fixed $y \in U$

$$\begin{aligned}
\left(\psi_*^{-1}\frac{\partial}{\partial t}\right) q_\epsilon(t,x,y) &= \epsilon^{\sum_{i=1}^n w_i} \left.\left(\psi_*^{-1}\frac{\partial}{\partial t}\right)\right|_{(t,x,y)} (p(\psi(t,x),\delta_\epsilon y))\\
&= \epsilon^{\sum_{i=1}^n w_i}\ \psi_*\left(\left.\left(\psi_*^{-1}\frac{\partial}{\partial t}\right)\right|_{(t,x,y)}\right) p|_{(\psi(t,x),\delta_\epsilon y)}\\
&= \epsilon^{\sum_{i=1}^n w_i}\ \left.\frac{\partial}{\partial t}\right|_{(\psi(t,x),\delta_\epsilon y)} p|_{(\psi(t,x),\delta_\epsilon y)}\\
&= \epsilon^{\sum_{i=1}^n w_i}\ L|_{(\psi(t,x),\delta_\epsilon y)}\ p|_{(\psi(t,x),\delta_\epsilon y)},
\end{aligned}$$

where the last equality follows since p is the fundamental solution of the operator $\frac{\partial}{\partial t} - L$. Applying the same computations to the fields $\psi_*^{-1}f_i$ for $i = 0,\ldots,k$, we find that q_ϵ satisfies

$$\psi_*^{-1}\frac{\partial}{\partial t}q_\epsilon(t,x,y) = \psi_*^{-1}f_0(q_\epsilon(t,x,y)) + \frac{1}{2}\sum_{i=1}^k \left(\psi_*^{-1}f_i\right)^2 (q_\epsilon(t,x,y))$$

and hence q_ϵ is a solution for the operator in (3.12).

Let us prove the second property of a fundamental solution. Here it becomes clear that the constant of normalization $\epsilon^{\sum_{i=1}^n w_i}$ in the definition of q_ϵ is exactly the parameter that we need in order to construct the other solutions of the partial differential equation by convolution with the fundamental solution. Indeed let us prove that for any $\varphi_0 \in L^2(U)$, it holds

$$\lim_{t\searrow 0}\int_U q_\epsilon(t,x,y)\varphi_0(y)\mu(y) = \varphi_0(x).$$

This follows by a change of variable and the same property valid for the fundamental solution $p(t,x,y)$:

$$\begin{aligned}
\lim_{t\searrow 0}\int_U q_\epsilon(t,x,y)\varphi_0(y)\mu(y) &= \lim_{t\searrow 0}\int_U \epsilon^{\sum_{i=1}^n w_i} p(\epsilon^\alpha t, \delta_\epsilon x, \delta_\epsilon y)\varphi_0(y)\mu(y)\\
&= \lim_{t\searrow 0}\int_M \epsilon^{\sum_{i=1}^n w_i} p(\epsilon^\alpha t, \delta_\epsilon x, \delta_\epsilon y)\varphi_0(y)\mu(y).
\end{aligned}$$

Here we integrate on M, by considering φ_0 as a function on M that is zero outside U. Now let us do a change of variable with $z = \delta_\epsilon y$. As computed in (3.11), the volume is transformed as $\mu(z) = \epsilon^{\sum_{i=1}^n w_i}\mu(y)$. Then

$$\begin{aligned}
\lim_{t\searrow 0}\int_U q_\epsilon(t,x,y)\varphi_0(y)dy &= \lim_{t\searrow 0}\int_M \epsilon^{\sum_{i=1}^n w_i} p(\epsilon^\alpha t, \delta_\epsilon x, z)(\varphi_0\circ\delta_{1/\epsilon})(z)\frac{\mu(z)}{\epsilon^{\sum_{i=1}^n w_i}}\\
&= (\varphi_0\circ\delta_{1/\epsilon})(\delta_\epsilon x) = \varphi_0(x),
\end{aligned}$$

where the second equality follows because p is a fundamental solution. $\square$

Let us investigate better the behavior of the fields f_i under the action of the dilations. We write every component $f_i^{(j)}$ of the fields f_i in a Taylor expansion centered in $x_0 = 0$ for x in the coordinate neighborhood U

$$f_i^{(j)}(x)\frac{\partial}{\partial x_j} = f_i^{(j)}(0)\frac{\partial}{\partial x_j} + \sum_{l=1}^{n} \frac{\partial f_i^{(j)}(0)}{\partial x_l} x_l \frac{\partial}{\partial x_j} + o(|x|).$$

By the properties of a dilation acting on the coordinate functions and on the coordinate vector fields, (3.10), when we apply a dilation to the vector fields f_i, every component has a different degree with respect to ϵ. Depending on the value of the weights $(w_1, \dots, w_n)$ and on the coefficients of the Taylor expansion of the fields f_i, for every $i = 0, \dots, k$, there exist an integer α_i and a principal vector field $\hat{f}_i$ such that

$$\delta_{\frac{1}{\epsilon}*} f_i = \frac{1}{\epsilon^{\alpha_i}} \hat{f}_i + o\left(\frac{1}{\epsilon^{\alpha_i}}\right),$$

where $\hat{f}_i$ contains the components of every $f_i^{(j)} \frac{\partial}{\partial x_j}$ that is homogeneous of degree $-\alpha_i$ with respect to the dilations. Applying this formula to the dilated operator in (3.12), we find that the operator L rescales as

$$\delta_{1/\epsilon*} f_0 + \frac{1}{2}\sum_{i=1}^{k} \left(\delta_{1/\epsilon*} f_i\right)^2 = \frac{1}{\epsilon^{\alpha_0}} \hat{f}_0 + \frac{1}{2}\sum_{i=1}^{k} \frac{1}{\epsilon^{2\alpha_i}} \hat{f}_i^2 + o\left(\frac{1}{\epsilon^{\alpha}}\right)$$

where $\alpha := \max\{\alpha_0, 2\alpha_1, \dots, 2\alpha_k\}$.

The main task in our study is to find suitable coordinates and good weights w_i, so that all the principal parts of the vector field f_0 and of $f_1^2, \dots, f_k^2$ rescale with the same degree α under the dilations, but they keep "enough" information from the original vector fields. Let L_0 be the operator defined by the principal vector fields $\hat{f}_0, \hat{f}_1, \dots, \hat{f}_k$ as

$$L_0 := \hat{f}_0 + \frac{1}{2}\sum_{i=1}^{k} \hat{f}_i^2.$$

Notice that by definition L_0 is homogeneous under the dilation, and in particular

$$\partial_t - \epsilon^{\alpha}\delta_{1/\epsilon*} L_0 = \partial_t - L_0.$$

(Here and in what follows ∂_t is a shorthand to denote the derivation in the time variable). Let us assume for a moment that $\partial_t - L_0$ admits a fundamental solution q_0. By Proposition 3.4, for every $\epsilon > 0$ the fundamental solution of $\partial_t - \epsilon^{\alpha}\delta_{1/\epsilon*} L_0$ is

$$q_\epsilon(t, x, y) := \epsilon^{\sum_{i=1}^{n} w_i} q_0(\epsilon^{\alpha} t, \delta_\epsilon x, \delta_\epsilon y) = q_0(t, x, y), \tag{3.13}$$

where the last identity follows since the dilated operator is again $\partial/\partial_t - L_0$.

Let us split the operator L as

$$\frac{\partial}{\partial t} - L = \frac{\partial}{\partial t} - L_0 + (L_0 - L), \tag{3.14}$$

where we have underlined the principal part L_0 plus a modification $L_0 - L$. To an operator like this we can apply Duhamel's formula, that gives the asymptotics of the fundamental solution as a perturbation of the asymptotics of the fundamental solution of the principal operator.

3.2.1 Duhamel's formula

In this section we recall briefly a famous formula, called Duhamel's formula, which allows to find the asymptotics of the fundamental solution of a perturbed operator, once we have the explicit fundamental solution of its principal part. This method, also called *parametrix technique*, is a perturbative method that has been already introduced in Chapter 3 of [44], in [6] and more in general in [36]. There exist also powerful tools to have two-sided pointwise estimates on the fundamental solution. For the non degenerate case we refer to the monograph by [25], while there are some recent results for the hypoelliptic case of the Kolmogorov operators in [23] and [24].

Let $\mathcal{L}$ be an operator on a Hilbert space with fundamental solution $p(t,x,y)$ (in our setting $\mathcal{L} = \partial_t - L$) and let us define the following operator on the Hilbert space

$$e^{t\mathcal{L}}\varphi(x) = \int p(t,x,y)\varphi(y)dy.$$

By the properties of the fundamental solution this is an heat operator $e^{t\mathcal{L}}$, i.e. an operator such that

$$\frac{\partial e^{t\mathcal{L}}\varphi}{\partial t} = \mathcal{L}e^{t\mathcal{L}}\varphi \quad \text{and} \quad \lim_{t\to 0} e^{t\mathcal{L}}\varphi = \varphi.$$

Suppose that $\mathcal{L}$ can be decomposed in a sum,

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{X},$$

of a principal part, $\mathcal{L}_0$, and a perturbation, $\mathcal{X}$, and assume that $\mathcal{L}_0$ has a well defined heat operator $e^{t\mathcal{L}_0}$. Then Duhamel's formula allows to reconstruct the heat operator of $\mathcal{L}$ by a perturbation of the heat operator of $\mathcal{L}_0$ (see Chapter 3 of [44] for a proof), namely

$$e^{t\mathcal{L}} = e^{t\mathcal{L}_0} + \int_0^t e^{(t-s)\mathcal{L}}\mathcal{X}e^{s\mathcal{L}_0}ds = e^{t\mathcal{L}_0} + e^{t\mathcal{L}} * \mathcal{X}e^{t\mathcal{L}_0}, \tag{3.15}$$

where with $*$ we denote the convolution operator between two operators, $A(t)$ and $B(t)$, on the Hilbert space:

$$(A * B)(t) = \int_0^t A(t-s)B(s)ds.$$

Let $a(t,x,y)$ and $b(t,x,y)$ be the heat kernels of $A(t)$ and $B(t)$ respectively and let $\mathcal{X}$ be an operator. Then the heat kernel of $(A * \mathcal{X}B)(t)$ is

$$(a * \mathcal{X}b)(t,x,y) = \int_0^t \int_M a(s,x,z)\mathcal{X}_z b(t-s,z,y)dzds.$$

Indeed, for any function φ in the Hilbert space, we have

$$\begin{aligned}
\left[(A * \mathcal{X}B)(t)\varphi\right](x) &= \left[\int_0^t A(t-s)\mathcal{X}B(s)ds\ \varphi\right](x) \\
&= \left[\int_0^t A(t-s)\left[\mathcal{X}\int_M b(s,\cdot,y)\varphi(y)dy\right]ds\right](x) \\
&= \int_0^t \int_M a(t-s,x,z)\mathcal{X}_z\left(\int_M b(s,z,y)\varphi(y)dy\right)dzds \\
&= \int_M \left(\int_0^t \int_M a(t-s,x,z)\mathcal{X}_z b(s,z,y))dzds\right)\varphi(y)dy.
\end{aligned}$$

From (3.15) we can now derive an approximation of the heat kernel $p(t,x,y)$ of the perturbed operator $\mathcal{L}$, by means of the heat kernel $p_0(t,x,y)$ of the principal operator:

$$p(t,x,y) = p_0(t,x,y) + (p * \mathcal{X} p_0)(t,x,y). \tag{3.16}$$

3.2.2 The perturbative method

We can apply Duhamel's formula to the operator in (3.14). Indeed it is the sum of a principal operator

$$L_0 = \hat{f}_0 + \frac{1}{2}\sum_{i=1}^{k} \hat{f}_i^2$$

perturbed by $\mathcal{X} := L_0 - L$. If we find good coordinates and weights, so that $\frac{\partial}{\partial t} - L_0$ admits a fundamental solution $q_0(t,x,y)$, then by Duhamel's formula (3.16) the asymptotics of the fundamental solution $p(t,x,y)$ is

$$p(t,x,y) = q_0(t,x,y) + (p * \mathcal{X} q_0)(t,x,y). \tag{3.17}$$

Recall the homogeneity property of the function q_0 written in equation (3.13) and choose $\epsilon = t^{-1/\alpha}$, then

$$\begin{aligned} q_0(t,x,y) =& q_0(\epsilon^{-\alpha}1, x, y) = \epsilon^{\sum_{i=1}^n w_i} q_0(t, \delta_\epsilon x, \delta_\epsilon y) \\ =& \frac{1}{t^{\sum_{i=1}^n w_i/\alpha}} q_0(1, \delta_{1/t^\alpha} x, \delta_{1/t^\alpha} y). \end{aligned}$$

Let us choose $x = y = x_0$ in Eq. (3.17), and let t go to zero. Then

$$\begin{aligned} p(t,x_0,x_0) &= q_0(t,x_0,x_0) + (p * \mathcal{X} q_0)(1,x_0,x_0) \\ &= \frac{1}{t^{\sum_{i=1}^n w_i/\alpha}} \Big(q_0(1,x_0,x_0) + t^{\sum_{i=1}^n w_i/\alpha} (p * \mathcal{X} q_0)(t,x_0,x_0) \Big). \end{aligned}$$

Provided we can control the error term $t^{\sum_{i=1}^n w_i/\alpha}(p * \mathcal{X} q_0)(t,x,y)$, we have then found the desired asymptotics.

In conclusion, in the choice of the coordinates (U,x) and the weights $(w_1,\dots,w_n)$ it will be important that the principal parts, $\hat{f}_0, \hat{f}_1, \dots, \hat{f}_k$, of the vector fields make homogeneous the principal part L_0 of the dilated operator and, moreover, that they satisfy Hörmander condition, so that it is guaranteed the existence of a fundamental solution q_0 of the principal operator. Finally we will need to check that the remainder term in the asymptotics of p goes to zero, as t goes to zero.

3.3 Graded structure induced by a filtration

In this section we introduce some notation and recall the definition of local graded structure of a manifold, induced by a filtration. This terminology is essential in order to find the right coordinates to rescale the differential operator L and to compute the order of the asymptotics of the fundamental solution. We constantly refer to Bianchini and Stefani's paper [20].

3.3.1 Chart adapted to a filtration

Let M be an n-dimensional smooth manifold, let $f_0, f_1, \ldots, f_k$ be smooth vector fields on M, that satisfy the Hörmander condition (3.2), and consider the hypoelliptic operator on $\mathbb{R} \times M$ defined as

$$\frac{\partial}{\partial t} - f_0 - \sum_{i=1}^{k} f_i^2.$$

(In this part of the chapter the boundedness assumption on the fields is not necessary).

The role played by the drift field f_0 and the other vector fields, $f_1, \ldots, f_k$, in the sum of squares, is different, and in particular the fields $f_1, \ldots, f_k$ are applied twice as many times as the drift field is. For this reason we want to treat differently the two kinds of fields, by giving to them two different *weights*.

Let Lie$X$ be the Lie algebra generated by a set $\{X_0, X_1, \ldots, X_k\}$ of noncommutative indeterminates.

Definition 3.6. *For every bracket Λ in* Lie$X$ *we denote by $|\Lambda|_i$ the number of times that the indeterminate X_i appears in the definition of Λ. We will call this number the* length *of Λ with respect to X_i.*

For example, the bracket $\Lambda = [X_0, [X_2, X_0]]$ has lengths $|\Lambda|_0 = 2$, $|\Lambda|_1 = 0$ and $|\Lambda|_2 = 1$, and it has zero length with respect to any other indeterminate.

By fixing a weight, l_i, to every indeterminate $X_0, \ldots, X_k$ we can define the weight of a bracket Λ.

Definition 3.7. *Given a set of integers $(l_0, l_1, \ldots, l_k)$, we define the* weight *of a bracket $\Lambda \in$* Lie$X$ *as*

$$||\Lambda|| := \sum_{i=0}^{k} l_i |\Lambda|_i \qquad \text{if } \Lambda \neq 0$$

and we set $||0|| = 0$.

In order to give different importance to the drift field, with respect to the other vector fields, in the following we fix the integers to be

$$l_0 = 2 \quad \text{and} \quad l_1 = \ldots = l_k = 1.$$

This means that the indeterminate X_0 will have weight 2, while the other indeterminates will have weight 1. For more complex Lie brackets, we have for example that the weight of the bracket considered before $\Lambda = [X_0, [X_2, X_0]]$ is

$$||\Lambda|| = 2 \cdot |\Lambda|_0 + 1 \cdot |\Lambda|_2 = 5.$$

By means of the weight of the indeterminates we introduce now a filtration of the Lie algebra spanned by $f_0, f_1, \ldots, f_k$ in the following way. For every bracket Λ in Lie$X$ we denote by Λ_f the vector field on M obtained by replacing every indeterminate X_i with the corresponding field f_i for $0 \leq i \leq k$. Then we define an increasing filtration $\mathcal{G} = \{G_i\}_{i \geq 0}$ of Vec(M) by

$$G_i = \text{span}\{\Lambda_f : \Lambda \in \text{Lie}X, ||\Lambda|| \leq i\}. \tag{3.18}$$

In other words G_i is the subalgebra of $\mathrm{Vec}(M)$ that contains all the vector fields obtained from a bracket of weight less then or equal to i. In particular, following our choice of weights, the first subspaces of the filtration are

$$\begin{aligned}
G_0 &= \{0\} \\
G_1 &= \mathrm{span}\{f_1, \dots, f_k\} \\
G_2 &= \mathrm{span}\{f_i,\, [f_i, f_j],\, f_0 : i, j = 1, \dots, k\} \\
G_3 &= \mathrm{span}\{f_i,\, [f_i, f_j],\, f_0,\, [f_i, [f_j, f_h]],\, [f_0, f_i] : i, j, h = 1, \dots, k\} \\
&\ \ \vdots
\end{aligned}$$

Notice moreover that for every $i, j \geq 0$, the following properties hold

- $G_i \subset G_{i+1}$
- $[G_i, G_j] \subset G_{i+j}$
- $\bigcup_{i\geq 0} G_i = \mathrm{Lie}\{f_0, f_1, \dots, f_k\}$ and $\bigcup_{i\geq 0} G_i(x) = T_x M$, for every $x \in M$, since by assumption the family $\{f_0, \dots, f_k\}$ satisfies the weak Hörmander condition (3.2).

When we evaluate $\mathcal{G}$ at the stationary point x_0 we get a stratification of the tangent space $T_{x_0}M$ at x_0. Let

$$k_i := \dim G_i(x_0) \qquad \forall i \geq 0.$$

In particular, $k_0 = 0$ and $k_1 \leq k$. Moreover, by Hörmander condition (3.2), there exists a smallest integer m such that $G_m(x_0) = T_{x_0}M$. We call this number the *step* of the filtration $\mathcal{G}$ at x_0.

The filtration $\mathcal{G}$ induces a particular choice of coordinates centered at x_0, as proved by the following proposition.

Proposition 3.8 (Bianchini, Stefani [20]). *There exists a chart (U, x) centered at x_0 such that for every $1 \leq j \leq m$*

(i) $G_j(x_0) = \mathrm{span}\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{k_j}}\}$

(ii) $D x_h(x_0) = 0$ *for every differential operator* $D \in \mathcal{A}^j := \{Z_1 \cdots Z_l$ *with* $Z_s \in G_{i_s}$ *and* $i_1 + \cdots + i_l \leq j\}$ *and for every* $h > k_j$.

Definition 3.9. *We call a chart that satisfies the properties of Proposition 3.8 an* adapted chart to the filtration $\mathcal{G}$ at x_0.

Since this kind of coordinates will reveal to be very important in our study, we give here the proof of the proposition, which relies upon the following Lemma:

Lemma 3.10. *Let m be the step of the filtration $\mathcal{G}$ at x_0 and let $j < m$ be an integer. Let $\varphi \in C^\infty(M)$ be such that $d_{x_0}\varphi \neq 0$ and $Z\varphi(x_0) = 0$ for all $Z \in G_j$. Then there exists an open neighborhood U of x_0 and a function $\hat{\varphi} \in C^\infty(U)$ such that*

- $d_{x_0}\varphi = d_{x_0}\hat{\varphi}$
- $D\hat{\varphi}(x_0) = 0$, *for every* $D \in \mathcal{A}^j = \{Z_1 \cdots Z_l : \ Z_s \in G_{i_s}, i_1 + \cdots + i_l \leq j\}$.

Proof. Let $Y_1, \dots, Y_n$ be vector fields on M such that they form a basis of $T_{x_0}M$ at x_0 and such that

- $\{Y_1, \dots, Y_{k_i}\}$ are in G_i and form a basis of G_i at x_0, for every $i \leq j$,
- $Y_i\varphi(x_0) = 0$ for every $i \leq n-1$,
- $Y_n\varphi(x_0) = 1$.

We choose the chart $y = (y_1, \dots, y_n)$ as the local inverse of the map

$$(y_1, \dots, y_n) \mapsto e^{y_nY_n} \cdots e^{y_1Y_1}x_0.$$

Then the function $\hat{\varphi} := y_n$ is such that $d_{x_0}\varphi = d_{x_0}\hat{\varphi}$.

Let $D = Z_1 \cdots Z_l \in \mathcal{A}^j$ with $Z_s \in G_{i_s}$ and $i_1 + \cdots + i_l \leq j$ and let us prove the second property required for the function $\hat{\varphi}$, by induction on l. Since $d_{x_0}\varphi = d_{x_0}\hat{\varphi}$ and $Z\varphi(x_0) = 0$ for all $Z \in G_j$ by hypothesis, the property is satisfied for $l = 1$. For $l > 1$, since $Z_l \in G_{i_l}$ we can write $Z_l(x_0) = \sum_{i=1}^{k_{i_l}} a_iY_i(x_0)$ for some a_i so

$$D\hat{\varphi}(x_0) = \sum_{i=1}^{k_{i_l}} a_i \left(Z_{l-1} \cdot Y_i \cdot Z_{l-2} \cdots Z_1 + [Y_i, Z_{l-1}] \cdot Z_{l-2} \cdots Z_1\right) \hat{\varphi}(x_0)$$

The second component on the left side vanishes because, by the definition of the filtration, $[G_i, G_h] \in G_{i+h}$, so we can apply on this component the induction hypothesis. By applying again the same commutation we have

$$D\hat{\varphi}(x_0) = \sum_{i=1}^{k_{i_l}} a_iZ_{l-1} \cdots Z_1Y_i\hat{\varphi}(x_0).$$

Iterating the same procedure also to $Z_{l-1}, \dots, Z_1$ we can write $D\hat{\varphi}(x_0)$ as a linear combination of elements of the type

$$Y_{i_l} \cdots Y_{i_1}\hat{\varphi}(x_0),$$

with $1 \leq i_h \leq k_j < n$ for every $h = 1, \dots, l$. Therefore we get $D\hat{\varphi}(x_0) = Dy_n(x_0) = 0$. □

Proof of Proposition 3.8. Let (U, x) be any chart centered at x_0. We can get a chart with property (i) of the proposition by a linear change of coordinates. Let us still denote it by (U, x). For every $i \leq n$ let j be such that $k_j < i \leq k_{j+1}$. Then the coordinate function x_i satisfies the hypothesis of the Lemma with respect to the integer j. By applying the Lemma to each function of the chart we get the statement. □

In Section 3.4.1 we present an example of adapted chart.

3.3.2 Graded structure

For $1 \leq i \leq m$ we define the integers

$$d_i := k_i - k_{i-1},$$

which indicate the number of new coordinates achieved with every new layer i of the filtration at x_0.

Definition 3.11. *Let us denote a point* $x = (x_1, \dots, x_n) \in U$ *by the* m*-tuple* $(x^1, x^2, \dots, x^m) \in \mathbb{R}^{d_1} \oplus \mathbb{R}^{d_2} \oplus \dots \oplus \mathbb{R}^{d_m}$, *where each component* $x^i := (x_{k_{i-1}+1}, \dots, x_{k_i})$ *is a vector of length* d_i*. Then for every* $0 < \epsilon \leq 1$ *we define the* anisotropic dilations *around* x_0 *of factor* ϵ *as*

$$\delta_\epsilon(x) = \delta_\epsilon(x^1, \dots, x^m) := (\epsilon x^1, \epsilon^2 x^2, \dots, \epsilon^m x^m).$$

For $\epsilon > 1$ *we define* δ_ϵ *in the same way, but it will be defined only on* $\delta_{\frac{1}{\epsilon}} U$.

The dilations δ_ϵ act on every coordinate function and on the coordinate vector fields with a different weight, namely

$$x_j \circ \delta_\epsilon = \epsilon^i x_j, \quad (\delta_\epsilon)_* \frac{\partial}{\partial x_j} = \epsilon^i \frac{\partial}{\partial x_j} \qquad \forall k_{i-1} < j \leq k_i.$$

For every $1 \leq i \leq n$ let w_j be the order of expansion of the coordinate x_j, that is $w_j := i$ if $k_{i-1} < j \leq k_i$. We call w_j the *weight* of the coordinate x_j. Then the dilation δ_ϵ is a particular choice of dilations of Definition 3.3 with respect to the coordinates induced by the filtration $\mathcal{G}$ and the weights $(w_1, \dots, w_n)$.

Accordingly we define the *weight* of a monomial to be

$$\mathcal{W}(x_1^{\alpha_1} \cdots x_n^{\alpha_n}) := \sum_{j=1}^{n} \alpha_j w_j,$$

and the weight of a polynomial to be the greatest order of its monomials. Moreover we define the *graded order*, $\mathcal{O}(g)$, of a function $g \in C^\infty(U)$ to be the smallest weight of the monomials that appear in any Taylor approximation of g at x_0.

For example, let $n = 2$ and suppose that x_1 has weight 1 and x_2 has weight 2. Then the polynomial $x_1 x_2 - \frac{(x_1)^2 (x_2)^2}{6}$ has weight 6, because the two monomials composing it are $x_1 x_2$ of weight 3 and the rest of weight 6. On the other hand, $\sin(x_1 x_2) = x_1 x_2 - \frac{(x_1)^2(x_2)^2}{6} + o\left((x_1)^2 (x_2)^2\right)$ has graded order 3.

We extend these definitions to differential operators. We say that a polynomial vector field Z is homogeneous of *weight* i if

$$\mathcal{W}(Z\varphi) = \mathcal{W}(\varphi) - i \qquad \forall \text{ monomial } \varphi \text{ of weight } \mathcal{W}(\varphi).$$

In other words Z subtracts weight i to every function. Then the weight of a polynomial vector field is the smallest weight of its homogeneous components. We define the *graded order*, $\mathcal{O}(D)$, of a differential operator D by saying that

$$\mathcal{O}(D) \leq j \quad \text{if and only if} \quad \mathcal{O}(D\varphi) \geq \mathcal{O}(\varphi) - j \qquad \forall \text{ polynomial } \varphi,$$

that is D subtracts at most weight j from the functions. For example the graded order of a vector field like $(x_1^{\alpha_1} \cdots x_n^{\alpha_n}) \frac{\partial}{\partial x_h}$ is

$$\mathcal{O}\left((x_1^{\alpha_1} \cdots x_n^{\alpha_n}) \frac{\partial}{\partial x_h} \right) = w_h - \left(\sum_{j=1}^{n} \alpha_j w_j \right).$$

Coming back to the previous example, the graded order of a field like $\sin(x_1 x_2) \frac{\partial}{\partial x_1}$ is obtained as $1 - \mathcal{O}(\sin(x_1 x_2)) = -2$.

By means of the graded order we can give a generalization of the concept of Taylor approximation of a function up to weight h. Namely, for any $\varphi \in C^\infty(U)$ and every integer $h \geq 0$, there is a unique polynomial $\varphi_{(h)}$ of weight h such that $\mathcal{O}(\varphi - \varphi_{(h)}) \geq h$.

Definition 3.12. *The polynomial* $\varphi_{(h)}$ *is called the* graded approximation of weight h of φ *and it is the sum of the polynomials of weight less then or equal to* h *in the formal Taylor expansion of* φ *at* x_0.

For every vector field $V \in \mathrm{Vec}(U)$ and each integer $h \leq m$ there is a polynomial vector field $V_{(h)}$ of weight h such that $\mathcal{O}(V - V_{(h)}) \leq h - 1$.

Definition 3.13. $V_{(h)}$ *is called the* graded approximation of weight h of V *and it is the sum of the homogeneous vector fields of weight greater than or equal to* h *in the formal Taylor expansion of* V *at* x_0.

Notice that, since $V_{(h)}$ is a polynomial vector field, we can consider it as defined on the whole Euclidean space $\mathbb{R}^n$.

We will see in the next sections how to apply this graded structure in order to underline the most important properties of the operator in (3.1), concerning the small time asymptotics of its fundamental solution.

3.4 Nilpotent approximation and the order of the dilations

In this section we apply the graded structure, that we have just developed, to define a special class of vector fields, which approximate the original one $f_0, f_1, \ldots, f_k$ and we show an example to clarify the setting. Finally we compute how the dilations change the volume form and we introduce the order that will appear in the asymptotics of the heat kernel.

3.4.1 Nilpotent approximation

Let (x, U, w) be the graded structure around x_0 introduced in Section 3.3 and $f_0, f_1, \ldots, f_k$ be the vector fields used to define the filtration of $T_{x_0}M$. Then as proved in [20] Theorem 3.1, for every $f \in G_i$, we have a bound on the graded order, namely $\mathcal{O}(f) \leq i$, where $\mathcal{O}$ is the graded order associated to the graded structure (x, U, w) defined in Section 3.3.

Recall the integers $l_0 = 2$ and $l_1 = \ldots = l_k = 1$ introduced in Section 3.3 to define the filtration and denote by $\hat{f}_i$ the graded approximation of weight l_i of f_i. In other words, $\hat{f}_i$ has weight l_i and $\mathcal{O}(f_i - \hat{f}_i) < l_i$.

Definition 3.14. *The vector fields* $\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_k$ *are called the* nilpotent approximation *of* $f_0, f_1, \ldots, f_k$.

The fields $\hat{f}_i$, for $0 \leq i \leq k$, are polynomials, so they can be defined on $\mathbb{R}^n$.

We can describe more precisely the structure of the approximating fields $\hat{f}_i$:

- $\hat{f}_0$ contains the terms of weight 2; therefore every component $\hat{f}_0^j$ of $\hat{f}_0$ depends only linearly on the coordinates of weight $w_j - 2$ and more than linearly on the coordinates of less weight, but does not depend on the coordinates of weight greater than or equal to $w_j - 1$, that are x_h with $h > k_{j-2}$.

- $\hat{f}_i$ contains the terms of weight 1 for $i = 1, \dots, k$; therefore every component $\hat{f}_i^j$ of $\hat{f}_i$ depends only linearly on the coordinates of weight $w_j - 1$ and more than linearly on the coordinates of less weight, but do not depend on the coordinates of weight greater than or equal to w_j, that are x_h with $h > k_{j-1}$.

To make the construction more clear we end this subsection with an example, in which we present the filtration in x_0, the induced adapted chart and the graded structure, and we find the related nilpotent approximation.

Example 1. *Let $M = \mathbb{R}^2$, and let the number of controlled vector fields be $k = 1$. Define the vector fields*

$$f_1 := \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_2} \quad \textit{and} \quad f_0 := \sin(x_1^2) \frac{\partial}{\partial x_2}$$

and recall the choice of weights $l_0 = 2$ and $l_1 = 1$. The non vanishing Lie brackets that contribute to span the tangent space in any point are given by

$$[f_1, f_0] = 2x_1 \cos(x_1^2) \frac{\partial}{\partial x_2}, \qquad [f_1, [f_1, f_0]] = (2\cos(x_1^2) - 4x_1^2 \sin(x_1^2)) \frac{\partial}{\partial x_2}.$$

Then Hörmander assumption (3.2) *holds in any point and the filtration defined in* (3.18) *is equal to*

- $G_1 = \text{span}\{f_1\}$
- $G_2 = \text{span}\{f_1, f_0\}$
- $G_3 = \text{span}\{f_1, f_0, [f_1, f_0]\}$
- $G_4 = \text{span}\{f_1, f_0, [f_1, f_0], [f_1, [f_1, f_0]]\}$

Let $x_0 \in \mathbb{R}^2$ be a stationary point of the drift field, and center the coordinates so that $x_0 = (0, 0)$. The filtration in x_0 is given by

$$G_1(x_0) = G_2(x_0) = G_3(x_0) = \text{span}\{\frac{\partial}{\partial x_1}\} \quad \textit{and} \quad G_4(x_0) = \mathbb{R}^2 \tag{3.19}$$

and the dimensions are: $k_1 = k_2 = k_3 = 1$, $k_4 = 2$.

Let us find an adapted chart to the filtration at $x_0 = (0, 0)$. As one can easily see, the coordinates (x_1, x_2) are not adapted, since $f_1^2(x_2)\big|_{x_0} = 1 \neq 0$ and the second property of the adapted chart then fails. Following the constructive proof of Lemma 3.10, one can find that the new coordinates (y_1, y_2) defined by

$$\begin{cases} y_1 = x_1 - \frac{x_1^2}{2} + x_2 \\ y_2 = -\frac{x_1^2}{2} + x_2 \end{cases}$$

give an adapted chart at $(0, 0)$. In these coordinates the two vector fields are written as

$$f_1 = \frac{\partial}{\partial y_1} \quad \textit{and} \quad f_0 = \sin((y_1 - y_2)^2) \left(\frac{\partial}{\partial y_1} + \frac{\partial}{\partial y_2} \right).$$

For $\epsilon > 0$ the dilations defined in Definition 3.11 are

$$\delta_\epsilon : (y_1, y_2) \mapsto (\epsilon y_1, \epsilon^4 y_2). \tag{3.20}$$

Then the weights of the coordinate functions are $\mathcal{W}(y_1) = 1$ and $\mathcal{W}(y_2) = 4$, and the weights of the coordinate vector fields are $\mathcal{W}(\frac{\partial}{\partial y_1}) = 1$ and $\mathcal{W}(\frac{\partial}{\partial y_2}) = 4$.

Finally, let us write the Taylor expansion of the two vector fields f_1, f_0:

$$\begin{aligned} f_1 &= \frac{\partial}{\partial y_1} \\ f_0 &= \left(y_1^2 - 2y_1y_2 + y_2^2 + o(|(y_1, y_2)|^2)\right) \frac{\partial}{\partial y_2} \end{aligned}$$

We can see that f_1 has already weight 1, while the only part of weight 2 in f_0 is $y_1^2 \frac{\partial}{\partial y_2}$. We can therefore conclude that the nilpotent approximation of f_0, f_1 is given by

$$\hat{f}_0 = y_1^2 \frac{\partial}{\partial y_2} \quad \text{and} \quad \hat{f}_1 = \frac{\partial}{\partial y_1}.$$

3.4.2 Order of the dilations

We analyze here the order of the dilations, that is the order of homogeneity of the volume form under the action of the dilations. This number will be crucial to find the order of degeneracy of the fundamental solution of the operator (3.1).

Let us consider the dilations δ_ϵ. They were defined by introducing the notation $x = (x^1, \ldots, x^m)$, where each component x^i is a vector of length $d_i = k_i - k_{i-1}$. Then we set $\delta_\epsilon(x^1, x^2, \ldots, x^m) = (\epsilon x^1, \epsilon^2 x^2, \ldots, \epsilon^m x^m)$. Let $\mathcal{N}$ be the order of homogeneity of the volume form $dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n$ around the point x_0, that is a number such that

$$(\delta_\epsilon)_*(dx_1 \wedge dx_2 \wedge \ldots \wedge x_n) = \epsilon^{\mathcal{N}} dx_1 \wedge dx_2 \wedge \ldots \wedge dx_n.$$

Then $\mathcal{N}$ is given by

$$\mathcal{N} := \sum_{i=1}^{m} i \cdot d_i = \sum_{i=1}^{m} i \left(\dim G_i(x_0) - \dim G_{i-1}(x_0)\right). \tag{3.21}$$

Since this number is very important we give here some examples to understand its meaning.

Example 2 (Continuation of Example 1)**.** *As a first example we consider the one given in Example 1. We have already computed the filtration in Eq.* (3.19)*, so we already know that the integers $d_i := \dim G_i - \dim G_{i-1}$ are*

$$d_1 = 1 \qquad d_2 = d_3 = 0 \quad \text{and} \quad d_4 = 1.$$

Therefore the order of the dilations is $\mathcal{N} = 1 \cdot 1 + 4 \cdot 1 = 5$, as one can compute directly from the explicit expression of the dilations in (3.20)*.*

Example 3 (Sub-Riemannian manifold)**.** *Let us assume that the operator in* (3.1) *is induced by a sub-Riemannian structure. In other words, we consider an operator without drift field, and the vector fields $f_1, \ldots, f_k$ generate a completely non-holonomic distribution,*

Δ, of step m. Recall that the growth vector of the distribution is defined as the vector, at any point q of the manifold, given by

$$(\Delta(q), \Delta^2(q), \dots, \Delta^m(q)) \qquad \textit{where } \Delta^{i+1} := \Delta^i + [\Delta, \Delta^i].$$

Then the filtration $G_i = \Delta^i$ for every i and the integers d_i related to the filtration are the same defined by the growth vector, i.e. $d_i = \dim(\Delta^i) - \dim(\Delta^{i-1})$. Finally the number $\mathcal{N}$ is exactly the homogeneous dimension $\mathcal{Q}$ of the manifold. More explicitly

$$\mathcal{N} = \mathcal{Q} = 1 \cdot d + 2 \cdot d_2 + \cdots m \cdot d_m = \sum_{i=1}^{m} i(\dim(\Delta^i) - \dim(\Delta^{i-1})).$$

In particular for the $2n+1$ dimensional Heisenberg group, for which $\dim \Delta = 2n$ and $[\Delta, \Delta]_q = T_qM$ for every $q \in M$, the homogeneous dimension is $\mathcal{Q} = 2n+2$.

Example 4 (Linear case)**.** *As a last example, we consider an involutive distribution $\mathcal{D}$, spanned locally by k constant vector fields, and assume that the drift field is linear in a neighborhood of x_0.*

The operators arising from such a structure include the Kolmogorov equations, that appear in diffusion theory, probability and finance. Important results on this type of equations have been achieved by Lanconelli, Pascucci and Polidoro [31, 32]. See also [22] for an analysis for continuity methods.

Without loss of generality we can assume that

$$f_i = \frac{\partial}{\partial x_i} \quad \forall 1 \le i \le k \qquad \textit{and} \qquad f_0 = \sum_{i,j=1}^{n} A_{ij} x_i \frac{\partial}{\partial x_j}$$

for some constants A_{ij}. Under these assumptions, the only Lie brackets different from zero are the one involving only one vector field of the distribution and the drift field. Let us call A the $n \times n$ matrix with entries equal to A_{ij} and B the $n \times k$ matrix that is the identity in the first k rows and is equal to zero in the last $n-k$ rows. Then Hörmander's condition of hypoellipticity (3.2) *is also called in this linear setting* Kalman's condition of controllability *for linear control systems and becomes the following condition on the rank of Kalman's $n \times (nk)$ matrix*

$$\operatorname{rank}[B, AB, A^2B, \dots, A^{n-1}B] = n. \tag{3.22}$$

The filtration is then completely determined and we have

$$G_{2i-1}(x_0) = G_{2i}(x_0) = \operatorname{span}\{A^j B : 0 \le j \le i-1\}.$$

Consequently the numbers $k_{2i-1} = k_{2i}$ are determined by the rank of the Kalman's matrix in (3.22)*, where we stop the series of matrices at $A^{i-1}B$. The numbers d_j are zero if j is even, while if $j = 2i-1$ they are the number of new linearly independent columns obtained by adding the matrix $A^{i-1}B$ to the previous one. The step of the distribution is then an odd number $2\tilde{m}-1$ and $\mathcal{N}$ is equal to an odd sum of integers:*

$$\mathcal{N} = \sum_{i=1}^{\tilde{m}} (2i-1) d_{2i-1} = 1 \cdot d_1 + 3 \cdot d_3 + 5 \cdot d_5 + \cdots + (2\tilde{m}-1) d_{2\tilde{m}-1}.$$

3.4.3 Comparison between $\mathcal{N}$ and the sub-Riemannian dimension $\mathcal{Q}$

Let us consider the case in which $f_1, \ldots, f_k$ satisfy the strong Hörmander condition, i.e., the Lie algebra doesn't require the drift field to generate the tangent space (see (sHC)). As explained in the Introduction, in this case it is well defined a distance function, called the *sub-Riemannian* or *CC-distance*, and the sub-Riemannian homogeneous dimension $\mathcal{Q}$. We want to compare the dimension $\mathcal{Q}$ and the integer $\mathcal{N}$ defined by the filtration $\mathcal{G}$ at a stationary point x_0 of the drift field f_0.

Let $\Delta_x := \text{span}\{f_1, \ldots, f_k\}_x$ be the horizontal distribution at $x \in M$ and for $i \geq 2$ let

$$\Delta^i := \Delta^{i-1} + [\Delta^{i-1}, \Delta]$$

be the filtration defined by the distribution, where $\Delta^1 = \Delta$. Let m_Δ be the *step* of the distribution Δ at x_0, i.e. the smallest integer j such that $\Delta^j_{x_0} = T_{x_0}M$. In other words the subspace $\Delta^i_x \subset T_xM$ is spanned by all the Lie brackets up to length i between the fields $f_1, \ldots, f_k$:

$$\Delta^i = \text{span}\{[f_{i_1}, \ldots, [f_{i_{l-1}}, f_{i_l}]] : \text{ for all } 1 \leq i_1, \ldots, i_l \leq k \text{ and } l \leq i\}.$$

At the same time we can built the filtration $\mathcal{G}$ at x_0 defined in (3.18), which involves also the drift field f_0. In general it holds the inclusion

$$\Delta^i_{x_0} \subset G_i(x_0). \tag{3.23}$$

Therefore the dimensions k_i of $G_i(x_0)$ are greater or equal than the dimensions $\tilde{k}_i$ of $\Delta^i_{x_0}$ and the step m of the filtration $\mathcal{G}$ is smaller or equal than the step m_Δ of the distribution. To gain an inequality between the integers $\mathcal{N}$ and $\mathcal{Q}$ notice that we can rewrite the sum

$$\mathcal{N} := \sum_{i=1}^{m} i d_i = \sum_{j=1}^{m} \sum_{i=j}^{m} d_i = \sum_{j=1}^{m} (n - k_{j-1}),$$

where we recall that $d_i = k_i - k_{i-1}$. The same identities can be written for the sum defining $\mathcal{Q}$. Therefore it always holds the inequality

$$\mathcal{N} := \sum_{j=1}^{m} (n - k_{j-1}) \leq \sum_{j=1}^{m} (n - \tilde{k}_{j-1}) \leq \sum_{j=1}^{m_\Delta} (n - \tilde{k}_{j-1}) = \mathcal{Q}.$$

It becomes an equality if only if $\Delta^i_{x_0} = G_i(x_0)$ for every layer i.

Remark 3.15. *The identities $\Delta^i_{x_0} = G_i(x_0)$ are verified in particular if $f_0 \in \Delta^2$, because f_0 can be written equivalently as a combination of Lie brackets between the fields $f_1, \ldots, f_k$ up to length 2 and then it does not play any role in the construction of the $G_i(x_0)$. This is the case studied by Ben Arous in [15] and recalled in* (3.3)*.*

If $f_0 \in \Delta^i \setminus \Delta^2$ for some $i > 2$, then the inclusion in (3.23) could be strict, because the Lie brackets between f_0 and some other fields could generate new dimensions in the layers of the filtration $\mathcal{G}$ that could be reached by the filtration of Δ only with longer combination of Lie brackets. For example, let us take on $\mathbb{R}^2$ the fields $f_1 = \frac{\partial}{\partial x_1}$, $f_2 = \frac{x_1^3}{6}\frac{\partial}{\partial x_2}$ and $f_0 = [f_1,[f_1,f_2]] = x_1\frac{\partial}{\partial x_2}$, and consider the two filtrations at the origin x_0. We have

$$\Delta^1_{x_0} = \Delta^2_{x_0} = \Delta^3_{x_0} = \operatorname{span}\left\{\frac{\partial}{\partial x_1}\right\} \quad \text{and} \quad \Delta^4_{x_0} = \mathbb{R}^2,$$

while the filtration $\mathcal{G}$ at x_0 as only 3 layers, namely

$$G_1(x_0) = G_2(x_0) = \operatorname{span}\left\{\frac{\partial}{\partial x_1}\right\} \quad \text{and} \quad G_3(x_0) = \operatorname{span}\{f_1,[f_1,f_0]\}_{x_0} = \mathbb{R}^2,$$

therefore $\mathcal{Q}$ is strictly bigger then $\mathcal{N}$.

Notice that the behavior of this last example can not happen if $m_\Delta \leq 3$, then in this case it always holds $\mathcal{N} = \mathcal{Q}$.

3.5 Small time asymptotics on the diagonal

We come back now to the perturbative method explained in Section 3.2. Let

$$\frac{\partial \varphi}{\partial t} - f_0(\varphi) - \frac{1}{2}\sum_{i=1}^{k} f_i^2(\varphi) \qquad \forall \varphi \in C^\infty(\mathbb{R}\times M)$$

be the differential operator (3.1) on $\mathbb{R}^+ \times M$ and assume that $f_0, f_1, \ldots, f_k$ satisfy the Hörmander condition (3.2) and are bounded with bounded derivatives of any order.

The role played by the drift field f_0 and the other vector fields, $\{f_1,\ldots,f_k\}$, in the sum of squares, is different, and in particular the fields $\{f_1,\ldots,f_k\}$ are applied twice as many times as the drift field is. Therefore we give to f_0 weight 2 and to $f_1,\ldots,f_k$ weight 1. Consequently fix the corresponding graded structure, (U,x,w), around x_0, that is induced by the filtration $\mathcal{G}$ introduced in Section 3.3.

The fields $f_0, f_1, \ldots, f_k$ can be written in terms of the nilpotent approximation as

$$f_0 = \hat{f}_0 + g_0, \qquad\qquad f_i = \hat{f}_i + g_i \qquad 1 \leq i \leq k,$$

where g_0 and g_i are vector fields of order less than or equal to 1 and 0 respectively. Let

$$L_0 := \hat{f}_0 + \frac{1}{2}\sum_{i=1}^{k} \hat{f}_i^2$$

and write $\partial_t - L = \partial_t - L_0 + (L_0 - L)$. To apply Duhamel's formula (3.16) to this kind of operator, we need to prove that there exists the fundamental solution of the principal operator $\partial/\partial_t - L_0$. As we will prove now, this follows by the property of hypoellipticity of the original operator, that are preserved by the nilpotent approximation, that defines L_0. The same statement can be found also in the paper by Bianchini and Stefani [20].

Proposition 3.16. *Let $\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_k$ be the nilpotent approximation of the fields $f_0, f_1, \ldots, f_k$ defined in Definition 3.14.*

(i) For every bracket Λ *such that* $||\Lambda|| = i$*, then* $(\Lambda_f - \Lambda_{\hat{f}}) \in G_{i-1}(x_0)$ *and* $\Lambda_{\hat{f}}(x_0) = 0$ *whenever* $\Lambda_f(x_0) \in G_{i-1}(x_0)$*, where* $||\Lambda||$ *denotes the weight of the bracket* Λ *defined in Definition 3.7.*

(ii) Assume $G_m(x_0) = \mathrm{Lie}_{x_0}\{f_0, f_1, \ldots, f_k\}$*, then*

$$\mathrm{Lie}_{x_0}\{f_0, f_1, \ldots, f_k\} = \mathrm{Lie}_{x_0}\{\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_k\}.$$

Proof. Let us prove the first statement. Let (U, x) be coordinates around x_0 adapted to the filtration $\{G_i\}_i$. Then for every i,

$$G_i(x_0) = \mathrm{span}\left\{ \left.\frac{\partial}{\partial x_1}\right|_{x_0}, \ldots, \left.\frac{\partial}{\partial x_{k_i}}\right|_{x_0} \right\}. \tag{3.24}$$

Let Λ be a bracket such that $||\Lambda|| = i$, then as proved in [20] Theorem 3.1, $\mathcal{O}(\Lambda_f) \leq i$, where $\mathcal{O}$ is the graded order associated to the graded structure induced by the filtration. Therefore there exist constants a_j such that

$$\Lambda_f(x_0) = \sum_{j \leq k_i} a_j \left.\frac{\partial}{\partial x_j}\right|_{x_0}. \tag{3.25}$$

Notice that if two vector fields h_1, h_2 are homogeneous of graded order respectively n_1 and n_2, then their Lie bracket is either zero or homogeneous of order $n_1 + n_2$. Then the Lie bracket $\Lambda_{\hat{f}}$ is either zero or homogeneous of order $||\Lambda|| = i$. Therefore by equation (3.25), we have

$$\Lambda_{\hat{f}}(x_0) = \sum_{k_{i-1} < j \leq k_i} a_j \left.\frac{\partial}{\partial x_j}\right|_{x_0}. \tag{3.26}$$

By subtracting (3.26) to (3.25), we find that $(\Lambda_f - \Lambda_{\hat{f}}) \in G_{i-1}(x_0)$, because $G_i(x_0)$ is obtained as in (3.24).

Let us prove the second statement. Let (U, x) be as before and let

$$\hat{V}(x_0) \in \mathrm{Lie}_{x_0}\{\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_k\}.$$

Then $\hat{V}(x_0) = \Lambda_{\hat{f}}(x_0)$, for some bracket Λ with $||\Lambda|| = j$ equal to the graded order of $\hat{V}(x_0)$. Then by expression (3.24), there exist α_i such that $\hat{V}(x_0) = \sum_{i \leq k_j} \alpha_i \frac{\partial}{\partial x_i}$. Since $G_j(x_0) \subset G_m(x_0) = \mathrm{Lie}_{x_0}\{f_0, f_1, \ldots, f_k\}$, we have that $\hat{V}(x_0) \in \mathrm{Lie}_{x_0}\{f_0, f_1, \ldots, f_k\}$.

We prove the other inclusion by proving that $G_i(x_0) = \mathrm{span}\{\Lambda_{\hat{f}}(x_0) : ||\Lambda|| \leq i\}$, for every i. We prove it by induction on i.

For $i = 1$, $G_1(x_0) = \mathrm{span}\{\Lambda_f(x_0) : ||\Lambda|| \leq 1\}$. Let $\Lambda_f(x_0) \in G_1(x_0)$, then by statement (i), $(\Lambda_f - \Lambda_{\hat{f}})(x_0) \in G_0(x_0) = \{0\}$. Then $\Lambda_f(x_0) = \Lambda_{\hat{f}}(x_0)$ and the statement is true for $i = 1$.

Assume that the statement is true for $i - 1$. Recall that

$$G_i(x_0) = \mathrm{span}\{\Lambda_f(x_0) : ||\Lambda|| \leq i\}$$

and $(\Lambda_f - \Lambda_{\hat{f}})(x_0) \in G_{i-1}(x_0)$. By the induction hypothesis, there exists $g \in \text{span}\{\Lambda_{\hat{f}}(x_0) : ||\Lambda|| \leq i-1\}$ such that

$$\Lambda_f(x_0) = \Lambda_{\hat{f}}(x_0) + g.$$

And the statement is proved also for i.

We conclude, since $\text{Lie}_{x_0}\{f_0, f_1, \dots, f_k\} = G_m(x_0) \subset \text{Lie}_{x_0}\{\hat{f}_0, \hat{f}_1, \dots, \hat{f}_k\}$. □

Corollary 3.17. *Recall that x_0 is a point where $f_0(x_0) \in \mathcal{D}^2_{x_0} = \text{span}_{x_0}\{f_i, [f_i, f_j] : 1 \leq i, j \leq k\}$. Then the operator $\partial/\partial_t - L_0$ is hypoelliptic on $\mathbb{R}^n$.*

Proof. By Hörmander's condition of hypoellipticity we know that

$$\text{Lie}_{x_0}\{f_0, f_1, \dots, f_k\} = \mathbb{R}^n.$$

Then the hypothesis of statement (ii) of the Proposition are fulfilled and also the nilpotent approximation is Lie bracket generating. To guarantee the hypoellipticity of $\partial/\partial_t - L_0$, however, we need that the field f_0 gives a contribution in the generating process only if it is applied to a Lie bracket with some other vector fields. In other words, we want that f_0 alone gives no contribution.

Let us suppose first that x_0 is an equilibrium point for the drift. Then $f_0(x_0) = 0 = \hat{f}_0(x_0)$ and the weak Hörmander's condition is immediately satisfied, i.e.,

$$\text{span}_{x_0}\{\frac{\partial}{\partial t} - \hat{f}_0, \hat{f}_1, \dots, \hat{f}_k\} = \mathbb{R}^{n+1}. \tag{3.27}$$

More in general, let $f_0(x_0) \in \mathcal{D}^2_{x_0}$. By point (i) in the previous proposition, $\hat{f}(x_0) = f(x_0) + g_1$ for a vector $g_1 \in \text{span}_{x_0}\{f_1, \dots, f_k\} = \text{span}_{x_0}\{\hat{f}_1, \dots, \hat{f}_k\}$. On the other hand, by hypothesis $f_0(x_0) = [f_i, f_j]_{x_0}$ for some $1 \leq i, j \leq k$. Then the proposition implies that there exists $g_2 \in \text{span}_{x_0}\{\hat{f}_1, \dots, \hat{f}_k\}$ such that $f_0(x_0) = [\hat{f}_i, \hat{f}_j]_{x_0} + g_2$. In conclusion we have proved that $\hat{f}(x_0) = [\hat{f}_i, \hat{f}_j]_{x_0} + g_1 + g_2$, and then $\hat{f}_0(x_0)$ alone does not give any contribution in the Lie bracket generating condition. So (3.27) holds again.

Using the lower semi-continuity of the rank, we can find a small neighborhood U of x_0 where the Hörmander condition holds at any point.

Now by the homogeneity of the approximating system we know that

$$\delta_{\epsilon *}\hat{f}_0 = \epsilon^2 \hat{f}_0 \qquad \text{and} \qquad \delta_{\epsilon *}\hat{f}_i = \epsilon \hat{f}_i \quad \forall 1 \leq i \leq k.$$

Therefore, since the differential operator commutes with the Lie brackets, we can extend Hörmander condition, which holds on a neighborhood of x_0, to the whole Euclidean space $\mathbb{R}^n$ and the operator $\partial/\partial_t - L_0$ is hypoelliptic on $\mathbb{R}^n$. □

Remark 3.18. *In the proof of this corollary, the assumption on x_0 is important, because it permits to say that the approximating system $\hat{f}_0, \hat{f}_1, \dots, \hat{f}_k$ satisfies not only the strong Hörmander condition (that is always guarantied), but also the weak one* (3.2). *Indeed if it fails, there are cases in which the approximating fields do not satisfy condition* (3.2) *and L_0 is not hypoelliptic, even if L is.*

For example, on $\mathbb{R}^2$ the fields

$$f_1 = \frac{\partial}{\partial x_1} \qquad f_0 = (1 + x_1)\frac{\partial}{\partial x_2}$$

satisfy the Hörmander condition (3.2)*, but this fails for their nilpotent approximation* $\hat{f}_1 = \frac{\partial}{\partial x_1}$ *and* $\hat{f}_0 = \frac{\partial}{\partial x_2}$*. Indeed, even if* $\partial_t - f_0 - \frac{1}{2}f_1^2$ *is hypoelliptic, its principal part* $\partial_t - \frac{\partial}{\partial x_2} - \frac{1}{2}\frac{\partial^2}{\partial x_1^2}$ *is not and there does not exist a heat kernel of the principal operator.*

By the corollary we can conclude that the principal operator $\partial_t - L_0$ admits a well defined heat kernel, $q_0(t, x, y)$. Indeed, even if the approximating fields are not bounded in general, all the components $\hat{f}_i^j$ depend only on the coordinates $x_1, \ldots, x_{j-1}$ and not on the subsequent coordinates (for example f_i^1 are just constant). This implies that the approximating system satisfies the weaker hypothesis of Remark 2.12 and then it admits a fundamental solution. This is characterized as the density function of the solution, $\xi(t)$, of the stochastic differential equation in Stratonovich form

$$\begin{aligned} d\xi_t &= \hat{f}_0(\xi_t)dt + \sum_{i=1}^{k} \hat{f}_i(\xi_t) \circ dw_i(t) \\ \xi(0) &= x \end{aligned}$$

where $w_i(t)$ is a 1-dim Brownian motion for every $1 \leq i \leq k$. The solution q_0 satisfies the following important homogeneity property.

Lemma 3.19. *For every* $\epsilon > 0$*, for every* $t > 0$ *and for all* $x, y \in \mathbb{R}^n$ *it holds*

$$q_0(t, x, y) = \epsilon^{\mathcal{N}} q_0\left(\epsilon^2 t, \delta_\epsilon x, \delta_\epsilon y\right).$$

In particular, for $\epsilon = 1/\sqrt{t}$ *we have the following identity*

$$q_0(t, x, y) = \frac{1}{t^{\mathcal{N}/2}} q_0\left(1, \delta_{1/\sqrt{t}} x, \delta_{1/\sqrt{t}} y\right) \qquad \forall t > 0, x, y \in \mathbb{R}^n.$$

Proof. This lemma is indeed a corollary of Proposition 3.4, since by definition of $\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_k$ we have

$$\epsilon^2 \delta_{\frac{1}{\epsilon}*} L_0 = L_0.$$

□

We can then apply the procedure introduced in Section 3.2 and we conclude by giving the asymptotics on the diagonal in x_0 of the fundamental solution $p(t, x, y)$.

Theorem 3.20. *Let* x_0 *be a point where the drift field lies in* $\mathcal{D}_{x_0}^2$*. Assume that* $q_0(1, x_0, x_0)$ *is strictly positive. Then the short time asymptotics on the diagonal of the fundamental solution,* $p(t, x, y)$*, of* (3.1) *is given by*

$$p(t, x_0, x_0) = \frac{q_0(1, x_0, x_0)}{t^{\mathcal{N}/2}}(1 + o(1)), \tag{3.29}$$

where $\mathcal{N}$ *is the degree of homogeneity of the volume form* $dx_1 \wedge \ldots \wedge dx_n$ *under the action of the dilations* δ_ϵ *computed in* (3.21)*.*

Proof. Let us write $L = L_0 - (L_0 + L)$. By Duhamel's formula (3.16) the asymptotics on the diagonal of the fundamental solution p for small time to the perturbed operator $\partial_t - L$ is

$$\begin{aligned} p(t,x_0,x_0) &= q_0(t,x_0,x_0) + p*(L_0-L)q_0(t,x_0,x_0) \\ &= \frac{1}{\sqrt{t}^{\mathcal{N}}}\left(q_0(1,x_0,x_0) + \sqrt{t}^{\mathcal{N}} p*(L_0-L)q_0(t,x_0,x_0)\right), \end{aligned} \tag{3.30}$$

provided that the remainder term $\sqrt{t}^{\mathcal{N}} p*(L_0-L)q_0$ is negligible for t small.

To prove this, let us study the function $(L_0 - L)q_0(t,x,y)$:

$$\begin{aligned} |(L_0-L)q_0(t,x,y)| &= \frac{1}{\sqrt{t}^{\mathcal{N}}}|(L_0-L)q_0(1,\delta_{\frac{1}{\sqrt{t}}}x,\delta_{\frac{1}{\sqrt{t}}}y))| \\ &= \frac{1}{\sqrt{t}^{\mathcal{N}}}|\delta_{\frac{1}{\sqrt{t}}*}(L_0-L)q_0(1,x,y)| \\ &\le \frac{1}{\sqrt{t}^{\mathcal{N}}}\left|\frac{C}{\sqrt{t}}(L_0-L)q_0(1,x,y)\right| \end{aligned}$$

for some constant C. The last inequality follows because L_0 is the part of order -2 of L and the difference $L_0 - L$ has order -1. Therefore for small t the operator $\delta_{\frac{1}{\sqrt{t}}*}(L_0 - L)$ rescales at most as $\frac{1}{\sqrt{t}}(L_0 - L)$. Moreover $(L_0 - L)q_0(1,x,y)$ is uniformly bounded in U, because it is evalueted in $t = 1$ where the function is C^∞.

Let us use this relation to prove that $(p*(L_0-L)q_0)(t,x_0,x_0) = O\left(\frac{\sqrt{t}}{\sqrt{t}^{\mathcal{N}}}\right)$. Indeed

$$\begin{aligned} &\lim_{t\searrow 0} t^{\frac{\mathcal{N}-1}{2}}(p*(L_0-L)q_0)(t,x_0,x_0)| \\ &\qquad \le \lim_{t\searrow 0} t^{\frac{\mathcal{N}-1}{2}}\int_0^t\int_U |p(s,x_0,y)(L_0-L)q_0(t-s,y,x_0)|dyds \\ &\qquad \le \lim_{t\searrow 0} t^{\frac{\mathcal{N}-1}{2}}\int_0^t \frac{C}{(t-s)^{\frac{\mathcal{N}+1}{2}}}\int_U p(s,x_0,y)dyds \\ &\qquad \le \lim_{t\searrow 0} t^{\frac{\mathcal{N}-1}{2}}\int_0^t C(t-s)^{-\frac{\mathcal{N}+1}{2}}ds \\ &\qquad = C_2 \end{aligned}$$

for a constant C_2 that comes from the exact value of the limit. Moreover the third inequality is true since p is a fundamental solution and hence has integral ≤ 1.

We have controlled the error, so we can conclude that the desired small time asymptotics on the diagonal is determined by the asymptotics (3.30) and we find

$$p(t,x_0,x_0) = \frac{q_0(1,x_0,x_0)}{t^{\mathcal{N}/2}}(1 + O(\sqrt{t})),$$

which is well defined since by hypothesis the leading term $q_0(1,x_0,x_0)$ doesn't vanish. □

3.6 The principal operator and the associated control system

In this section we are going to investigate the conditions for the positivity of the heat kernel, $q_0(t,x,y)$, that we have introduced in the last section.

Let $f_0, f_1, \dots, f_k$ satisfy Hörmander condition (3.2) and consider the principal operator

$$\frac{\partial}{\partial t} - \hat{f}_0 + \frac{1}{2}\sum_{i=1}^{k} \hat{f}_i^2 \;=:\; \frac{\partial}{\partial t} - L_0 \tag{3.31}$$

defined by the approximating system of the original vector fields. As already pointed out, it admits a smooth fundamental solution given by the probability density, $q_0(t,x,y)$, of the process ξ_t to be at time t in the point y starting from the point x, where ξ_t is the solution of the stochastic differential equation

$$d\xi_t = \hat{f}_0(\xi_t)dt + \sum_{i=1}^{k} \hat{f}_i(\xi_t) \circ dw_i(t). \tag{3.32}$$

In their famous work [47] Stroock and Varadhan characterized the support of $q_0(t,x,y)$ and they showed that it is the set of the reachable points from x of the following associated control problem:

$$\dot{x} = \hat{f}_0(x) + \sum_{i=1}^{k} u_i \hat{f}_i(x), \tag{3.33}$$

where $x : [0,t] \to \mathbb{R}^n$ is a curve in $\mathbb{R}^n$ and $u = (u_1, \dots, u_k) \in L^\infty([0,t];\mathbb{R}^k)$ are bounded controls.

A short proof of Stroock and Varadhan's theorem can be found in [37], while a recent approach by rough paths is given in [26]. For a generalization to the C^α norm see [16].

Unfortunately, Stroock and Varadhan's result holds only for globally bounded vector fields, with bounded derivatives of any order. Since our vector fields are polynomial, they don't satisfy such assumptions and we can not directly apply the result of the support theorem. We will see in a moment how we can adapt their procedure to our system, but first we introduce two simple remarks that will simplify our study.

Remark 3.21. *Let $q_0(t,x,y)$ be the probability density of the solution to equation (3.32). By Lemma 3.19 we have the following equivalence*

$$q_0(1,x_0,x_0) > 0 \quad \Longleftrightarrow \quad q_0(t,x_0,x_0) \quad \forall t > 0.$$

Lemma 3.22. *Consider the control problem (3.33) on $\mathbb{R}^n$. Let $y_1, y_2 \in \mathbb{R}^n$ and $T > 0$ be fixed and assume there exists a curve $y : [0,T] \to \mathbb{R}^n$ that satisfies the control problem (3.33) for some control function $u \in L^\infty$ and such that $y(0) = y_1$ and $y(T) = y_2$.*

Then for any $M > 0$ the curve $x(t) := \delta_M\left(y\left(\frac{t}{M^2}\right)\right)$ is an admissible curve for the control problem defined on $[0, M^2T]$ with control $\tilde{u}(t) := \frac{1}{M}u\left(t/M^2\right)$, that connects $x_1 := \delta_M(y_1)$ with $x_2 := \delta_M(y_2)$ in time M^2T.

In particular, if the control problem (3.33) is controllable in a neighborhood U of x_0 in time T, then it is controllable in $\delta_M(U)$ in time M^2T.

Proof. The boundary conditions are easily satisfied, since $x(0) = \delta_M(y(0)) = \delta_M(y_1) = x_1$ and $x(M^2T) = \delta_M(y(T)) = \delta_M(y_2) = x_2$. Moreover, $x(t)$ is an admissible curve for the control problem with control $\frac{1}{M}u_i\left(t/M^2\right)$, indeed by the homogeneity of the approximating system we have

$$\begin{aligned}
\dot{x}(t) &= \frac{1}{M^2}\delta_{M*}\left(\dot{y}\left(t/M^2\right)\right) \\
&= \frac{1}{M^2}\delta_{M*}\left[\hat{f}_0\left(y\left(t/M^2\right)\right) + \sum_{i=1}^{k} u_i\left(t/M^2\right)\hat{f}_i\left(y\left(t/M^2\right)\right)\right] \\
&= \frac{1}{M^2}\left[M^2\hat{f}_0\left(y\left(t/M^2\right)\right) + \sum_{i=1}^{k} Mu_i\left(t/M^2\right)\hat{f}_i\left(y\left(t/M^2\right)\right)\right] \\
&= \hat{f}_0\left(x(t)\right) + \sum_{i=1}^{k}\frac{1}{M}u_i\left(t/M^2\right)\hat{f}_i\left(x(t)\right).
\end{aligned}$$

□

Our generalization of Stroock and Varadhan's support theorem holds because of the very particular structure of the approximating system. Indeed, since $\hat{f}_0$ has weight 2 and $\hat{f}_i$ has weight 1, for every $1 \leq i \leq k$, every component of these fields depends only on the coordinates of less weight, i.e., $\hat{f}_i^{(j)}$ does not depend on $x_j, x_{j+1}, \ldots, x_n$. This structure is enough to modify the proof of the support theorem in order to prove it in our case.

Definition 3.23. *Consider the control problem* (3.33) *and let us call $x_u(t)$ the solution corresponding to a control u. The* reachable set *of the control problem in time t from x is the set,*

$$\mathcal{A}_t(x) := \left\{y \in \mathbb{R}^n \ : \exists u \in L^\infty([0,t];\mathbb{R}^k) \text{ such that } x_u(0) = x \text{ and } x_u(t) = y\right\}.$$

Proposition 3.24. *Let $X_0, X_1, \ldots, X_k$ be smooth vector fields on $\mathbb{R}^n$, that satisfy the Hörmander condition, and such that every j-th component $X_i^{(j)}$ of X_i, for $0 \leq i \leq k$, does not depend on the coordinates $x_j, \ldots, x_n$, but only on the first coordinates $x_1, \ldots, x_{j-1}$. Let ξ_t be the solution of the stochastic differential equation in Itô form*

$$d\xi_t = \tilde{X}_0(\xi_t)dt + \sum_{i=1}^{k} X_i(\xi_t)dw_i(t)$$

where by the Itô expression of the equation[1]*, the drift field X_0 is changed in $\tilde{X}_0$, which stands for the vector field whose j-component is given by*

$$\tilde{X}_0^{(j)} = X_0^{(j)} + \frac{1}{2}\sum_{i=1}^{k}\sum_{l=1}^{n} X_i^{(l)}\frac{\partial X_i^{(j)}}{\partial x_l}. \tag{3.34}$$

[1]See Section 2.1

Let $p(t,x,y)$ be the probability density of ξ_t to be in y at time t starting from the point x and $\mathcal{A}_t(x)$ be the reachable set at time t from x of the associated control problem

$$\dot{x} = X_0(x) + \sum_{i=1}^{k} u_i(t) X_i(x) \tag{3.35}$$

where $u = (u_1, \ldots, u_k) \in L^\infty([0,t];\mathbb{R}^k)$ is a control function. Then

$$\operatorname{supp}(p(t,x,\cdot)) = \overline{\mathcal{A}_t(x)}.$$

Proof. First of all notice that also the field $\tilde{X}_0$ has the same particular structure as the other fields. Indeed since $X_i^{(j)}$ does not depend on $x_j, \ldots, x_n$, then in (3.34) the sum in l runs only from 1 to $j-1$. Then also $\tilde{X}_0^{(j)}$ depends only on $x_1, \ldots, x_{j-1}$.

Stroock and Varadhan have proved this theorem under the assumption that the fields are Lipschitz and globally bounded, together with their derivatives of first and second order. Following their proof in [47], we have to show that for a dense set of controls u and $\forall \epsilon > 0$

$$P_x(||\xi_t - x_t|| < \epsilon) := P(\,||\xi_t - x_t|| < \epsilon|\, \xi_0 = x) > 0, \tag{3.36}$$

where x_t is the solution of (3.35) starting at x. In particular let us take $\psi \in C^2(\mathbb{R}^+;\mathbb{R}^k)$, with $\psi(0) = 0$, and let x_t be the solution of (3.35) starting at x with control $u_i(t) := \dot{\psi}_i(t)$. Then for all $\epsilon > 0$ we show that

$$P_x\left(||\xi_t - x_t|| < \epsilon \mid\, ||w_t - \psi_t|| < \delta\right) \to 1 \tag{3.37}$$

as $\delta \searrow 0$. This is enough to prove (3.36), since

$$P_x\left(||\xi_t - x_t|| < \epsilon\right) = P_x\left(||\xi_t - x_t|| < \epsilon \mid\, ||w_t - \psi_t|| < \delta\right) \cdot P\left(||w_t - \psi_t|| < \delta\right)$$

and $P\left(||w_t - \psi_t|| < \delta\right) > 0$ for every $\delta > 0$.

Stroock and Varadhan proved (3.37) under the boundedness assumption that we do not have directly, but we will recover it by iterating a conditional probability. Indeed, notice that by our assumption the first component of every vector field, $X_i^{(1)}$, does not depend on any coordinate, so they are actually constant and they trivially satisfy Stroock and Varadhan's assumptions. Then the limit in (3.37) holds for the process $||\xi_t^{(1)} - x_t^{(1)}||$.

Moreover let us assume, by induction, that the first $j-1$ components of ξ_t live in a bounded set. Then the components $X_i^{(j)}$ are Lipschitz and bounded, together with their derivatives of any order, and we can apply Stroock and Varadhan's theorem to the j-th component of ξ_t, then

$$P_x\left(||\xi_t^{(j)} - x_t^{(j)}|| < \epsilon \mid ||w_t - \psi_t|| < \delta, ||\xi_t^{(l)} - x_t^{(l)}|| < \epsilon \;\forall 1 \leq l < j\right) \to 1$$

as $\delta \searrow 0$.

The proof of (3.37) now follows using an iterated conditional probability, indeed in general for every measurable sets $A_1, \dots, A_n, B$, it holds

$$\begin{aligned} P\left(\bigcap_{j=1}^{n} A_j \middle| B\right) &= P\left(\bigcap_{j=2}^{n} A_j \middle| B \cap A_1\right) P\left(A_1 | B\right) \\ &= P\left(\bigcap_{j=3}^{n} A_j \middle| B \cap A_1 \cap A_2\right) P\left(A_2 | B \cap A_1\right) P\left(A_1 | B\right) \\ &\vdots \\ &= \prod_{j=1}^{n} P\left(A_j \middle| B \cap \bigcap_{l=1}^{j-1} A_l\right). \end{aligned}$$

Then

$$\begin{aligned} &P\left(||\xi_t - x_t|| < \epsilon \mid ||w_y - \psi_t|| < \delta\right) = \\ &= \prod_{j=1}^{n} P\left(||\xi_t^{(j)} - x_t^{(j)}|| < \epsilon \mid ||w_t - \psi_t|| < \delta, ||\xi_t^{(l)} - x_t^{(l)}|| < \epsilon \; \forall 1 \le l < j\right) \to 1 \end{aligned}$$

as $\delta \searrow 0$ and we have proved (3.37) in our case. □

We are now ready to show a condition for the positivity of the fundamental solution of the approximating differential operator.

Theorem 3.25. *Let $q_0(t,x,y)$ be the fundamental solution of* (3.31). *If the reachable set $\mathcal{A}_t(x_0)$ of the associated control problem* (3.33) *is a neighborhood of x_0 for some $t > 0$, then $q_0(1, x_0, x_0) > 0$.*

Proof. By Remark 3.21 it is enough to prove that $q_0(T, x_0, x_0) > 0$ for some $T > 0$. We will choose $T = 2t$. Moreover, by Lemma 3.22, if $\mathcal{A}_t(x_0)$ is a neighborhood of x_0 for some $t > 0$, it is a neighborhood for every $t > 0$.

Assume by contradiction that $q_0(2t, x_0, x_0) = 0$. By Chapman-Kolmogorov equation we know that

$$0 = q_0(2t, x_0, x_0) = \int_{\mathbb{R}^n} q_0(t, x_0, y) q_0(t, y, x_0) dy = \int_{\overline{\mathcal{A}_t(x_0)}} q_0(t, x_0, y) q_0(t, y, x_0) dy,$$

where we can restrict the space of integration, since by Lemma 3.24, $\mathrm{supp}(q_0(t, x_0, \cdot)) = \overline{\mathcal{A}_t(x_0)}$. Then for all $y \in \overline{\mathcal{A}_t(x_0)}$ we have $q_0(t, y, x_0) = 0$. As shown in Section 2.2.1, the function $\tilde{q}(t, x, y) := q_0(t, y, x)$ is the fundamental solution of the adjoint operator to L_0, that is

$$L_0^* = -\hat{f}_0 + \frac{1}{2}\hat{f}_i^2.$$

Then $\tilde{q}(t, x, y)$ is the probability density function of the stochastic process $\tilde{\xi}_t$ solution of the stochastic equation

$$\begin{cases} d\tilde{\xi}_t = -\hat{f}_0(\tilde{\xi}_t)dt + \sum_{i=1}^{k} \hat{f}_i(\tilde{\xi}_t) \circ dw_i(t) \\ \tilde{\xi}_0 = x. \end{cases}$$

By contradiction we have assumed that $\tilde{q}(2t, x_0, x_0) = 0$, then again by Chapman-Kolmogorov equation we have

$$0 = \tilde{q}(2t, x_0, x_0) = \int_{\tilde{\mathcal{A}}_t(x_0)} \tilde{q}(t, x_0, y)\tilde{q}(t, y, x_0)dy$$

where $\tilde{\mathcal{A}}_t(x_0)$ is the reachable set in time t from the point x_0 of the associated control problem

$$\dot{x} = -\hat{f}_0(x) + \sum_{i=1}^{k} u_i(t)\hat{f}_i(x).$$

It follows that $q_0(t, x_0, y) = \tilde{q}(t, y, x_0) = 0$ for all $y \in \tilde{\mathcal{A}}_t(x_0)$. Since the point x_0 is a stationary point for the control problem, then $x_0 \in \tilde{\mathcal{A}}_t(x_0)$, for every $t > 0$. By Krener's theorem (see Chapter 8 in [4]), x_0 is in the closure of $\mathrm{int}(\tilde{\mathcal{A}}_t(x_0))$, then $\mathcal{A}_t(x_0) \cap \tilde{\mathcal{A}}_t(x_0)$ has non zero measure and $q_0(t, x_0, z) = 0$ for all z in this intersection. This is a contradiction to the support theorem.

We conclude that, if $\mathcal{A}_t(x_0)$ is a neighborhood of x_0, then $q_0(t, x_0, x_0) > 0$ for all $t > 0$. □

Remark 3.26. *In view of Theorem 3.20 and Theorem 3.25 we can conclude the following properties about the asymptotics of the fundamental solution p of the operator* (3.1).

(i) If the control problem associated with the original system:

$$\dot{x} = f_0(x) + \sum_{i=1}^{k} u_i(t)f_i(x) \tag{3.38}$$

is not controllable around x_0, that is $\mathcal{A}_t(x_0)$ is not a neighborhood of x_0, then

$$p(t, x_0, x_0) = 0 \qquad \forall t > 0.$$

Indeed by the support theorem $\mathrm{supp}(p(t, x_0, \cdot)) = \overline{\mathcal{A}_t(x_0)}$, *therefore x_0 is on the boundary of the support. Since $p(t, x_0, \cdot)$ is smooth, the conclusion follows.*

(ii) If the control problem (3.38) *is controllable in x_0, then we study the controllability of its nilpotent approximation, defined by the fields $\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_k$ introduced in Section 3.4.*

(ii.1) If the approximating control problem

$$\dot{x} = \hat{f}_0(x) + \sum_{i=1}^{k} u_i(t)\hat{f}_i(x) \tag{3.39}$$

is controllable around x_0, then the asymptotics is given in Theorem 3.20, where we see that the fundamental solution on the diagonal in x_0 blows up for small t as the rational polynomial $\frac{c_0}{t^{\mathcal{N}/2}}$, for a positive constant c_0 depending on the chosen volume and on the approximating system (3.39). *The order $\mathcal{N}$ is determined by the Lie algebra generated by the fields $f_0, f_1, \ldots, f_k$ at x_0 as explained in formula* (3.21).

(ii.2) If the approximating control problem (3.39) *is not controllable, then* $p(t, x_0, x_0)$ *goes to infinity for small t faster than* $\frac{1}{t^{\mathcal{N}/2}}$ *as shown in the following Proposition. In [17] the authors show an example, where the asymptotics goes to infinity even exponentially fast.*

Proposition 3.27. *Assume that the control problem* (3.38) *is controllable around* x_0*, but the approximating control problem* (3.39) *is not. If* $p(t, x_0, x_0) > 0$*, then* $p(t, x_0, x_0)$ *goes to infinity faster then* $\frac{c}{t^{\mathcal{N}}}$*, where* $\mathcal{N}$ *is defined in* (3.21).

Proof. To show this we need to prove that all the coefficients in the asymptotics (3.29) are zero. The asymptotics was found introducing the fundamental solution q_0 and by using Duhamel's formula (3.16). By iterating it we can achieve a better approximation of the asymptotics and find the higher coefficients. Indeed,we obtain

$$p = q_0 + \sum_{i=1}^{j} q_0 \left(*\mathcal{X} q_0\right)^i + p \left(*\mathcal{X} q_0\right)^{j+1}$$

for every $j \in \mathbb{N}$, where $\mathcal{X} := L_0 - L$ and $\left(*\mathcal{X} q_0\right)^i$ means that we iterate the convolution i times. We have to show that all the terms $q_0 \left(*\mathcal{X} q_0\right)^i$ vanish at the point $(1, x_0, x_0)$ for every $i \in \mathbb{N}$.

By Lemma 3.24 we know that $q_0(t, x_0, y) = 0$ for every $y \in \overline{\mathcal{A}_t(x_0)}^c$ and all $t > 0$. Let us assume by induction that for an $i \in \mathbb{N}$

$$q_0 \left(*\mathcal{X} q_0\right)^i (t, x_0, y) = 0 \qquad \forall y \in \overline{\mathcal{A}_t(x_0)}^c,\ \forall t > 0. \tag{3.40}$$

Then for every $y \in \overline{\mathcal{A}_t(x_0)}^c$ we have

$$q_0 \left(*\mathcal{X} q_0\right)^{i+1} (t, x_0, y) = \int_0^t \int_{\overline{\mathcal{A}_s(x_0)}} q_0 \left(*\mathcal{X} q_0\right)^i (s, x_0, z)\ \mathcal{X} q_0(t-s, z, y) dz ds,$$

where the integral can be computed just on $\overline{\mathcal{A}_s(x_0)}$ by the induction hypothesis. But $q_0(t-s, z, y) \equiv 0$ on $\overline{\mathcal{A}_s(x_0)}$ by Chapman-Kolmogorov equation, then also all its derivatives vanish there. Then the integral is zero and we have proved (3.40) for every i.

Since q_0 is smooth and x_0 is on the boundary of $\overline{\mathcal{A}_t(x_0)}^c$, then equation (3.40) holds also in the point (t, x_0, x_0) for every $t > 0$, that means that all the coefficients of the asymptotics (3.29) are zero. □

3.7 Examples

We end this chapter with a study of some known examples, to understand better the meaning of Theorem 3.20.

Example 5 (Continuation of Example 1)**.** *We complete the study of Example 1. We have already computed the principal part of the operator* $\partial_t - L$*, that is*

$$\frac{\partial}{\partial t} - \hat{f}_0 - \frac{1}{2}\hat{f}_1^2 = \frac{\partial}{\partial t} - x_1^2 \frac{\partial}{\partial x_2} - \frac{1}{2}\frac{\partial^2}{\partial x_1^2}.$$

This operator is indeed hypoelliptic since $\hat{f}_1 = \frac{\partial}{\partial x_1}$ *and* $[\hat{f}_1, [\hat{f}_1, \hat{f}_0]] = 2\frac{\partial}{\partial x_2}$ *span the whole tangent space in every point. Let* $q_0(t, x, y)$ *be the density function of the solution* $\xi(t)$ *of the stochastic equation* (3.28)*, that in coordinates is given by*

$$d\xi_1 = dw_1, \qquad d\xi_2 = x_1^2 dt.$$

We can see that the second coordinate is actually deterministic and has positive derivative, so it can only increase. Consequently, if a path starts in $x_0 = (0,0)$ *the solution* $\xi(t)$ *will almost surely never come back to* x_0 *again, indeed* x_0 *is on the boundary of the support of* $q_0(t, x_0, \cdot)$*. Then the hypothesis of Theorem 3.20 that requires* $q_0(1, x_0, x_0) > 0$ *is not fulfilled.*

Since the original control problem is controllable, this is an example of type (ii.2).

Example 6 (Sub-Riemannian manifold: continuation of Example 3)**.** *The study of the asymptotics on the diagonal of the heat kernel on 3D contact sub-Riemannian manifolds has already been performed by Barilari in [6] and more in general by Ben Arous in [15]. The nilpotent approximation of a sub-Riemannian manifold of dimension 3 is isometric to the Heisenberg group. Let us represent the Heisenberg group as* $\mathbb{R}^3$ *with coordinates* (x, y, z)*, then the approximating system can be written as*

$$\hat{f}_1 = \frac{\partial}{\partial x} + \frac{y}{2}\frac{\partial}{\partial z} \qquad \textit{and} \qquad \hat{f}_2 = \frac{\partial}{\partial y} - \frac{x}{2}\frac{\partial}{\partial z}$$

As one can easily verify, the order of homogeneity of the volume form is given by $\mathcal{N} = 4$*, as computed also with the general formula* (3.21)*.*

The principal part L_0 *of the operator is hypoelliptic and symmetric, the associated control problem is then controllable, so there exists a well-defined symmetric heat kernel, that is positive for every* $t > 0$ *as seen in Theorem 3.25. The hypothesis of Theorem 3.20 are then fulfilled and we find that the asymptotics on the diagonal of the original heat kernel* $p(t, x, y)$ *has the following order:*

$$p(t, x, x) = \frac{a_0(x) + o(1)}{\sqrt{t}^4} \qquad \textit{for a smooth function } a_0(x) > 0 \textit{ on the manifold.}$$

This was the same order found by Ben Arous in [15] and by Barilari in [6].

This example is of type (ii.1).

Example 7 (Ben Arous and Léandre)**.** *We consider here an example studied by Ben Arous and Léandre in [17]. Consider the space* $\mathbb{R}^2$ *with coordinates* (x_1, x_2) *and let*

$$f_0 = x_1^a \frac{\partial}{\partial x_2}, \qquad f_1 = \frac{\partial}{\partial x_1}, \qquad f_2 = x_1^b \frac{\partial}{\partial x_2},$$

where a *and* b *are positive integers. Then* $x_0 = (0, x_2)$ *is a stationary point of the drift field for any* $x_2 \in \mathbb{R}$*. The operator* $L = f_0 + \frac{1}{2}(f_1^2 + f_2^2)$ *satisfies even the strong Hörmander condition, i.e. the fields* f_1, f_2 *alone are Lie bracket generating.*

In [18], it is shown the complete behavior of the heat kernel $p(t, x, y)$ *of this operator on the diagonal. We summarize here the most interesting properties for our study:*

Theorem 3.28. *1. If $b \le a+1$, then there exists a constant $K(a,b) > 0$ such that*

$$p(t, x_0, x_0) \sim \frac{K(a,b)}{\sqrt{t}^{b+2}}. \tag{3.41}$$

2. If $b > a+1$ and a is even, then the fundamental solution $p(t, x_0, x_0)$ decreases with exponential velocity.

The results found in this paper agree with the statement of the theorem. Indeed, let us consider the filtration $\mathcal{G}$ given by the vector fields $f_0, f_1, \ldots, f_k$ at a point $x_0 = (0, x_2)$. Let $m := \min\{a+2; b+1\}$. It is easy to verify that the subspaces of the filtration $\mathcal{G}$ are

$$G_i(x_0) = \begin{cases} \mathbb{R} \times \{0\} & \text{if } 1 \le i < m \\ \mathbb{R}^2 & \text{if } i = m. \end{cases}$$

Accordingly the coordinate x_1 has weight 1, while the coordinate x_2 has weight m and the order of homogeneity of the volume form is given by

$$\mathcal{N} = 1 + m = \begin{cases} b+2 & \text{if } b \le a+1 \\ a+3 & \text{if } b \ge a+1. \end{cases}$$

To determine the nilpotent approximation, it is convenient to divide the study in 3 cases, depending on the value of a w.r.t. b.

If $b < a+1$, then $m = b+1$. The nilpotent approximation is obtained by taking the Taylor expansion of the field f_0 of order 2 and the Taylor expansion of order 1 of f_1 and f_2. Then we find that $\hat{f}_1 = f_1$, $\hat{f}_2 = f_2$ and $\hat{f}_0 = 0$. The principal part L_0 of the operator L is $\frac{1}{2}(\hat{f}_1^2 + \hat{f}_2^2)$, that is hypoelliptic. Then there exists a well-defined heat kernel, $q_0(t,x,y)$, and, since the associated control system is controllable, $q_0(t, x_0, x_0) > 0$ for every $t > 0$. Then the hypothesis of Theorem 3.20 are fulfilled and we find that the small time asymptotics of the fundamental solution of L has order $\mathcal{N}/2 = \frac{b+2}{2}$, that is exactly the one given in (3.41).

If $b = a+1$, then $m = b+1$ and the nilpotent approximation is equal to the fields f_1, f_2, f_0 themselves. Since f_1, f_2 are Lie bracket generating the associated control system

$$\dot{x} = f_0 + u_1 f_1 + u_x f_2$$

is still controllable, then the heat kernel $q_0(t, x_0, x_0)$ is positive for every $t > 0$, and we obtain again the statement (3.41).

If $b > a+1$, then $\hat{f}_1 = f_1$, $\hat{f}_0 = f_0$ and $\hat{f}_2 = 0$. The principal operator $\hat{f}_0 + \frac{1}{2}\hat{f}_1^2$ is still hypoelliptic, but if a is even, then the heat kernel q_0 is zero in $x_0 = (0, x_2)$ for any $t > 0$. This is because a.e. path starting from x_0 will never come back to x_0 again, since the drift f_0 makes the first coordinate increase, if x_1 becomes different from 0. Then we can not apply Theorem 3.20 and indeed Ben Arous and Léandre have shown an exponential decrease in this case.

Chapter 4

Curvature terms in small time heat kernel expansion for a model class of hypoelliptic Hörmander operators

In this chapter, which is based on the results of [9], we perform the first step in the characterization from a geometric viewpoint of the coefficients of the heat kernel asymptotics on the diagonal. In particular we focus on the model case of linear hypoelliptic operators on $\mathbb{R}^n$ of the form

$$L = \sum_{j=1}^{n} (Ax)_j \frac{\partial}{\partial x_j} + \frac{1}{2} \sum_{j,h=1}^{n} (BB^*)_{jh} \frac{\partial^2}{\partial x_j \partial x_h},$$

where $A = (a_{jh})$ and $B = (b_{ij})$ are respectively $n \times n$ and $n \times k$ constant matrices, that satisfy Hörmander condition of hypoellipticity. In this setting it reduces to the assumption that

$$\mathrm{rk}[B, AB, A^2B, \ldots, A^{m-1}B] = n. \tag{4.1}$$

As explained in Section 2.4, the heat equation associated to L admits a smooth fundamental solution $p(t, x, y) \in C^\infty(\mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^n)$ that can be computed explicitly as follows

$$p(t, x, y) = \frac{e^{-\frac{1}{2}(y - e^{tA}x)^* D_t^{-1} (y - e^{tA}x)}}{(2\pi)^{n/2} \sqrt{\det D_t}},$$

where

$$D_t = e^{tA} \left(\int_0^t e^{-\tau A} BB^* e^{-\tau A^*} d\tau \right) e^{tA^*}.$$

By condition (4.1), the matrix D_t is invertible for every $t > 0$.

These operators are the simplest class of hypoelliptic, but not elliptic, operators satisfying (wHC) and are classical in the literature, starting from the pioneering work of Hörmander [29] (see also [32] for a detailed discussion on this class of operators). As already pointed out by Stroock and Varadhan [47] in the study of the support of the diffusion, the properties of ξ_t are strongly related with the solutions to the control problem

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^k. \tag{4.2}$$

The relation between the solution of the heat equation and the associated control problem is classical, see for instance [21]. In this chapter we further investigate this relation and we show that the study of the minimizers of the *cost functional*

$$J_T(u) = \frac{1}{2}\int_0^T |u(s)|^2 ds$$

reveals some geometric-like properties of the heat kernel. In other words, for every fixed $x_1, x_2 \in \mathbb{R}^n$ and $T > 0$, one is interested in computing

$$S_T(x_1, x_2) := \inf\{J_T(u) : u \in L^\infty([0,T]; \mathbb{R}^k), x_u(0) = x_1, x_u(T) = x_2\},$$

where $x_u(\cdot)$ is the solution of (4.2) associated with the control u. The condition (4.1) (also known as *Kalman condition*) ensures that the control system (4.2) is controllable, i.e., $S_T(x_1, x_2) < +\infty$ for all $x_1, x_2 \in \mathbb{R}^n$ and $T > 0$.

For $x_0 \in \mathbb{R}^n$ fixed, let $x_{\bar{u}}(t)$ be an optimal trajectory starting at x_0, i.e., a minimizer of the cost functional. The *geodesic cost* associated with $x_{\bar{u}}$ is the family of functions

$$c_t(x) := -S_t(x, x_{\bar{u}}(t)) \qquad \text{for } t > 0, x \in \mathbb{R}^n.$$

From the asymptotics of the second derivative of c_t, one can highlight some "curvature-like" invariants of the cost, which define a family of symmetric operators

$$\mathcal{I} : \mathbb{R}^k \to \mathbb{R}^k, \qquad \mathcal{Q}^{(i)} : \mathbb{R}^k \to \mathbb{R}^k, \qquad i \geq 0.$$

These operators, that are in principle associated with an optimal trajectory, in the case of a linear-quadratic optimal control problem are constant.

The operator $\mathcal{I}$ is connected to the flag generated by the brackets along the optimal trajectory. The operators $\mathcal{Q}^{(i)}$ play the role of curvature invariants for the optimal control problem (see Chapter 4.3 and [2] for more details).

To state the main results let us introduce the filtration $E_1 \subset E_2 \subset \ldots \subset E_m = \mathbb{R}^n$ as

$$E_i = \text{span}\{A^j Bx \,|\, x \in \mathbb{R}^k,\ 0 \leq j \leq i-1\}. \tag{4.4}$$

This is the linear counterpart to the filtration $\mathcal{G}$ of the previous chapter (see Eq. (3.5)), indeed $E_i = G_{2i-1}(x_0) = G_{2i}(x_0)$. In particular, the operator $\mathcal{I}$ is connected with the order of the asymptotics, $\mathcal{N}$, indeed

$$\mathcal{N} = \text{tr}(\mathcal{I}) = \sum_{i=1}^{m}(2i-1)(\dim E_i - \dim E_{i-1}). \tag{4.5}$$

We prove the following results: when x_0 is an equilibrium of the drift field, we can compute and characterize all the coefficients in the small time asymptotic expansion, providing a characterization of the coefficients that is analogous to the one obtained on a Riemannian manifold, see (1.1).

Theorem 4.1. *Assume that* $Ax_0 = 0$. *Then*

$$p(t, x_0, x_0) = \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}} \left(\sum_{i=0}^{m} a_i t^i + O(t^m) \right), \qquad \text{for } t \to 0,$$

where $\mathcal{N} = \mathrm{tr}(\mathcal{I})$ *is defined in* (4.5) *and* c_0 *is a positive constant. Moreover there exist universal polynomials* P_i *of degree* i *such that*

$$a_i = P_i(\mathrm{tr}A, \mathrm{tr}\mathcal{Q}^{(0)}, \dots, \mathrm{tr}\mathcal{Q}^{(i-2)}).$$

In particular for $i = 1, 2, 3$, *we have*

$$a_1 = -\frac{\mathrm{tr}A}{2}, \qquad a_2 = \frac{(\mathrm{tr}A)^2}{8} + \frac{\mathrm{tr}\mathcal{Q}^{(0)}}{4},$$

$$a_3 = -\frac{\mathrm{tr}\mathcal{Q}^{(1)}}{12} - \frac{\mathrm{tr}A\,\mathrm{tr}\mathcal{Q}^{(0)}}{8} - \frac{(\mathrm{tr}A)^3}{48}.$$

We stress that the explicit structure of any higher order coefficient can be a priori computed by a simple Taylor expansion, as it follows from the proof, cf. Section 4.4.

More in general, one has an expansion of $p(t, x, y)$, at every pair of points x, y, relating the heat kernel with the optimal cost functional and the same geometric coefficients of Theorem 4.1.

Corollary 4.2. *For any pair of points,* $x, y \in \mathbb{R}^n$,

$$p(t, x, y) = \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}} e^{-S_t(x,y)} \left(\sum_{i=0}^{m} a_i t^i + O(t^m) \right), \qquad \text{for } t \to 0,$$

where the coefficients a_i *are characterized as in Theorem 4.1.*

This corollary is a direct consequence of the previous theorem and is proved in Section 4.4.2.

Next we consider the case when x_0 is not a zero of the drift field. In this case one can observe different behaviors depending on the smallest level of the filtration (4.4) to which the vector Ax_0 belongs. Indeed the cost of the constant trajectory x_0 is strictly positive and the asymptotics depends on the exponential term appearing in Corollary 4.2.

Theorem 4.3. *Assume that* $Ax_0 \neq 0$. *Then*

(i) if $Ax_0 \in E_1$, *we have the polynomial decay*

$$p(t, x_0, x_0) = \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}} \left[1 - \left(\frac{\mathrm{tr}A}{2} + \frac{|Ax_0|^2}{2} \right) t + O(t^2) \right], \qquad \text{for } t \to 0;$$

(ii) if $Ax_0 \in E_i \setminus E_{i-1}$ *for some* $i > 1$, *then* $p(t, x_0, x_0)$ *has exponential decay to zero. More precisely there exists* $C > 0$ *such that*

$$p(t, x_0, x_0) = \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}} \exp\left(-\frac{C + O(t)}{t^{2i-3}} \right), \qquad \text{for } t \to 0.$$

We stress that claim (i) is the analogue of the Riemannian expansion (1.2): here no scalar curvature appears since all brackets of horizontal fields are zero. For the same reason, in claim (ii), the condition $Ax_0 \notin E_1$ means that the drift field is not in $\mathcal{D}^2$, and indeed we observe an exponential decay as already experienced by Ben Arous and Léandre [17, 18] (cf. Chapter 1). The fact that the order of decay of the exponential could be faster then $\frac{1}{t}$ in this case is not in contrast with their result, since here the strong Hörmander condition does not hold.

4.1 Linear quadratic optimal control problems

Let us consider the optimal control problem associated with the operator L

$$\begin{cases} \dot{x} = Ax + Bu \\ J_T(u) = \frac{1}{2}\int_0^T \sum_{i=1}^k |u_i(s)|^2 ds \to \min \end{cases} \tag{4.6}$$

Here $u \in L^\infty([0,T];\mathbb{R}^k)$ is the control and J_T is the optimal cost to be minimized. A curve $x(t) \in \mathbb{R}^n$ is called *admissible* for the control problem (4.6) if there exists a control function $u \in L^\infty([0,T];\mathbb{R}^k)$ such that $\dot{x}(t) = Ax(t) + Bu(t)$ for a.e. $t \in [0,T]$.

The solution of the differential equation (4.6) corresponding to the control u will be denoted by $x_u : [0,T] \to \mathbb{R}^n$ and for a fixed initial point $x_1 \in \mathbb{R}^n$ is given by:

$$x_u(t) = e^{tA}x_1 + e^{tA}\int_0^t e^{-\tau A}Bu(\tau)d\tau. \tag{4.7}$$

Among all trajectories x_u starting at x_1 and arriving in a point $x_2 \in \mathbb{R}^n$ in time T we want to minimize the cost functional J_T: for every fixed $x_1, x_2 \in \mathbb{R}^n$ and $T > 0$, we define the *optimal cost*

$$S_T(x_1,x_2) := \inf\{J_T(u) | u \in L^\infty([0,T];\mathbb{R}^k), x_u(0) = x_1, x_u(T) = x_2\}. \tag{4.8}$$

A control $\bar{u}$ that realizes the minimum in (4.8) is called an *optimal control*, and the corresponding trajectory $x_{\bar{u}} : [0,T] \to \mathbb{R}^n$ is called an *optimal trajectory* of the control problem (4.6).

It is well-known (see for example [4]) that the optimal trajectories of the control problem (4.6) can be obtained as the projection of the extremals of an Hamiltonian flow in $T^*\mathbb{R}^n$. Namely, let

$$H(p,x) = p^*Ax + \frac{1}{2}p^*BB^*p \qquad \forall\,(p,x) \in T^*\mathbb{R}^n$$

be the Hamiltonian function associated with the optimal control problem. All the optimal trajectories are the projection $x(t)$ of the solution $(p(t), x(t)) \in T^*\mathbb{R}^n \cong \mathbb{R}^{2n}$ of the Hamiltonian system associated with H

$$\begin{cases} \dot{p} = -A^*p \\ \dot{x} = Ax + BB^*p. \end{cases}$$

Moreover, the control realizing the optimal trajectory is uniquely recovered by $\bar{u}(t) = B^*p(t)$. Thus the solution corresponding to the initial condition $(p_0, x_0) \in T^*_{x_0}\mathbb{R}^n$ can be found explicitly

$$\begin{cases} p(t) = e^{-tA^*}p_0 \\ x(t) = e^{tA}\left(x_0 + \int_0^t e^{-\tau A}BB^*e^{-\tau A^*}d\tau p_0\right). \end{cases} \tag{4.9}$$

Let us denote by Γ_t the matrix

$$\Gamma_t := \int_0^t e^{-\tau A}BB^*e^{-\tau A^*}d\tau. \tag{4.10}$$

By Kalman's condition (4.1), it follows that Γ_t is invertible for every $t > 0$.

Remark 4.4. *Fix $x_1, x_2 \in \mathbb{R}^n$ and $T > 0$. By the explicit formulas* (4.9) *there exists a unique initial covector p_0 such that the corresponding extremal $x(t)$ satisfies $x(0) = x_1$ and $x(T) = x_2$. It is equal to*

$$p_0 = \Gamma_T^{-1}\left(e^{-TA}x_2 - x_1\right).$$

Since the optimal control is given by $\bar{u}(t) = B^*p(t)$, we can also write the optimal cost to go from x_1 to x_2, namely

$$S_T(x_1, x_2) = \frac{1}{2}p_0^*\Gamma_T p_0.$$

It follows that the cost function is smooth in $(T, x_1, x_2) \in \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^n$, where $\mathbb{R}^+ = \{T \in \mathbb{R} : T > 0\}$.

4.2 The flag and growth vector of an admissible curve

Let $x_u : [0, T] \to \mathbb{R}^n$ be an admissible curve such that $x_u(0) = x_0$, associated with the control u. Let $P_{0,t}$ be the flow defined by u, i.e., for every $y \in \mathbb{R}^n$

$$P_{0,t}(y) := x_u(t; y) \quad \text{s.t.} \quad x_u(0; y) = y.$$

At any point of $\mathbb{R}^n$ we split the tangent space $T_x\mathbb{R}^n \cong \mathbb{R}^n = D \oplus D^\perp$, where D is the k-dimensional subspace generated by the columns of B and $D^\perp$ is its orthogonal complement, and we define the following family of subspaces of $T_{x_0}\mathbb{R}^n$:

$$\mathcal{F}_{x_u}(t) := (P_{0,t})_*^{-1}D \subset T_{x_0}\mathbb{R}^n.$$

In other words, the family $\mathcal{F}_{x_u}(t)$ is obtained by translating in $T_{x_0}\mathbb{R}^n$ the subspace D along the trajectory x_u using the flow $P_{0,t}$.

Definition 4.5. *The* flag of the admissible curve $x_u(t)$ *is the sequence of subspaces*

$$\mathcal{F}^i_{x_u}(t) := \operatorname{span}\left\{ \frac{d^j}{dt^j}v(t) \,\middle|\, v(t) \in \mathcal{F}_{x_u}(t) \text{ smooth, } j \leq i-1 \right\} \subset T_{x_0}\mathbb{R}^n, \quad i \geq 1.$$

By construction, this is a filtration of $T_{x_0}\mathbb{R}^n$, i.e., $\mathcal{F}^i_{x_u}(t) \subset \mathcal{F}^{i+1}_{x_u}(t)$, for all $i \geq 1$.

Definition 4.6. *Denote by* $k_i(t) := \dim \mathcal{F}^i_{x_u}(t)$*. The* growth vector *of the admissible* curve $x_u(t)$ *is the sequence of integers*

$$\mathcal{G}_{x_u}(t) = \{k_1(t), k_2(t), \dots\}.$$

Remark 4.7. *For any* $s \in [0,T]$ *we can define the family of subspace* $\mathcal{F}_{x_{u,s}}(t)$ *associated with the admissible curve* x_u *starting at time* s*. Namely, let* $P_{s,t}$ *be the flow defined by* u *starting at time* $s < t$*, i.e., for every* $y \in \mathbb{R}^n$

$$P_{s,t}(y) := x_u(t;y) \qquad s.t.\ x_u(s;y) = y,$$

and let $x_{u,s}(t) := x_u(s+t)$ *be the shifted curve by time* s*. Then we introduce the family of subspaces* $\mathcal{F}_{x_{u,s}}(t) := (P_{s,s+t})_*^{-1}\mathbb{R}^k$ *with base point* $x_u(s)$*.*

The relation $\mathcal{F}_{x_{u,s}}(t) = (P_{0,s})_*\mathcal{F}_{x_u}(s+t)$ *implies that the growth vector of the original curve at time* t *can be equivalently computed via the growth vector at time* 0 *of the shifted curve* $x_{u,t}$*, i.e.,* $k_i(t) = \dim \mathcal{F}^i_{x_{u,t}}(0)$ *and* $\mathcal{G}_{x_u}(t) = \mathcal{G}_{x_{u,t}}(0)$*.*

The growth vector of a curve x_u at time 0 can be easily computed. Indeed, by the explicit expression (4.7) of the flow of u, the flag of the curve x_u starting from a point x_0 is

$$\mathcal{F}^i_{x_u}(0) = \operatorname{span}\{B, AB, \dots, A^{i-1}B\}.$$

In particular the flag at time 0 is independent on the control u and the initial point x_0. By Remark 4.7 the growth vector of any curve x_u is then independent also from the time and is equal to $\mathcal{G}_{x_u}(t) = \{k_1, k_2, \dots\}$, where the indices k_i are

$$k_i := \dim \operatorname{span}\{B, AB, \dots, A^{i-1}B\}.$$

By Kalman's condition (4.1) there exists a minimal integer $1 \leq m \leq n$ such that $k_m = n$. We call such m the *step* of the control problem (independent of the admissible curves). Moreover notice that $k_1 = k$.

Lemma 4.8. *Let* $d_i := k_i - k_{i-1}$*. Then* $d_1 \geq d_2 \geq \dots \geq d_m$*.*

Proof. The linear map $\widehat{A} : \mathcal{F}^i_{x_u}(t) \to \mathcal{F}^{i+1}_{x_u}(t)/\mathcal{F}^i_{x_u}(t)$ defined by $\widehat{A}v := Av$ for every $v \in \mathcal{F}^i_{x_u}(t)$ is surjective and $\operatorname{Ker}\widehat{A} = \mathcal{F}^{i-1}_{x_u}(t)$. Then

$$\dim \mathcal{F}^i_{x_u}(t) - \dim \mathcal{F}^{i-1}_{x_u}(t) \geq \dim \mathcal{F}^{i+1}_{x_u}(t) - \dim \mathcal{F}^i_{x_u}(t),$$

which concludes the proof. □

To any curve x_u we associate a tableau with m columns of length d_i for $i = 1, \dots, m$. By the previous Lemma, the height of the columns is decreasing from left to right. We call n_j the length of the j−th row, for $j = 1, \dots, k$ (for example $n_1 = m$, see Figure 4.1).

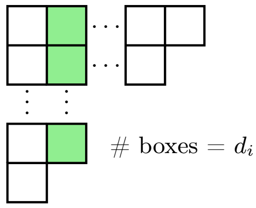


Figure 4.1: Young diagram

4.3 Geodesic cost and curvature invariants

Definition 4.9. *Let $x_0 \in \mathbb{R}^n$ and fix an optimal trajectory $x_{\bar{u}} \in \mathbb{R}^n$ starting at x_0 of the optimal control problem* (4.6). *The* geodesic cost *associated with $x_{\bar{u}}$ is the family of functions $\{c_t\}_{t>0}$ defined by*

$$c_t(x) := -S_t(x, x_{\bar{u}}(t)) \qquad x \in \mathbb{R}^n,$$

where S_t is the optimal cost function defined in (4.8).

Notice that thanks to Remark 4.4 and the smoothness of optimal trajectories, the geodesic cost is smooth in $\mathbb{R}^+ \times \mathbb{R}^n$.

Moreover, for any $t > 0$ and $x \in \mathbb{R}^n$ there exists a unique minimizer of the cost functional among all the trajectories that connect x with $x_{\bar{u}}(t)$ in time t. By the explicit expressions of the extremals in (4.9) and of the optimal control $\bar{u}$, we can write the explicit formula:

$$c_t(x) = -\frac{1}{2}p_0^*\Gamma_t p_0 + p_0^*(x - x_0) - \frac{1}{2}(x - x_0)^*\Gamma_t^{-1}(x - x_0),$$

where $x_{\bar{u}}$ is a solution of the Hamiltonian system with initial data (p_0, x_0) and Γ_t is the invertible matrix defined in (4.10).

Then we define the following family of quadratic forms, $\mathcal{Q}(t)$, on $\mathbb{R}^k$:

$$\mathcal{Q}(t) := B^* \left(d^2_{x_0}\dot{c}_t\right) B = -\frac{d}{dt}B^*\Gamma_t^{-1}B. \tag{4.11}$$

This family of operators is the linear quadratic counterpart of the more general family of operators introduced in [2, Chapter 4] for the wider class of non linear control systems that are affine in the control.

Remark 4.10. *Actually the family of operators $\mathcal{Q}(t)$ does not depend on the initial data (p_0, x_0) of the optimal trajectory and is the same for any geodesic. This is saying that it is an intrinsic object of the pair control system and cost.*

Moreover, let us stress that $\mathcal{Q}(t)$ is an intrinsic object of the optimal control problem, i.e., it does not depend on the chosen coordinate on $\mathbb{R}^n$.

Indeed let $y = Cx$, with C an $n \times n$ invertible matrix, be a linear change of coordinates on $\mathbb{R}^n$. In the new coordinates the dynamical system (4.6) *is rewritten as*

$$\dot{y} = \tilde{A}y + \tilde{B}u$$

where $\tilde{A} = CAC^{-1}$ *and* $\tilde{B} = CB$. *Since the matrix* C *is invertible the dimensions* d_i *remain unchanged, and in particular* $\operatorname{rk} B = \operatorname{rk} \tilde{B} = k$. *The matrix* $\tilde{\Gamma}_t$ *for the new coordinate system can be easily recovered by definition* (4.10) *and we find* $\tilde{\Gamma}_t = C\Gamma_t C^*$. *This implies in particular that the definition of* $\mathcal{Q}(t)$ *in* (4.11) *is independent on the coordinates. See also [2] for a general discussion about the geodesic cost.*

The next theorem shows the asymptotic behavior of the family $\mathcal{Q}(t)$ for small $t > 0$. The proof of this result, in its general setting, can be found in [2, Chapter 4].

Theorem 4.11 (Theorem A in [2]). *Let* $x_{\bar{u}} : [0,T] \to \mathbb{R}^n$ *be an optimal trajectory for the problem* (4.6). *The function* $t \mapsto t^2\mathcal{Q}(t)$ *defined in* (4.11) *can be extended to a smooth family of symmetric operators on* $\mathbb{R}^n$ *for all* $t \geq 0$. *In particular, for every fixed* $h \in \mathbb{N}$, *one has the following Laurent expansion for* $t \to 0$

$$\mathcal{Q}(t) = \frac{1}{t^2}\mathcal{I} + \sum_{i=0}^{h} \mathcal{Q}^{(i)} t^i + O\left(t^{h+1}\right). \tag{4.12}$$

Moreover, the matrices $\mathcal{I}$ *and* $\mathcal{Q}^{(i)}$ *for* $i \geq 0$ *are symmetric and*

$$\operatorname{tr}\mathcal{I} = \sum_{i=1}^{m} (2i-1)(k_i - k_{i-1}), \tag{4.13}$$

where $k_i := \dim \operatorname{span}\{B, AB, \ldots, A^{i-1}B\}$.

The expansion (4.12) defines a sequence of symmetric operators (or matrices) $\mathcal{I}$ and $\mathcal{Q}^{(i)}$, for $i \in \mathbb{N}$. These operators are canonically associated to the optimal control problem.

4.4 Small time asymptotics at an equilibrium point

In this section we prove Theorem 4.1, concerning the small time asymptotics of the fundamental solution $p(t,x,x)$ at a point $x \in \operatorname{Ker} A$. We will write this expansion in terms only of the drift field and the invariant $\mathcal{Q}^{(i)}$ introduced in Theorem 4.11.

Remark 4.12. *The exponent* $\mathcal{N}$ *giving the order of the asymptotics appearing in Theorem 4.1 is computed in [43] for a wider class of hypoelliptic operators. In particular, for the linear case the number* $\mathcal{N}$ *is determined only by the numbers* $k_i = \operatorname{rk}\{B, AB, \ldots, A^{i-1}B\}$, *for every* $1 \leq i \leq m-1$, *and* $k_0 = 0$. *Indeed* $\mathcal{N}$ *coincides with the trace of* $\mathcal{I}$ *and is computed as in* (4.13).

Recall now the expression of the fundamental solution $p(t,x,y)$, that we have recovered in Section 2.4, Eq. (2.11):

$$p(t,x,y) = \frac{e^{-\frac{1}{2}(y-e^{tA}x)^* D_t^{-1}(y-e^{tA}x)}}{(2\pi)^{n/2}\sqrt{\det D_t}},$$

where the matrix D_t is defined as

$$D_t = e^{tA}\int_0^t e^{-\tau A}BB^* e^{-\tau A^*} d\tau e^{tA^*} = e^{tA}\Gamma_t e^{tA^*}. \tag{4.14}$$

A change of variable in the integral defining Γ_t gives easily that $D_t = -\Gamma_{-t}$.

If a point x is an equilibrium point for the drift field (i.e., $x \in \text{Ker}A$), then the asymptotics of $p(t,x,x)$ on the diagonal is determined uniquely by the Taylor expansion of $\sqrt{\det D_t}^{-1}$, since $e^{tA}x = x$ for every t.

Let $d(t) := \det D_t$, then $d(t)$ satisfies

$$d'(t) = d(t)\text{tr}(D_t^{-1}\dot{D}_t) = d(t)(2\text{tr}(A) + \text{tr}(B^*D_t^{-1}B)). \tag{4.15}$$

Moreover, since $d(t)$ has a simple pole at $t=0$ of order $\mathcal{N}$ we can write the determinant as $d(t) = t^{\mathcal{N}}f(t)$, for some smooth function f non-vanishing at zero. Substituting this expression in (4.15) one gets

$$\frac{d'(t)}{d(t)} = \frac{\mathcal{N}}{t} + \frac{f'(t)}{f(t)} = 2\text{tr}(A) + \text{tr}(B^*D_t^{-1}B). \tag{4.16}$$

Combining (4.11) and (4.12) of Theorem 4.11, one obtains the asymptotic expansion of $-\frac{d}{dt}\text{tr}(B^*\Gamma_t^{-1}B)$ in terms of the invariants $\mathcal{I}$ and $\mathcal{Q}^{(i)}$. Its integral is

$$-\text{tr}(B^*\Gamma_t^{-1}B) = -\frac{\mathcal{N}}{t} + c + \sum_{i=0}^{h}\text{tr}(\mathcal{Q}^{(i)})\frac{t^{i+1}}{i+1} + O\left(t^{h+2}\right)$$

for some constant c, coming from the integration. Recall that $D_t = -\Gamma_{-t}$. Thus in Eq. (4.16) we have

$$\frac{f'(t)}{f(t)} = 2\text{tr}(A) + c + \sum_{i=0}^{h}(-1)^{i+1}\text{tr}(\mathcal{Q}^{(i)})\frac{t^{i+1}}{i+1} + O(t^{h+2}).$$

Then we can write the determinant of D_t in the following exponential form depending on the invariants $\mathcal{Q}^{(i)}$

$$\det D_t = c_0 t^{\mathcal{N}} e^{(c+2\text{tr}A)t + \sum_{i=0}^{h}(-1)^{i+1}\text{tr}\mathcal{Q}^{(i)}\frac{t^{i+2}}{(i+1)(i+2)} + O(t^{h+3})},$$

for some constant c_0 and the constant c. In particular, one can easily find the first terms of the Taylor expansion of $\det D_t$ at $t=0$. More precisely

$$\det D_t = c_0 t^{\mathcal{N}}\left[1 + (c+2\text{tr}A)t + \frac{(c+2\text{tr}A)^2}{2}t^2 - \frac{\text{tr}\mathcal{Q}^{(0)}}{2}t^2 + o(t^2)\right], \tag{4.17}$$

and the asymptotic expansion is determined up to the constant c.

Remark 4.13. *Notice that, in the above argument, it is crucial that the order $\mathcal{N}$ of* $\det D_t$ *coincides with the first coefficient in the asymptotics of* $\text{tr}\,\mathcal{Q}(t)$.

4.4.1 The first term in the expansion

To find the constant c in Eq. (4.17) let us compute the first terms of the expansion of $\det D_t$. The derivative matrix $\dot{\Gamma}_t = e^{-tA}BB^*e^{-tA^*}$ is positive semi-definite and can be written as

$$\dot{\Gamma}_t = V(t)V(t)^*,$$

for $V(t) = e^{-tA}B$. Let $v_i(t)$ denote the columns of $V(t)$ and define the filtration $E_1 \subset E_2 \subset \ldots \subset E_m = \mathbb{R}^n$ as

$$E_i = \operatorname{span}\{v_j^{(l)}(0),\ 1 \leq j \leq k, 0, \leq l \leq i-1\}. \tag{4.18}$$

Therefore E_i is the subspace of $\mathbb{R}^n$ defined by the columns of the matrices A^jB for $0 \leq j \leq i-1$ and has dimension k_i. Choose coordinates on $\mathbb{R}^n$ adapted to this filtration, i.e., associated with a basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ such that $\operatorname{span}\{e_1, \ldots, e_{k_i}\} = E_i$. In these coordinates $V(t)$ has a peculiar structure, namely

$$V(t) = \begin{pmatrix} \widehat{v}_1 \\ t\widehat{v}_2 \\ \vdots \\ t^{m-1}\widehat{v}_1 \end{pmatrix} + \begin{pmatrix} t\widehat{w}_1 \\ t^2\widehat{w}_2 \\ \vdots \\ t^m\widehat{w}_1 \end{pmatrix} + \begin{pmatrix} O(t^2) \\ O(t^3) \\ \vdots \\ O(t^{m+1}) \end{pmatrix}, \tag{4.19}$$

where $\widehat{v}_i$ and $\widehat{w}_i$ are $(k_i - k_{i-1}) \times k$ constant matrices and every $\widehat{v}_i$ has maximal rank. Let $\widehat{V}(t)$ and $\widehat{W}(t)$ be the first and second principal parts of $V(t)$, then the Taylor series of the matrix Γ_t can be found as

$$\Gamma_t = \int_0^t V(\tau)V(\tau)^* d\tau = \int_0^t \widehat{V}(\tau)\widehat{V}(\tau)^* + \left(\widehat{V}(\tau)\widehat{W}(\tau)^* + \widehat{W}(\tau)\widehat{V}(\tau)^*\right) d\tau + r(t),$$

where $r(t)$ is a remainder term. In components we write Γ_t as a $m \times m$ block matrix, whose block $\Gamma_{ij}(t)$ is the $(k_i - k_{i-1}) \times (k_j - k_{j-1})$ matrix with Taylor expansion

$$\begin{aligned}\Gamma_{ij}(t) &= \left(\frac{\widehat{v}_i\widehat{v}_j^*}{i+j-1}\right) t^{i+j-1} + \left(\frac{\widehat{v}_i\widehat{w}_j^* + \widehat{w}_i\widehat{v}_j^*}{i+j}\right) t^{i+j} + O\left(t^{i+j+1}\right) \\ &= \mathcal{X}_{ij}t^{i+j-1} + \mathcal{Y}_{ij}t^{i+j} + O\left(t^{i+j+1}\right),\end{aligned}$$

where $\mathcal{X}$ and $\mathcal{Y}$ are $m \times m$ block matrices implicitly defined by this formula. Moreover $\mathcal{X}$ is invertible. Let $J_{\sqrt{t}}$ be the $n \times n$ diagonal matrix whose j-th element is equal to $\sqrt{t}^{2i-1}$ for $k_{i-1} < j \leq k_i$, then

$$\Gamma_t = J_{\sqrt{t}}\left(\mathcal{X} + t\mathcal{Y} + O(t^2)\right) J_{\sqrt{t}}. \tag{4.20}$$

The Taylor expansion of the determinant of Γ_t is computed in terms of $\mathcal{X}$ and $\mathcal{Y}$ as follows

$$\det \Gamma_t = t^{\mathcal{N}} \det(\mathcal{X})(1 + \operatorname{tr}(\mathcal{X}^{-1}\mathcal{Y})t + o(t)).$$

Now we are ready to find the determinant of D_t. This follows from the two identities $D_t = e^{tA}\Gamma_t e^{tA^*} = -\Gamma_{-t}$. On one hand we have

$$\begin{aligned}\det D_t &= \det(e^{tA}\Gamma_t e^{tA^*}) = \det(e^{2tA}) \det \Gamma_t \\ &= \det(\mathcal{X})t^{\mathcal{N}}\left[1 + \left(2\operatorname{tr}A + \operatorname{tr}(\mathcal{X}^{-1}\mathcal{Y})\right)t + o(t)\right].\end{aligned} \tag{4.21}$$

On the other hand

$$\begin{aligned}\det D_t &= \det(-\Gamma_{-t}) = (-1)^n(-t)^{\mathcal{N}} \det(\mathcal{X})[1 - t\operatorname{tr}(\mathcal{X}^{-1}\mathcal{Y}) + o(t)] \\ &= \det(\mathcal{X})t^{\mathcal{N}}\left[1 - \operatorname{tr}\left(\mathcal{X}^{-1}\mathcal{Y}\right)t + o(t)\right],\end{aligned} \tag{4.22}$$

where in the last identity we used that n and $\mathcal{N}$ have the same parity, since

$$\mathcal{N} = \sum_{i=1}^{m}(2i-1)k_i = 2M - n,$$

where $M = \sum_{i=1}^{m} ik_i$.

$$\mathcal{N} = \sum_{i=1}^{m}(2i-1)k_i = 2M - n,$$

where $M = \sum_{i=1}^{m} ik_i$. Comparing equations (4.21) and (4.22) we find $\mathrm{tr}(\mathcal{X}^{-1}\mathcal{Y}) = -\mathrm{tr}A$, that means

$$\det D_t = \det(\mathcal{X})t^{\mathcal{N}}(1 + (\mathrm{tr}A)t + o(t)).$$

It follows that in formula (4.17) we have $c = -\mathrm{tr}A$ and $c_0 = \det(\mathcal{X}) > 0$. This allows us to conclude that the asymptotics of the fundamental solution in $x = y \in \ker A$ for small time is

$$\begin{aligned} p(t,x,x) &= \frac{1}{(2\pi)^{n/2}\sqrt{\det D_t}} \\ &= \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}} e^{-\frac{\mathrm{tr}A}{2}t} e^{\frac{1}{2}\sum_{i=0}^{h}(-1)^i \mathrm{tr}\mathcal{Q}^{(i)}\frac{t^{i+2}}{(i+1)(i+2)} + O(t^{h+3})} \\ &= \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}}\left(1 + \sum_{i=1}^{h} a_i t^i + O(t^{h+1})\right), \end{aligned} \tag{4.23}$$

where the coefficients a_i can be explicitly computed from the expansion of the exponential. It follows from (4.23) that every a_i is a polynomial in the components of $\mathrm{tr}A$ and $\mathrm{tr}\mathcal{Q}^{(j)}$ for $j \leq i-2$ of order i and does not depend on x. In particular the first coefficients are computed as follows

$$a_1 = -\frac{\mathrm{tr}A}{2}, \qquad a_2 = \frac{(\mathrm{tr}A)^2}{8} + \frac{\mathrm{tr}\mathcal{Q}^{(0)}}{4},$$

$$a_3 = -\frac{\mathrm{tr}\mathcal{Q}^{(1)}}{12} - \frac{\mathrm{tr}A\,\mathrm{tr}\mathcal{Q}^{(0)}}{8} - \frac{1}{48}(\mathrm{tr}A)^3.$$

4.4.2 Proof of Corollary 4.2

The proof of Corollary 4.2 is now an easy consequence of the previous analysis.

Indeed by Remark 4.4, the minimizer that connects x to y in time t has initial covector $p_0 = \Gamma_t^{-1}\left(e^{-tA}y - x\right)$, therefore

$$S_t(x,y) = \frac{1}{2}\left(e^{-tA}y - x\right)^* \Gamma_t^{-1}\left(e^{-tA}y - x\right) = \frac{1}{2}(y - e^{tA}x)^* D_t^{-1}(y - e^{tA}x).$$

This is exactly the quantity appearing at the exponential of the heat kernel, that therefore can be written in terms of the optimal cost as

$$p(t,x,y) = \frac{e^{-S_t(x,y)}}{(2\pi)^{n/2}\sqrt{\det D_t}}.$$

The statement is then a consequence of the Taylor expansion of $\sqrt{\det D_t}^{-1}$, given in (4.23).

4.5 Small time asymptotics out of the equilibrium

In this section we prove Theorem 4.3, concerning the small time asymptotics of the fundamental solution at a point, x_0, where the drift field is not zero.

By a translation of the origin, we can assume that $x_0 = 0$. This produces a no more linear, but affine drift field, i.e., $X_0(x) = Ax + \alpha$, where $\alpha := Ax_0$ is a column vector equal to the value of the drift at x_0. Then we can study the original asymptotics at x_0, through the asymptotics at the origin of the heat kernel of the linear pde, where to the drift field we add the constant value α. As shown in Section 2.4 its fundamental solution is

$$p(t,x,y) = \frac{e^{\varphi(t,x,y)}}{(2\pi)^{n/2}\sqrt{\det D_t}},$$

where

$$\varphi(t,x,y) = -\frac{1}{2}\left(y - e^{tA}\left(x + \int_0^t e^{-sA} ds\alpha\right)\right)^* D_t^{-1}\left(y - e^{tA}\left(x + \int_0^t e^{-sA} ds\alpha\right)\right)$$

and $D_t = e^{tA}\Gamma_t e^{tA^*}$ is the same covariance matrix as in (4.14). The original asymptotics of the fundamental solution at x_0 is given by the asymptotics of $(\det D_t)^{-1/2}$, found in the previous section, and the asymptotics of φ in $x = y = 0$, i.e.,

$$\varphi(t,0,0) = -\frac{1}{2}\alpha^*\left(\int_0^t e^{-sA} ds\right)^* \Gamma_t^{-1}\left(\int_0^t e^{-sA} ds\right)\alpha.$$

Let E_i be the subspaces defined in (4.18). If $\alpha = Ax_0 \in E_i \setminus E_{i-1}$, then $A^j\alpha \in E_{i+j}$ (actually it could possibly live in the previous subspaces, but not in a bigger one). Moreover $\int_0^t e^{-sA} ds = \sum_{i=1}^m -\frac{(-t)^i}{i!}A^{i-1} + O(t^{m+1})$. Therefore in coordinates adapted to the filtration $\{E_j\}_j$, the column vector $\int_0^t e^{-sA} ds\,\alpha$ can be written in m blocks of height d_j, as

$$\int_0^t e^{-sA} ds\,\alpha = \begin{pmatrix} t\alpha_1 \\ t\alpha_2 \\ \vdots \\ t\alpha_i \\ t^2\alpha_{i+1} \\ \vdots \\ t^{m-i+1}\alpha_m \end{pmatrix} + \begin{pmatrix} O(t^2) \\ O(t^2) \\ \vdots \\ O(t^2) \\ O(t^3) \\ \vdots \\ O(t^{m-i+2}) \end{pmatrix},$$

where α_j is a vector of length d_j for every $1 \le j \le m$ and α_i is not zero. The matrix Γ_t^{-1} can be written as a product

$$\Gamma_t^{-1} = J_{1/\sqrt{t}}(\mathcal{X}^{-1} + O(t))J_{1/\sqrt{t}},$$

where $J_{1/\sqrt{t}}$ and $\mathcal{X}$ are the matrices introduced in (4.20). Notice that

$$J_{1/\sqrt{t}}\int_0^t e^{-sA}ds\,\alpha = \begin{pmatrix} \sqrt{t}\alpha_1 \\ \frac{1}{\sqrt{t}}\alpha_2 \\ \vdots \\ \frac{1}{\sqrt{t}^{2i-3}}\alpha_i \\ \frac{1}{\sqrt{t}^{2i-3}}\alpha_{i+1} \\ \vdots \\ \frac{1}{\sqrt{t}^{2i-3}}\alpha_m \end{pmatrix} + \begin{pmatrix} O(t\sqrt{t}) \\ O(\sqrt{t}) \\ \vdots \\ O(\frac{1}{\sqrt{t}^{2i-1}}) \\ O(\frac{1}{\sqrt{t}^{2i-1}}) \\ \vdots \\ O(\frac{1}{\sqrt{t}^{2i-1}}) \end{pmatrix}.$$

From the last identity we see immediately that, since $\mathcal{X}^{-1}$ has maximal rank and is positive definite, the scalar product

$$\left(J_{1/\sqrt{t}}\int_0^t e^{-sA}ds\,\alpha\right)^* \mathcal{X}^{-1}\left(J_{1/\sqrt{t}}\int_0^t e^{-sA}ds\,\alpha\right)$$

has a simple pole of order $-(2i-3)$ at 0. Thus, if $\alpha \in E_i \setminus E_{i-1}$ with $i > 1$, then $\varphi(t,0,0)$ blows up as $t^{-(2i-3)}$ for $t \to 0$, i.e., the heat kernel decreases with exponential rate precisely as

$$p(t,x_0,x_0) = \frac{t^{-\mathcal{N}/2}}{(2\pi)^{n/2}\sqrt{c_0}}\exp\left(-\frac{C_1 + O(t)}{t^{2i-3}}\right) \qquad \text{for } t \to 0,\ Ax_0 \in E_i \setminus E_{i-1},$$

for a positive constant C_1. This proves the second part of Theorem 4.3. .

Let us consider now the case $\alpha \in \operatorname{span} B$. With this assumption the function $\varphi(t,0,0)$ is smooth in $t = 0$ and we want to find its exact value at the first order.

With a change of variables we can assume that the matrix B is the identity matrix in the first k rows and zero on the last $n-k$ rows. Moreover let $y \in \mathbb{R}^k$ be such that $Ax_0 = By$. We claim that

$$\varphi(t,0,0) = -t\frac{|y|^2}{2} + o(t) = -t\frac{|Ax_0|^2}{2} + o(t).$$

We can write the function φ as

$$\varphi(t,0,0) = -\frac{1}{2}y^*\left(\int_0^t e^{-sA}Bds\right)^* \Gamma_t^{-1}\left(\int_0^t e^{-sA}Bds\right)y$$

and we need to characterize $e^{-sA}B$. Using the results of the previous section, in coordinates adapted to the filtration we can write

$$\int_0^t e^{-sA}Bds = \begin{pmatrix} t\widehat{v}_1 \\ t^2\frac{\widehat{v}_2}{2} \\ \vdots \\ t^m\frac{\widehat{v}_m}{m} \end{pmatrix} + \begin{pmatrix} O(t^2) \\ O(t^3) \\ \vdots \\ O(t^{m+1}) \end{pmatrix},$$

where the $\widehat{v}_i$ are defined in (4.19). They are $d_i \times k$ matrices of maximal rank and $\widehat{v}_1$ is the $k \times k$ identity matrix. We can highlight the dependence on t by multiplying on the left by the diagonal matrix K_t with j-th entry equal to t^i, for $k_{i-1} < j \le k_i$. Also the matrix Γ_t^{-1} can be written as a product (see (4.20)), then

$$\int_0^t e^{-sA}Bds = K_t \left(\begin{pmatrix} \mathbb{I}_k \\ C \end{pmatrix} + O(t) \right) \quad \text{and} \quad \Gamma_t^{-1} = J_{1/\sqrt{t}} \left(\mathcal{X}^{-1} + O(t) \right) J_{1/\sqrt{t}},$$

where C is the $(n-k) \times k$ matrix composed by the last $n-k$ rows of the principal part of $\int_0^t e^{-sA}Bds$. Since $K_t \cdot J_{1/\sqrt{t}} = \sqrt{t}\,\mathbb{I}_n$, then

$$\varphi(t,0,0) = -\frac{t}{2} y^* \begin{pmatrix} \mathbb{I}_k \\ C \end{pmatrix}^* \mathcal{X}^{-1} \begin{pmatrix} \mathbb{I}_k \\ C \end{pmatrix} y + O(t^2). \tag{4.24}$$

Notice that $[\mathcal{X}]_{11} = \mathbb{I}_k$ and $[\mathcal{X}]_{i1} = \frac{\widehat{v}_i}{i} = [\mathcal{X}]_{1i}^*$, for every $2 \le i \le m$, hence $\mathcal{X}$ can be written as a block matrix

$$\mathcal{X} = \begin{pmatrix} \mathbb{I}_k & C^* \\ C & E \end{pmatrix},$$

where E is a $(n-k) \times (n-k)$ matrix.

Lemma 4.14. *The inverse matrix $\mathcal{X}^{-1}$ is the block matrix*

$$\mathcal{X}^{-1} = \begin{pmatrix} [\mathcal{X}^{-1}]_{11} & [\mathcal{X}^{-1}]_{12} \\ [\mathcal{X}^{-1}]_{12}^* & [\mathcal{X}^{-1}]_{22} \end{pmatrix},$$

where

$$[\mathcal{X}^{-1}]_{11} = \mathbb{I}_k + C^*(E - CC^*)^{-1}C,$$
$$[\mathcal{X}^{-1}]_{12} = -C^*(E - CC^*)^{-1}, \qquad [\mathcal{X}^{-1}]_{22} = (E - CC^*)^{-1}.$$

Proof. This is the general expression of the inverse of a block matrix, provided $[\mathcal{X}]_{11}$ and $E - CC^*$ are not singular. $[\mathcal{X}]_{11}$ is the identity matrix. Let us show that $E - CC^*$ is not singular. Assume $x \in \mathbb{R}^{n-k}$ satisfies $(E - CC^*)x = 0$. Then the column vector (of dimension n) equal to $(-(C^*x)^*, x^*)^*$ is in the kernel of $\mathcal{X}$. Therefore $x = 0$, since $\mathcal{X}$ is not singular. □

Applying the Lemma to Eq. (4.24) we find that

$$\begin{aligned} \varphi(t,0,0) &= -\frac{t}{2} y^* \left([\mathcal{X}^{-1}]_{11} + ([\mathcal{X}^{-1}]_{12}C)^* + [\mathcal{X}^{-1}]_{12}C + C^*[\mathcal{X}^{-1}]_{22}C \right) y + O(t^2) \\ &= -\frac{t}{2} y^* \,\mathbb{I}_k\, y + O(t^2) = -\frac{|Ax_0|^2}{2} t + O(t^2). \end{aligned}$$

Taking into account the asymptotics of $(\det D_t)^{-1/2}$ found in the previous section, this completes the proof of Theorem 4.3.

Chapter 5

Volume geodesic distorsion and Ricci curvature for Hamiltonian dynamics

In this chapter, based on the results of [1], we study the variation of a smooth volume under the flow associated to a quadratic Hamiltonian. We introduce a main invariant describing the interaction of the measure and the dynamics and we show how this invariant, together with curvature-like invariants of the dynamics introduced in [3], appear in this expansion. This generalizes the well-known result in Riemannian geometry and includes all sub-Riemannian manifolds.

5.1 Introduction

One of the possible ways of introducing curvature in Riemannian geometry is by looking for the variation of a smooth volume under the geodesic flow. Indeed, given a point x on a Riemannian manifold (M, g) and a tangent unit vector $v \in T_xM$, it is well-known that the asymptotic expansion of the Riemannian volume vol_g in the direction of v depends on the Ricci curvature at x.

More precisely, let us consider a geodesic $\gamma(t) = \exp_x(tv)$ with initial tangent vector v starting at x. For a fixed orthonormal basis $e_1, \ldots, e_n$ in T_xM let

$$\partial_i|_{\gamma(t)} := (d_{tv} \exp_x)(e_i), \qquad 1 \leq i \leq n.$$

be the image of e_i through the differential at tv of the Riemannian exponential map $\exp_x : T_xM \to M$. Once we take a set of normal coordinates centered in x, the vector fields ∂_i are the coordinate vector fields at $\gamma(t)$ and the volume element, that is written as $\mathrm{vol}_g = \sqrt{\det g_{ij}} dx_1 \ldots dx_n$, satisfies the expansion for $t \to 0$

$$\sqrt{\det g_{ij}(\exp_x(tv))} = 1 - \frac{1}{6}\mathrm{Ric}_g(v, v)t^2 + O(t^3), \tag{5.1}$$

where Ric_g is the Ricci curvature tensor associated with g.

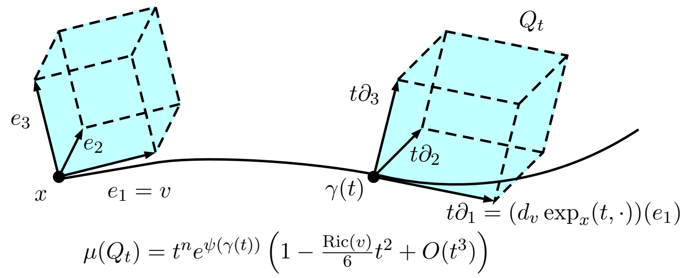


Figure 5.1: Volume distortion on a Riemannian manifold with volume $\mu = e^{\psi}\mathrm{vol}_g$

In particular, the left hand side of (5.1) measures the Riemannian volume of the parallelotope in $\gamma(t)$ with edges ∂_i, more explicitly

$$\mathrm{Vol}_g\left(\bigwedge_{i=1}^{n} \partial_i|_{\gamma(t)}\right) = \sqrt{\det g_{ij}(\gamma(t))}.$$

In the sub-Riemannian setting this construction cannot be directly generalized. Indeed, the sub-Riemannian exponential map is not a local diffeomorphism at zero and to compute the volume of small balls one should have a precise knowledge of the structure of the cut locus, which is not easy. Nevertheless the geodesic flow on the Riemannian manifold can be seen as the projection of a Hamiltonian flow on the cotangent bundle, associated to a non-degenerate quadratic Hamiltonian. On a sub-Riemannian manifold, and more in general even for structures deriving from an affine control system, this approach can be developed. Indeed the Hamiltonian flow is defined in a similar way and in particular, if the structure is sub-Riemannian, the restriction of the Hamiltonian to any fiber is a degenerate non-negative quadratic form. The projection on the manifold, M, of its integral curves are geodesics, but, contrary to the Riemannian case, in general not all the geodesics can be obtained in this way. These projected geodesics are then parametrized by the initial covector in the cotangent bundle and if they are sufficiently regular (ample and equiregular geodesics), it is possible to compute the variation of the volume in a "smooth" way by looking at the measure as an n-form in the cotangent space T^*M, which has dimension $2n$, restricted to the fiber T^*_xM.

To generalize this analysis to a sub-Riemannian structure, let us consider again the Riemannian case.

Let (M, g) be a Riemannian manifold, endowed with a smooth volume μ and let ψ be the smooth function such that $\mu = e^{\psi}\mathrm{vol}_g$. It is convenient to express the exponential map on M in terms of the Hamiltonian flow. Indeed, let $\exp_x(t, v)$ denote the point reached by a curve at time t starting from x with velocity v, i.e., $\exp_x(t, v) = \exp_x(tv)$. The metric g induces a canonical identification between T_xM and the cotangent space T^*_xM. So the exponential map can be seen as a Hamiltonian flow, indeed

$$\exp_x(tv) = \exp_x(t, v) = \pi\left(e^{t\vec{H}}V\right), \tag{5.2}$$

where in the last expression V denotes the element in T_x^*M corresponding to v. Then

$$(d_v \exp_x(t,\cdot))\,(e_i) = \left.\frac{\partial}{\partial s}\right|_{s=0} \exp_x(tv + tse_i) = t\,\partial_i|_{\gamma(t)}\,. \tag{5.3}$$

As pictorically expressed in Fig. 5.1, for $t \searrow 0$ the volume of the parallelotope, Q_t, with edges $t\partial_i$ has the following expansion deriving from (5.1) and depending on the choice of μ

$$\mu\left(\bigwedge_{i=1}^{n} t\partial_i|_{\gamma(t)}\right) = t^n e^{\psi(\gamma(t))}\left(1 - \frac{1}{6}\mathrm{Ric}(v,v)t^2 + O(t^3)\right). \tag{5.4}$$

We can interpret the last identity from the Hamiltonian point of view, see Fig. 5.2. Indeed, let λ be the initial cotangent vector of γ. In other words, λ is the element of T_x^*M associated to $\dot\gamma(0)$. For every e_i let E_i denote the associated cotangent vector in T_x^*M. Then by (5.3) and (5.2) $t\partial_i = (\pi \circ e^{t\vec{H}})_* E_i$. So the left hand side of (5.4) can be written as

$$\begin{aligned}\mu(Q_t) &= \langle \mu_{\gamma(t)}, (t\partial_1, \dots, t\partial_n)\rangle \\ &= \langle \mu_{\pi(e^{t\vec{H}}(V))}, \Big((\pi\circ e^{t\vec{H}})_* E_1, \dots, (\pi\circ e^{t\vec{H}})_* E_n\Big)\rangle \\ &= \langle (\pi\circ e^{t\vec{H}})^*\mu, (E_1,\dots,E_n)\rangle_\lambda.\end{aligned}$$

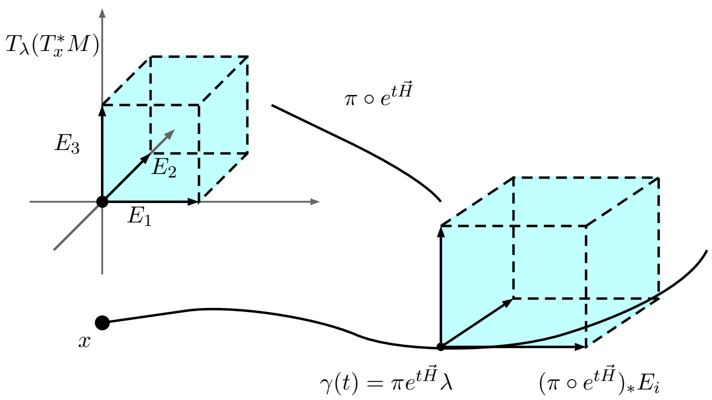


Figure 5.2: Equivalent volume distortion as variation under the Hamiltonian flow

Observe that the pull-back $(\pi\circ e^{t\vec{H}})^*\mu$ defines an n-form on the cotangent bundle T^*M, that has dimension $2n$. The quantity that we compute is the restriction of this form to the n-dimensional fiber T_x^*M. Moreover, the volume μ_x defines naturally a volume μ_λ on the fiber T_x^*M. With this Hamiltonian interpretation, the classical Riemannian asymptotics (5.4) determines the variation of $(\pi\circ e^{t\vec{H}})^*\mu$ restricted to the fiber T_x^*M, with respect to the volume μ_λ, i.e.,

$$\left.(\pi\circ e^{t\vec{H}})^*\mu\right|_{T_x^*M} = t^n\, e^{\int_0^t \psi'(\gamma(\tau))d\tau}\left(1 - \frac{1}{6}\mathrm{Ric}_g(v,v)t^2 + O(t^3)\right)\mu_\lambda. \tag{5.5}$$

Eq. (5.5) underlines geometric properties of the variation of the volume, as well as its measure properties, separated in distinct parts. Indeed, we see that the order term t^n depends only on the dimension of the manifold. The asymptotics in the brackets contains only geometric information, that depend on the metric g on M. The constant term $e^{\psi(x)}$ depending on μ at the initial point is contained in the associated volume μ_λ. Finally, the measure information is encoded in the exponential term. Indeed it represents the variation of μ along the geodesic and is equal to the exponential of $\int_0^t \psi'(\tau)d\tau = \int_0^t \langle \mathrm{grad}\psi, \dot\gamma(\tau)\rangle d\tau$. In particular, it defines a measure invariant function ρ at every initial cotangent vector λ:

$$\rho(\lambda)\mu_\lambda := \frac{d}{dt}\bigg|_{t=0} \left(t^{-n} \, (\pi \circ e^{t\vec{H}})^* \mu|_\lambda \Big|_{T_x^*M} \right), \qquad \lambda \in T_x^*M. \tag{5.6}$$

This function depends on the particular volume μ and on the underlying geometry g of M. More explicitly, let $\gamma(t) = \pi(e^{t\vec{H}}(\lambda))$ for $\lambda \in T_x^*M$ and let T denote a vector field on M that extends $\dot\gamma(t)$ along the curve, then

$$\rho(\lambda) = \mathrm{div}_\mu \mathsf{T}_x,$$

where div_μ denotes the divergence with respect to the volume μ.

Figure 5.3 illustrates the variation of the volume from the metric measure viewpoint. Indeed let $\Omega \subset T_x^*M$ be a small neighborhood of λ and let $\Omega_{x,t} := \pi \circ e^{t\vec{H}}(\Omega)$ be its image on M with respect to the Hamiltonian flow. For every t it is a neighborhood of $\gamma(t)$. Then

$$\mu(\Omega_{x,t}) = \int_\Omega (\pi \circ e^{t\vec{H}})^* \mu,$$

and (5.5) represents the variation of the volume element along γ.

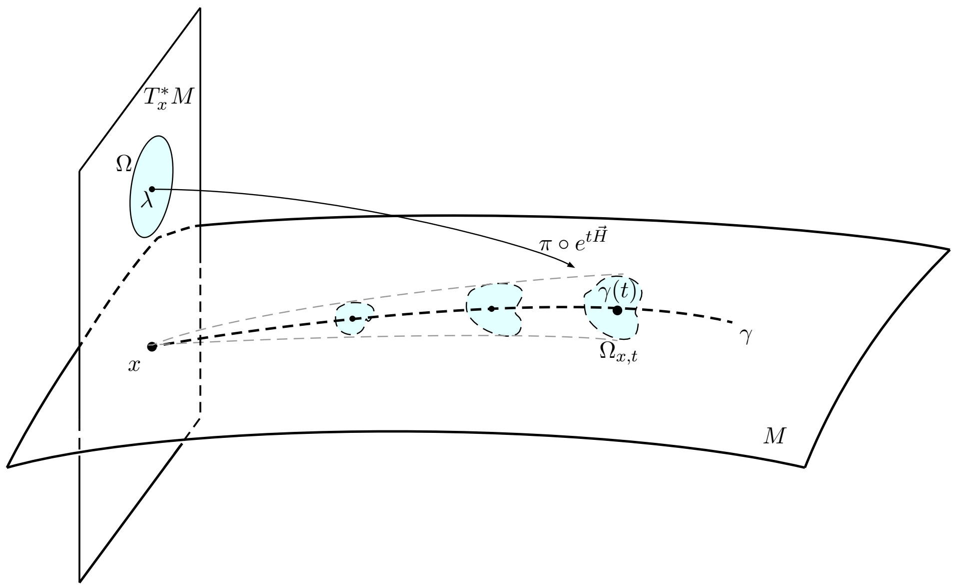


Figure 5.3: Variation of volume

In this chapter we generalize the asymptotics (5.5) to a sub-Riemannian structure, and more in general to any structure arising from a quadratic Hamiltonian. Let M be a smooth manifold and let $\vec{H}$ denote a quadratic, possibly degenerate, Hamiltonian. A special class of dynamics is given by the Hamiltonian, whose restriction to a fiber T_x^*M is a degenerate homogeneous quadratic form (i.e., without linear or constant terms). Then this case recovers the sub-Riemannian structures on the manifold M.

Fix $\lambda \in T_x^*M$ and let $\gamma(t) = \pi(e^{t\vec{H}}\lambda)$ be the associated geodesic on M. The asymptotics that we obtain is expressed as in (5.5) and we interpret every component as a generalization of the corresponding Riemannian element. In particular, the Hamiltonian at λ generates a constant leading term c_0 and influences the order of the asymptotic. Indeed, we observe that the order of the asymptotics is not constant, but depends on the particular geodesic. Indeed it is equal to the geodesic dimension, $\mathcal{N}(\lambda)$, of γ (cf. Definition 5.6), which for an n-dimensional Riemannian manifold, is independent on the curve and is always equal to n.

Moreover, the asymptotics depends on two geometric invariants, that are rational functions in the initial covector λ. The first one is a modification of the Ricci tensor, that is substituted now by the trace of a curvature operator in the direction of λ. This curvature operator, $\mathcal{R}_\lambda$, is a generalization of the sectional curvature and is defined in [3] for the wide class of geometric structures arising from affine control systems. The second invariant is the generalization of the measure invariant, $\rho(\lambda)$, introduced in (5.6). Indeed it is defined as

$$\rho(\lambda)\mu_\lambda := \frac{d}{dt}\bigg|_{t=0} \left(t^{-\mathcal{N}(\lambda)} \, (\pi \circ e^{t\vec{H}})^* \mu|_\lambda \Big|_{T_x^*M} \right),$$

where $\mathcal{N}$ is the geoedesic dimension of $\gamma(t) = \pi(e^{t\vec{H}}(\lambda))$. It is a measure metric invariant and represents how the volume changes along the curve with respect to a reference n-dimensional form given by the Hamiltonian. It depends, obviously on the fixed volume μ, as in the Riemannian case, but also on the symbol of the geodesic (cf. Definition 5.7), that represents the microlocal nilpotent approximation of $\gamma(t)$ with zero curvature. The symbol of any curve in a Riemannian manifold is trivial, so this behaviour was not evident in (5.5). If the structure in strictly non-Riemannian, we show an explicit formula to determine ρ, which involves the symbol of the curve and the variation of μ along the curve. In particular, we compute it for the special case of contact manifolds endowed with Popp's volume.

The precise statement of our theorem is as follows.

Theorem 5.1. *Let μ be a smooth volume on M and γ an equiregular ample geodesic, with initial covector $\lambda \in T_x^*M$. Let μ_λ be the dual volume on T_x^*M. Then the pull back of μ with respect to the projection of the Hamiltonian flow, and restricted to T_x^*M has the following asymptotic expansion*

$$(\pi \circ e^{t\vec{H}})^* \mu_\lambda \Big|_{T_x^*M} = c_0 t^{\mathcal{N}} \, e^{\int_0^t \rho(\lambda(\tau))d\tau} \left(1 - t^2 \frac{\mathrm{tr}\mathcal{R}_\lambda}{6} + o(t^2) \right) \mu_\lambda^*$$

where $\mathcal{N}$ is the geodesic dimension associated to γ (cf. Definition 5.6), $\mathcal{R}_\lambda$ is the $k \times k$ curvature matrix associated to λ and $\rho(\lambda(t))$ is a rational operator in $\lambda(t)$ depending on μ and the symbol of the curve γ.

5.2 The general setting

Let M be an n-dimensional connected manifold and $f_0, f_1 \dots, f_k \in \text{Vec}(M)$ smooth vector fields, with $k \leq n$. We consider the following *affine control system* on M

$$\dot{x}(t) = f_0(x(t)) + \sum_{i=1}^{k} u_i(t) f_i(x(t)), \qquad x \in M, u \in L^\infty(\mathbb{R};\mathbb{R}^k). \tag{5.9}$$

The essentially bounded function u is called *control* function. A Lipschitz curve $\gamma : [0,T] \to M$ is said to be *admissible* for the system if there exists a control $u \in L^\infty([0,T],\mathbb{U})$ such that γ satisfies (5.9) for a.e. $t \in [0,T]$. The pair (γ, u) of an admissible curve γ and its control u is called *admissible pair.*

Remark 5.2. *The affine control system can be generalized to not globally trivializable vector bundles. Indeed it can be defined more generally as a pair* $(\mathbb{U}, f)$ *such that* $\mathbb{U}$ *is a smooth rank* k *vector bundle with base* M *and fiber* $\mathbb{U}_x$*, and* $f : \mathbb{U} \to TM$ *is a smooth affine morphism of vector bundles such that* $\pi \circ f(u) = x$*, for every* $u \in \mathbb{U}_x$*. Locally this system can be written as in* (5.9)*, by taking a local traivialization of* $\mathbb{U}$*, a basis in the fibers and define the map* f *as* $f(u) = f_0 + \sum_{i=1}^{k} u_i f_i$ *for* $u \in \mathbb{R}^k$*.*

We denote by $\mathcal{D} \subset TM$ the family of subspaces, $\mathcal{D}_x$, of T_xM spanned by the linear part of the control problem at $x \in M$, i.e.

$$\mathcal{D} = \{\mathcal{D}_x\}_{x\in M}, \qquad \text{where } \mathcal{D}_x := \text{span}\{f_1, \dots, f_k\}_x.$$

A vector field X is called *horizontal* if $X_x \in \mathcal{D}_x$ for every x and we denote by $\overline{\mathcal{D}}$ the set of sections of $\mathcal{D}$. In the following we will assume that the distribution $\mathcal{D}$ has constant rank, k.

Among all admissible trajectories (γ, u) that join two fixed points in time T, we want to minimize the *cost functional*

$$J_T(u) := \frac{1}{2}\int_0^T |u(\tau)|^2 + Q(x(\tau))d\tau,$$

where Q is a smooth function on M.

Since the distribution has constant rank, we endow $\mathcal{D}$ with a scalar product such that the fields $f_1, \dots, f_k$ are orthonormal. Then, if f_0 and Q are zero, the cost functional reduces to the definition of length of curves.

Definition 5.3. *For all fixed points* $x_1, x_2 \in M$ *and time* $T > 0$*, we define the* value function

$$S_T(x_1, x_2) := \inf \{J_T(u) | (\gamma, u) \text{ admissible pair}, \gamma(0) = x_1, \gamma(x_T) = x_2\}. \tag{5.10}$$

In the following we will assume that the system is controllable, i.e. for every fixed point $x_1, x_2 \in M$ and time $T > 0$ there exists an admissible curve joining x_1 and x_2 in time T. With this assumption the value function in always finite.

Important examples of affine control problems are sub-Riemannian structures. These are a triple $(M, \mathcal{D}, g)$, where M is a smooth manifold, $\mathcal{D}$ is a smooth, completely non-integrable vector sub-bundle of TM and g is a smooth scalar product on $\mathcal{D}$. The value function is now the sub-Riemannian distance, i.e. the infimum of the length of absolutely continuous admissible curves joining two points. Since its tangent vector is almost everywhere in the distribution, the length can be computed via the scalar product g. The totally non-holonomic assumption on $\mathcal{D}$ implies, by the Rashevskii-Chow theorem, the controllability of the system and that the distance is finite on every connected component of M. Moreover the metric topology coincides with the one of M. A more detailed introduction on sub-Riemannian geometry can be found in [40, 5].

In Riemannian geometry, it is well-known that the geodesic flow can be seen as a Hamiltonian flow on the cotangent bundle T^*M, associated with the Hamiltonian

$$H(p,x) = \frac{1}{2}\sum_{i=1}^{n}\langle p, f_i(x)\rangle^2, \qquad (p,x) \in T^*M,$$

where $X_1,\ldots,X_n$ is any local orthonormal frame for the Riemannian structure.

For an affine control system, the Hamiltonian is still defined and is a generalization of the previous one. Namely, for a local frame $f_1,\ldots,f_k$ for $\mathcal{D}$ we set

$$H(p,x) = \frac{1}{2}\sum_{i=1}^{k}\langle p, f_i(x)\rangle^2 + \langle p, f_0(x)\rangle - \frac{1}{2}Q(x), \qquad (p,x) \in T^*M.$$

Hamilton's equations are written as a flow on T^*M

$$\dot\lambda = \vec{H}(\lambda), \qquad \lambda \in T^*M,$$

where $\vec{H}$ is the Hamiltonian vector field associated with H. The projection $\pi : T^*M \to M$ of its integral curves are geodesics, i.e. locally minimizing curves. In the general case, some geodesics may not be recovered in this way. These are the so-called strictly abnormal geodesics [39], and they are related with hard open problems in sub-Riemannian geometry.

In what follows, with a slight abuse of notation, the term "geodesic" refers to the not strictly abnormal ones.

An integral line of the Hamiltonian vector field $\lambda(t) = e^{t\vec{H}}(\lambda) \in T^*M$, with initial covector λ is called *extremal*. Notice that the same geodesic may be the projection of two different extremals.

5.3 Geodesic flag and symbol

In this section we define the flag and symbol of a geodesic, that are elements carrying information about the germ of the distribution

Let $\gamma : [0,T] \to M$ be a geodesic and consider a smooth admissible extension of the tangent vector, namely a vector field $\mathsf{T} = f_0 + X$, with $X \in \overline{\mathcal{D}}$, such that $\mathsf{T}(\gamma(t)) = \dot\gamma(t)$ for every $t \in [0,T]$.

Definition 5.4. *The* flag of the geodesic $\gamma : [0,T] \to M$ *is the sequence of subspaces*

$$\mathcal{F}^i_{\gamma(t)} := \operatorname{span}\{\mathcal{L}^j_{\mathsf{T}}(X)|_{\gamma(t)} \mid X \in \overline{\mathcal{D}},\ j \leq i-1\} \subseteq T_{\gamma(t)}M, \qquad \forall\, i \geq 1,$$

for any fixed $t \in [0,T]$, *where* $\mathcal{L}_{\mathsf{T}}$ *denotes the Lie derivative in the direction of* T.

Definition 5.4 is well posed, namely does not depend on the choice of the admissible extension T (see [3, Sec. 3.4]). By definition, the flag is a filtration of $T_{\gamma(t)}M$, i.e. $\mathcal{F}^i_{\gamma(t)} \subseteq \mathcal{F}^{i+1}_{\gamma(t)}$, for all $i \geq 1$. Moreover, $\mathcal{F}^1_{\gamma(t)} = \mathcal{D}_{\gamma(t)}$. The *growth vector* of the geodesic $\gamma(t)$ is the sequence of integer numbers

$$\mathcal{G}_{\gamma(t)} := \{\dim \mathcal{F}^1_{\gamma(t)}, \dim \mathcal{F}^2_{\gamma(t)}, \dots\}.$$

A geodesic $\gamma(t)$, with growth vector $\mathcal{G}_{\gamma(t)}$, is said

- *equiregular* if $\dim \mathcal{F}^i_{\gamma(t)}$ does not depend on t for all $i \geq 1$,
- *ample* if for all t there exists $m \geq 1$ such that $\dim \mathcal{F}^m_{\gamma(t)} = \dim T_{\gamma(t)}M$.

Equiregular (resp. ample) geodesics are the microlocal counterpart of equiregular (resp. bracket-generating) distributions. Let $d_i := \dim \mathcal{F}^i_\gamma - \dim \mathcal{F}^{i-1}_\gamma$, for $i \geq 1$, be the increment of dimension of the flag of the geodesic at each step (with the convention $\dim \mathcal{F}^0 = 0$).

Lemma 5.5 ([3]). *For an equiregular, ample geodesic, $d_1 \geq d_2 \geq \ldots \geq d_m$.*

Definition 5.6. *The* geodesic dimension *of an ample, equiregular geodesic is*

$$\mathcal{N} := \sum_{i=1}^{m} (2i-1) d_i.$$

The generic geodesic is ample and equiregular. More precisely, the set of points $x \in M$ such that there exists a non-empty Zariski open set $A_x \subseteq T^*_x M$ of initial covectors for which the associated geodesic is ample and equiregular, is open and dense in M. See [3, 49] for more details.

Fix an ample equiregular geodesic $\gamma : [0, T] \to M$ and let T be an admissible extension of its tangent vector. For $X \in \mathcal{F}^i_{\gamma(t)}$, consider a smooth extension of X along γ such that $X_{\gamma(s)} \in \mathcal{F}^i_{\gamma(s)}$ for every $s \in [0, T]$ and define

$$\mathcal{L}_{\mathsf{T}}(X) := [\mathsf{T}, X]_{\gamma(t)} \bmod \mathcal{F}^i_{\gamma(t)}.$$

It is easy to see that this definition does not depend on the admissible extension T and on the extension X under the regularity assumption. So the maps

$$\mathcal{L}_{\mathsf{T}} : \mathcal{F}^i_{\gamma(t)} / \mathcal{F}^{i-1}_{\gamma(t)} \to \mathcal{F}^{i+1}_{\gamma(t)} / \mathcal{F}^i_{\gamma(t)}, \qquad i \geq 1$$

are well-defined and surjective. In particular $\mathcal{L}^i_{\mathsf{T}} : \mathcal{F}_{\gamma(t)} \to \mathcal{F}^{i+1}_{\gamma(t)} / \mathcal{F}^i_{\gamma(t)}$ are surjective linear maps from the distribution $\mathcal{D}_{\gamma(t)} = \mathcal{F}_{\gamma(t)}$.

Definition 5.7. *Given a curve γ which is ample and equiregular we define its* symbol *at $\gamma(t)$, denoted by $S_{\gamma(t)}$, as follows*

- $\mathrm{gr}_{\gamma(t)}(\mathcal{F}) = \oplus_{i=0}^{m-1} \mathcal{F}^{i+1}_{\gamma(t)} / \mathcal{F}^i_{\gamma(t)}$
- *the operator $\mathcal{L}^i_{\mathsf{T}} : \mathcal{D}_{\gamma(t)} \to \mathcal{F}^{i+1}_{\gamma(t)} / \mathcal{F}^i_{\gamma(t)}$ for $i \geq 1$*

where T is an horizontal extension of $\dot{\gamma}$.

Definition 5.8. *Two different symbols $S_{\gamma(t)}$ and $S_{\gamma'(t')}$ of $\gamma : [0,T] \to M$ and $\gamma' : [0,T'] \to M$ resp. are said to be* equivalent *if there exist a reparametrization $\sigma : [0,T] \to [0,T']$ such that $\sigma(t) = t'$ and a diffeomorphism $\phi \in C^\infty(M)$ such that*

- $\phi(\gamma(s)) = \gamma'(\sigma(s))$ *for every s in a neighborhood of t*
- $\phi_* : \mathcal{D} \to \mathcal{D}'$ *is an isometry in a neighborhood of $\gamma(t)$ and $\gamma'(t')$ and $\mathcal{L}_{\mathsf{T}'} \circ \phi_* = \phi_* \circ \mathcal{L}_{\mathsf{T}}$.*

Remark 5.9. *In Definition 5.8 the time in which the symbols are considered is not meaningful, while it is important that ϕ is a diffeomorphisms between the two curves. In particular, by a reparametrization of the curve γ' we can assume that the two curves, γ and γ', are defined on the same interval $[0,T]$ and that the reparametrization σ is the identity. So in the following we will avoid the choice of σ in Definition 5.8 and consider the two symbols evaluated at the same time $S_{\gamma(t)}$ and $S_{\gamma'(t)}$.*

Through the surjective maps $\mathcal{L}^i_{\mathsf{T}} : \mathcal{D}_{\gamma(t)} \to \mathcal{F}^{i+1}_{\gamma(t)}/\mathcal{F}^i_{\gamma(t)}$ the inner product on $\mathcal{D}$ naturally induces an inner product on $\mathrm{gr}_{\gamma(t)}(\mathcal{F})$ such that the norm of $Y \in \mathcal{F}^{i+1}_{\gamma(t)}/\mathcal{F}^i_{\gamma(t)}$ is

$$\|Y\|_{\mathcal{F}^{i+1}_{\gamma(t)}/\mathcal{F}^i_{\gamma(t)}} := \min\left\{\|X\|_{\mathcal{D}} \text{ s.t. } \mathcal{L}^i_{\mathsf{T}} X = Y \operatorname{mod} \mathcal{F}^i_{\gamma(t)}\right\}.$$

Lemma 5.10. *If two symbols $S_{\gamma(t)}$ and $S_{\gamma'(t)}$ are equivalent then they are isometric as inner product spaces.*

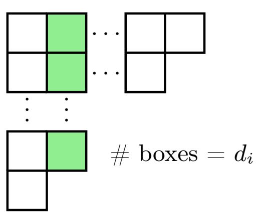
Proof. Let $\mathcal{F}^i_{\gamma(t)}$ and $\mathcal{F}^i_{\gamma'(t)}$ be the subspaces that define the symbols $S_{\gamma(t)}$ and $S_{\gamma'(t)}$ respectively and ϕ be the diffeomorphism of Definition 5.8. Let $X' \in \mathcal{D}'$ and $X \in \mathcal{D}$ such that $X' = \phi_* X$. By the commutation property of ϕ_* with $\mathcal{L}_{\mathsf{T}}$ it holds

$$\mathcal{L}^i_{\mathsf{T}'} X' = \mathcal{L}^i_{\mathsf{T}'}(\phi_* X) = \phi_*\left(\mathcal{L}^i_{\mathsf{T}} X\right),$$

therefore $\mathcal{F}^i_{\gamma'} = \phi_* \mathcal{F}^i_{\gamma}$ for every i and the curves have the same growth vectors and the same step m. Therefore ϕ_* descends to maps between every layer of the stratification in the following way. For $Y \in \mathcal{F}^i_{\gamma(t)}/\mathcal{F}^{i-1}_{\gamma(t)}$ let $X \in \mathcal{D}_{\gamma(t)}$ such that $Y = \mathcal{L}^{i-1}_{\mathsf{T}} X \operatorname{mod} \mathcal{F}^{i-1}_{\gamma(t)}$. Then we define $\phi^i_*(Y) := \mathcal{L}^{i-1}_{\mathsf{T}'}(\phi_*(X)) \in \mathcal{F}^i_{\gamma'(t)}/\mathcal{F}^{i-1}_{\gamma'(t)}$. Since ϕ_* is an isometry on the distribution and commutes with T, ϕ^i_* is an isometry on the quotient spaces. □

5.4 Young diagram, canonical frame and Jacobi fields

For an ample, equiregular geodesic we can build a tableau D with m columns of length d_i, for $i = 1, \ldots, m$, as follows:



The total number of boxes in D is $n = \dim M = \sum_{i=1}^{m} d_i$.

Consider an ample, equiregular geodesic, with Young diagram D, with k rows, of length $n_1, \ldots, n_k$. Indeed $n_1 + \ldots + n_k = n$. We are going to introduce a moving frame on $T_{\lambda(t)}(T^*M)$ indexed by the boxes of the Young diagram. The notation $ai \in D$ denotes the generic box of the diagram, where $a = 1, \ldots, k$ is the row index, and $i = 1, \ldots, n_a$ is the progressive box number, starting from the left, in the specified row. We employ letters $a, b, c, \ldots$ for rows, and $i, j, h, \ldots$ for the position of the box in the row.

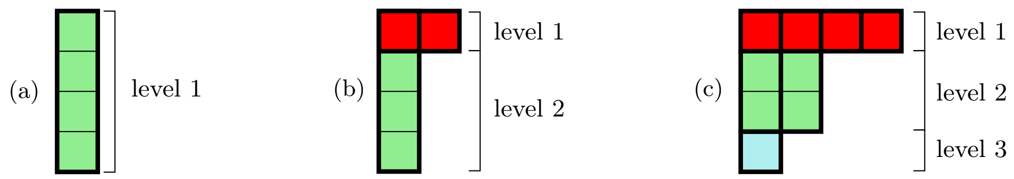


Figure 5.4: Levels (shaded regions) and superboxes (delimited by bold lines) for the Young diagram of (a) Riemannian, (b) contact, (c) a more general structure. The Young diagram for any Riemannian geodesic has a single level and a single superbox.

We collect the rows with the same length in D, and we call them *levels* of the Young diagram. In particular, a level is the union of r rows $D_1, \ldots, D_r$, and r is called the *size* of the level. The set of all the boxes $ai \in D$ that belong to the same column and the same level of D is called *superbox*. We use Greek letters $\alpha, \beta, \ldots$ to denote superboxes. Notice that two boxes ai, bj are in the same superbox if and only if ai and bj are in the same column of D and in possibly distinct rows but with same length, i.e. if and only if $i = j$ and $n_a = n_b$ (see Fig. 5.4).

The following theorem is proved in [49].

Theorem 5.11. *Assume $\lambda(t)$ is the lift of an ample and equiregular geodesic $\gamma(t)$ with Young diagram D. Then there exists a smooth moving frame $\{E_{ai}, F_{ai}\}_{ai\in D}$ along $\lambda(t)$ such that*

(i) $\pi_* E_{ai}|_{\lambda(t)} = 0$.

(ii) It is a Darboux basis, namely

$$\sigma(E_{ai}, E_{bj}) = \sigma(F_{ai}, F_{bj}) = \sigma(E_{ai}, F_{bj}) = \delta_{ab}\delta_{ij}, \qquad ai, bj \in D.$$

(iii) The frame satisfies structural equations

$$\begin{cases} \dot{E}_{ai} = E_{a(i-1)} & a = 1, \ldots, k, \quad i = 2, \ldots, n_a, \\ \dot{E}_{a1} = -F_{a1} & a = 1, \ldots, k, \\ \dot{F}_{ai} = \sum_{bj\in D} R_{ai,bj}(t) E_{bj} - F_{a(i+1)} & a = 1, \ldots, k, \quad i = 1, \ldots, n_a - 1, \\ \dot{F}_{an_a} = \sum_{bj\in D} R_{an_a,bj}(t) E_{bj} & a = 1, \ldots, k, \end{cases} \tag{5.11}$$

for some smooth family of $n\times n$ symmetric matrices $R(t)$, with components $R_{ai,bj}(t) = R_{bj,ai}(t)$, indexed by the boxes of the Young diagram D. The matrix $R(t)$ is normal *in the sense of [49].*

If $\{\widetilde{E}_{ai}, \widetilde{F}_{ai}\}_{ai\in D}$ *is another smooth moving frame along* $\lambda(t)$ *satisfying (i)-(iii), with some normal matrix* $\widetilde{R}(t)$*, then for any superbox* α *of size* r *there exists an orthogonal constant* $r\times r$ *matrix* O^α *such that*

$$\widetilde{E}_{ai} = \sum_{bj\in\alpha} O^\alpha_{ai,bj} E_{bj}, \qquad \widetilde{F}_{ai} = \sum_{bj\in\alpha} O^\alpha_{ai,bj} F_{bj}, \qquad ai\in\alpha.$$

Remark 5.12. *For* $a = 1,\dots,k$*, we denote by* E_a *the* n_a*-dimensional column vector* $E_a = (E_{a1}, E_{a2}, \dots, E_{an_a})^*$*, with analogous notation for* F_a*. Similarly,* E *denotes the* n*-dimensional column vector* $E = (E_1,\dots,E_k)^*$*, and similarly for* F*. Then, we rewrite the system* (5.11) *as follows*

$$\begin{pmatrix}\dot E\\ \dot F\end{pmatrix} = \begin{pmatrix} C_1^* & -C_2\\ R(t) & -C_1\end{pmatrix}\begin{pmatrix}E\\ F\end{pmatrix},$$

where $C_1 = C_1(D)$*,* $C_2 = C_2(D)$ *are* $n\times n$ *matrices, depending on the Young diagram* D*, defined as follows: for* $a,b = 1,\dots,k$*,* $i = 1,\dots,n_a$*,* $j = 1,\dots,n_b$*:*

$$[C_1]_{ai,bj} := \delta_{ab}\delta_{i,j-1},, \qquad [C_2]_{ai,bj} := \delta_{ab}\delta_{i1}\delta_{j1}.$$

It is convenient to see C_1 *and* C_2 *as block diagonal matrices:*

$$C_i(D) := \begin{pmatrix} C_i(D_1) & & \\ & \ddots & \\ & & C_i(D_k)\end{pmatrix}, \qquad i = 1,2,$$

the a*-th block being the* $n_a\times n_a$ *matrices*

$$C_1(D_a) := \begin{pmatrix} 0 & \mathbb{I}_{n_a-1}\\ 0 & 0\end{pmatrix}, \qquad C_2(D_a) := \begin{pmatrix} 1 & 0\\ 0 & 0_{n_a-1}\end{pmatrix},$$

where $\mathbb{I}_m$ *is the* $m\times m$ *identity matrix and* 0_m *is the* $m\times m$ *zero matrix. Notice that the matrices* C_1, C_2 *satisfy the Kalman rank condition*

$$\operatorname{rank}\{C_2, C_1C_2, \dots, C_1^{n-1}C_2\} = n. \tag{5.12}$$

Analogously, the matrices $C_i(D_a)$ *satisfy* (5.12) *with* $n = n_a$*.*

5.4.1 The Jacobi equations

For any vector field $V(t)$ along an extremal $\lambda(t)$ of the Hamiltonian flow, a dot denotes the Lie derivative in the direction of $\vec{H}$:

$$\dot V(t) := \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} e_*^{-\varepsilon\vec{H}} V(t+\varepsilon).$$

A vector field $\mathcal{J}(t)$ along $\lambda(t)$ is called a *Jacobi field* if it satisfies

$$\dot{\mathcal{J}} = 0. \tag{5.13}$$

The space of solutions of (5.13) is a $2n$-dimensional vector space. The projections $J = \pi_*\mathcal{J}$ are vector fields on M corresponding to one-parameter variations of $\gamma(t) = \pi(\lambda(t))$ through geodesics; in the Riemannian case (without drift field) they coincide with the classical Jacobi fields.

We intend to write (5.13) using the natural symplectic structure σ of T^*M and the canonical frame. First, observe that on T^*M there is a natural smooth sub-bundle of Lagrangian[1] spaces:

$$\mathcal{V}_\lambda := \ker \pi_*|_\lambda = T_\lambda(T^*_{\pi(\lambda)}M).$$

We call this the *vertical subspace*. Then, let $\{E_i(t), F_i(t)\}_{i=1}^n$ be a canonical frame along $\lambda(t)$. The fields $E_1, \dots, E_n$ belong to the vertical subspace. In terms of this frame, $\mathcal{J}(t)$ has components $(p(t), x(t)) \in \mathbb{R}^{2n}$:

$$\mathcal{J}(t) = \sum_{i=1}^n p_i(t)E_i(t) + x_i(t)F_i(t).$$

In turn, the Jacobi equation, written in terms of the components $(p(t), x(t))$, becomes

$$\begin{pmatrix} \dot{p} \\ \dot{x} \end{pmatrix} = \begin{pmatrix} -C_1(t) & -R(t) \\ C_2(t) & C_1(t)^* \end{pmatrix} \begin{pmatrix} p \\ x \end{pmatrix}. \tag{5.14}$$

This is a generalization of the classical Jacobi equation seen as first-order equation for fields on the cotangent bundle. Its structure depends on the Young diagram of the geodesic through the matrices $C_i(D)$, while the remaining invariants are contained in the curvature matrix $R(t)$. Notice that this includes the Riemannian case, where D is the same for every geodesic, with $C_1 = 0$ and $C_2 = \mathbb{I}$.

5.4.2 Geodesic cost and curvature operator

In this section we define the geodesic cost and the curvature operator associated to a geodesic γ. This operator generalizes the idea of sectional curvature.

Definition 5.13. *Let $x_0 \in M$ and consider a ample geodesic γ such that $\gamma(0) = x_0$. The* geodesic cost *associated to γ is the family of functions*

$$c_t(x) := -S_t(x, \gamma(t)), \qquad x \in M, t > 0,$$

where S is the value function defined in (5.10)*.*

Given an ample curve $\gamma(t) = \pi(e^{t\vec{H}}(\lambda))$, the geodesic cost function is smooth in a neighborhood of x_0 and for $t > 0$ sufficiently small. Moreover the differential $d_{x_0}c_t = \lambda$ for every t small, see [3]. Let $\dot{c}_t$ denote the derivative in t of the geodesic cost. Then $\dot{c}_t$ has a critical point in x_0 and its second differential $d^2_{x_0}\dot{c}_t : T_{x_0}M \to \mathbb{R}$ is defined as

$$d^2_{x_0}\dot{c}_t(v) = \left.\frac{d^2}{dt^2}\right|_{t=0} \dot{c}_t(\gamma(t)), \qquad \gamma(0) = x_0, \quad \dot{\gamma}(0) = v.$$

We restrict the second differential of $\dot{c}_t$ to the distribution $\mathcal{D}_{x_0}$ and we define the following family of symmetric operators $\mathcal{Q}_\lambda(t) : \mathcal{D}_{x_0} \to \mathcal{D}_{x_0}$, for t small, associated to $d^2_{x_0}\dot{c}_t$ through the scalar product defined on $\mathcal{D}_x$:

$$d^2_{x_0}\dot{c}_t(v) := \langle \mathcal{Q}_\lambda(t)v|v\rangle_{x_0}, \qquad t > 0, v \in \mathcal{D}_{x_0}. \tag{5.15}$$

[1]A Lagrangian subspace $L \subset \Sigma$ of a symplectic vector space (Σ, σ) is a subspace with $\dim L = \dim \Sigma/2$ and $\sigma|_L = 0$.

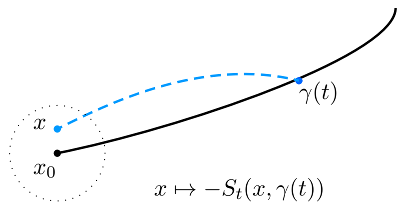


Figure 5.5: The geodesic cost function

The operators $\mathcal{Q}_\lambda(t)$ are smooth for small t. In Theorem A of [3] it is proved their small time asymptotic behavior, which

Theorem 5.14. *Let $\gamma : [0,T] \to M$ be an ample geodesic with initial covector $\lambda \in T^*_{x_0}M$ and let $\mathcal{Q}_\lambda(t) : \mathcal{D}_{x_0} \to \mathcal{D}_{x_0}$ be defined by (5.15). Then $t \mapsto t^2\mathcal{Q}_\lambda(t)$ can be extended to a smooth family of symmetric operators on $\mathcal{D}_{x_0}$ for small $t > 0$. Moreover*

$$\mathcal{I}_\lambda := \lim_{t\searrow 0} t^2\mathcal{Q}_\lambda(t) \geq \mathbb{I} > 0, \qquad \left.\frac{d}{dt}\right|_{t=0} t^2\mathcal{Q}_\lambda(t) = 0,$$

where $\mathbb{I}$ is the k-dimensional identity matrix. In particular, there exists a symmetric operator $\mathcal{R}_\lambda : \mathcal{D}_{x_0} \to \mathcal{D}_{x_0}$ such that

$$\mathcal{Q}_\lambda(t) = \frac{1}{t^2}\mathcal{I}_\lambda + \frac{1}{3}\mathcal{R}_\lambda + O(t), \qquad t > 0. \tag{5.16}$$

Definition 5.15. *We call the symmetric operator $\mathcal{R}_\lambda : \mathcal{D}_{x_0} \to \mathcal{D}_{x_0}$ in (5.16) the* curvature *at λ.*

Let γ be ample and equiregular and let $E(t), F(t)$ be a canonical frame along the lift $\lambda(t)$. Then the curvature operator $\mathcal{R}_\lambda$ can be written in terms of the smooth n-dimensional symmetric matrix $R(t)$, introduced in the canonical equations (5.11). Indeed, for $i, j \in \mathbb{N}$ let

$$\Omega(i,j) = \begin{cases} 0 & |i-j| \geq 2, \\ \frac{1}{4(i+j)} & |i-j| = 1, \\ \frac{i}{4i^2-1} & i = j, \end{cases}$$

then in [3] it is proved that $\mathcal{R}_\lambda$ depends only on the elements of $R_{a1,b1}(0)$ corresponding to the first column of the associated Young diagram, namely

$$(\mathcal{R}_\lambda)_{ab} = 3\Omega(n_a, n_b)R_{a1,b1}(0), \qquad a, b \in \{1, \dots, k\}. \tag{5.17}$$

5.5 The metric measure invariant ρ

Let $\gamma(t) = \pi e^{t\vec{H}}\lambda$ be ample and equiregular, and let $J_\lambda(t)$ be the associated Jacobi curve.

Lemma 5.16 ([3], Lemma 8.3)**.** *Let $\{E_{ai}(t), F_{ai}(t)\}_{ai\in D}$ be a canonical frame along the curve $J_\lambda(t)$. Then the set of vector fields along $\gamma(t)$*

$$X_{ai}(t) := \pi_* F_{ai}(t), \qquad ai \in D$$

is a basis for $T_{\gamma(t)}M$ adapted to the flag $\{\mathcal{F}^i_{\gamma(t)}\}_{i=1}^m$ and $\{X_{a1}(t)\}_{a=1}^k$ is an orthonormal basis for $\mathcal{D}_{\gamma(t)}$ along the geodesic.

Lemma 5.17 ([3], Lemma 8.5)**.** *For $t \in [0,T]$, the projections $X_{ai}(t)$ can be recovered as*

$$X_{ai}(t) = (-1)^{i-1}\mathcal{L}_{\mathsf{T}}^{i-1}(X_{a1}(t)) \operatorname{mod}\mathcal{F}^{i-1}_{\gamma(t)}, \qquad a = 1,\dots,k, \quad i = 1,\dots,n_a.$$

Let $\{\theta_{ai}(t)\}_{ai\in D} \in T^*_{\gamma(t)}M$ be the coframe dual to $X_{ai}(t)$ and define a volume form ω along γ as

$$\omega_{\gamma(t)} := \theta_{1,1}(t) \wedge \theta_{1,2}(t) \wedge \ldots \wedge \theta_{kn_k}(t). \tag{5.18}$$

Given a fixed volume μ on M, let $g_\lambda : [0,T] \to \mathbb{R}$ be the smooth function such that

$$\mu_{\gamma(t)} = e^{g_\lambda(t)}\omega_{\gamma(t)}. \tag{5.19}$$

Lemma 5.18. *Let $\gamma(t) = \pi e^{t\vec{H}}\lambda$ be an ample and equiregular geodesic. Then at the point $\gamma(t)$ it holds*

$$\dot{g}_\lambda(t) = \dot{g}_{\lambda(t)}(0), \qquad \forall t \in [0,T].$$

Proof. Let $\lambda(t) = e^{t\vec{H}}\lambda \in T^*M$ be the lifted extremal and denote by $\gamma_t(s) := \gamma(t+s)$ the rescaled curve. Then $\gamma_t(s) = \pi e^{s\vec{H}}\lambda(t)$. Moreover, if $(E_{\lambda(t+s)}, F_{\lambda(t+s)})$ is a canonical frame along $\lambda(t+s)$, it is a canonical frame also for $e^{s\vec{H}}\lambda(t)$. Then the form $\omega_{\gamma(t+s)}$, defined by wedging the dual forms to $\pi(F_{\lambda(t+s)})$ is equal to $\omega_{\gamma_t(s)}$ for every s where the frame is defined. By the sequence of identities

$$e^{g_\lambda(t+s)}\omega_{\gamma(t+s)} = \mu_{\gamma(t+s)} = \mu_{\gamma_t(s)} = e^{g_{\lambda(t)}(s)}\omega_{\gamma_t(s)}$$

it follows that $g_\lambda(t+s) = g_{\lambda(t)}(s)$ for every s. □

Definition 5.19. *Let $\mathcal{A} \subset T^*M$ be the set of covectors such that the corresponding geodesic is ample and equiregular. We define the function $\rho : \mathcal{A} \to \mathbb{R}$ as*

$$\rho(\lambda) := \dot{g}_\lambda(0).$$

Lemma 5.18 allows to write g as a function of ρ, namely

$$g_\lambda(t) = g_\lambda(0) + \int_0^t \dot{g}_\lambda(s)ds = g(0) + \int_0^t \dot{\rho}(\lambda(s))ds. \tag{5.20}$$

Let T be any admissible extension of $\dot{\gamma}$. By the classical identity $\operatorname{div}_{f\mu}X = \operatorname{div}_\mu X + X(\log f)$ for a volume μ, a smooth function f and a vector field X, we have

$$\rho(\lambda) = \operatorname{div}_\mu \mathsf{T}_x - \operatorname{div}_\omega \mathsf{T}_x \qquad \text{for } \lambda \in T^*_xM. \tag{5.21}$$

Proposition 5.20. *The quantity $\rho(\lambda)$ depends only on the symbol and μ along $\gamma(t)$ in the following sense: if two symbols $S_{\gamma(0)}$ and $S_{\gamma'(0)}$ are equivalent, i.e.*

- *$\phi(\gamma(t)) = \gamma'(t)$ for $t \geq 0$*
- *$\phi_* : \mathcal{D} \to \mathcal{D}'$ is an isometry in a neighborhood of $\gamma(0)$ and $\gamma'(0)$ and $\mathcal{L}_{\mathsf{T}'} \circ \phi_* = \phi_* \circ \mathcal{L}_{\mathsf{T}}$.*

and μ is invariant under ϕ, i.e.

- *$\phi^*\mu_{\gamma'(t)} = \mu_{\gamma(t)}$, for $t \geq 0$,* *then $\rho(\lambda) = \rho(\lambda')$, where λ and λ' are the initial covectors associated to γ and γ' respectively.*

Proof. Let (E, F) and (E', F') be canonical basis with respect to λ and λ' resp., and $X_{ai}(t) = \pi_*(F_{ai}(t))$, $X'_{ai}(t) = \pi_*(F'_{ai}(t))$ be the associated basis of $T_{\gamma(t)}M$ and $T_{\gamma'(t)}M$. Then

$$e^{g_\lambda(t)} = |\mu_{\gamma(t)}(X_{1,1}(t), \dots, X_{kn_k}(t))| \quad \text{and} \quad e^{g_{\lambda'}(t)} = |\mu_{\gamma'(t)}(X'_{1,1}(t), \dots, X'_{kn_k}(t)).$$

Recall that $\{X_{a1}\}_{a=1}^k$ is an orthonormal basis for $\mathcal{D}_{\gamma(t)}$ and the same for $X'_{a1}(t)$ at $\gamma'(t)$. Since ϕ_* is an isometry, there exists a family of orthonormal $k \times k$ matrices $O(t)$ such that

$$X'_{a1}(t) = \sum_{b=1}^k O_{ab}(t)\phi_*(X_{b1}), \qquad \text{for } a = 1, \dots, k.$$

Moreover for the other vector fields we know that

$$\begin{aligned} X'_{ai}(t) =& (-1)^{i-1}\mathcal{L}_{\mathsf{T}'}^{(i-1)}(X'_{a1}(t)) \mathrm{mod} \mathcal{F}^{i-1}_{\gamma'(t)} \\ =& (-1)^{i-1}\mathcal{L}_{\mathsf{T}'}^{(i-1)}\left(\sum_{b=1}^k O(t)_{ab}\phi_*(X_{b1}(t))\right) \mathrm{mod} \mathcal{F}^{i-1}_{\gamma'(t)} \\ =& (-1)^{i-1}\sum_{b=1}^k O(t)_{ab}\mathcal{L}_{\mathsf{T}'}^{(i-1)}\left(\phi_*(X_{b1}(t))\right) \mathrm{mod}\, \mathcal{F}^{i-1}_{\gamma'(t)}, \end{aligned}$$

where the last identity follows by the chain rule. Indeed when we derive the matrix $O(t)$, we obtain elements of $\mathcal{F}^{i-1}_{\gamma'(t)}$. Then

$$\begin{aligned} X'_{ai}(t) =& (-1)^{i-1}\sum_{b=1}^k O(t)_{ab}\phi_*\mathcal{L}_{\mathsf{T}}^{(i-1)}(X_{b1}(t)) \mathrm{mod} \mathcal{F}^{i-1}_{\gamma'(t)} \\ =& \sum_{b=1}^{d_i} O(t)_{ab}\phi_* X_{bi}(t) \,\mathrm{mod} \mathcal{F}^{i-1}_{\gamma'(t)}, \end{aligned}$$

where in the sum we consider only the indices b such that $bi \in \mathrm{D}$. Therefore there exists an orthogonal transformation that sends $\phi_* X_{ai}$ in X'_{ai}. Therefore

$$\begin{aligned} e^{g_{\lambda'}(t)} =& \left|\mu_{\gamma'(t)}\left(X'_{1,1}(t), \dots, X'_{kn_k}\right)\right| = \left|\mu_{\gamma'(t)}\left(\phi_* X_{1,1}(t), \dots, \phi_* X_{kn_k}\right)\right| \\ =& \left|(\phi^*\mu)_{\gamma(t)}\left(X_{1,1}(t), \dots, X_{kn_k}\right)\right| \\ =& e^{g_\lambda(t)} \end{aligned}$$

and the proof is complete. □

Corollary 5.21. *If the symbol is constant through a diffeomorphism ϕ and μ is preserved by ϕ, then $\rho(\lambda(t)) = 0$.*

Eq. (5.20) and the last Proposition say that indeed the whole function $g_\lambda(t)$ depends only on the symbol along the curve (and μ).

Lemma 5.22. *Let $\gamma(t) = \pi \circ \lambda(t)$ be an ample equiregular geodesic. If $e^{t\mathsf{T}}$ is an isometry of the distribution along $\gamma(t)$, then $\mathrm{div}_\omega \mathsf{T}_{\gamma(t)} = 0$ and $\rho(\lambda(t)) = \mathrm{div}_\mu \mathsf{T}$.*

Proof. If we show that $\mathrm{div}_\omega \mathsf{T}_{\gamma(t)} = 0$, then from Eq. (5.21) it immediately follows that $\rho(\lambda(t)) = \mathrm{div}_\mu T$ and ρ depends only on the variation of the volume μ along the curve.

Let $X_{ai}(t)$ be the basis of $\mathcal{D}_{\gamma(t)}$ induced by the canonical frame along the curve $\lambda(t)$. The divergence is computed as

$$\begin{aligned}\mathrm{div}_\omega \mathsf{T}_{\gamma(t)} \omega_{\gamma(t)}(X_{11}(t), \dots, X_{kn_k}(t)) =& \mathcal{L}_\mathsf{T} \omega(X_{11}, \dots, X_{kn_k})_{\gamma(t)} \\ =& \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \omega_{\gamma(\epsilon)}(e_*^{\epsilon\mathsf{T}} X_{11}, \dots, e_*^{\epsilon\mathsf{T}} X_{kn_k}).\end{aligned}$$

Since the flow of T is an isometry of the graded structure that defines the symbol, the last quantity is equal to 0, which proves that $\mathrm{div}_\omega \mathsf{T} = 0$ along the curve. □

Lemma 5.23. *The function ρ is rational in λ.*

Proof. The Hamiltonian $\vec{H}$ is fiber-wise polynomial, therefore for any vector field $V(t) \in T_{\lambda(t)}(T^*M)$, $\dot{V} = [\vec{H}, V]$ is a rational function of the initial covector λ. Then both E and F are rational in λ, and so are also the projections $X(t) = \pi_* F(t)$. We conclude that

$$e^{g_\lambda(t)} = \left|\mu_{\gamma(t)}(X_{a1}(t), \dots, X_{kn_k}(t))\right|$$

is rational in λ. Then all the coefficients of its Taylor expansion are rational in λ. □

5.6 A formula for ρ

In this section we give a formula for ρ in terms only on the volume μ and the maps $\mathcal{L}^i_\mathsf{T}$. It is then once more clear that $\rho(\lambda(t))$ depends only on the symbol of $\gamma(t) = \pi\lambda(t)$ and on μ along the geodesic.

Fix a smooth volume μ on M and let $Y_1, \dots, Y_k$ be an orthonormal basis of $\mathcal{D}$ in a neighborhood of x_0. We complete it to a basis of the tangent space by choosing $Y_{k+1}, \dots, Y_n$ such that $\mu(Y_1, \dots, Y_n) = 1$. We define a scalar product on the whole tangent space, by establishing that this basis is orthonormal. Of curse this scalar product depends on the chosen basis, but this choice does not affect our construction.

Let $\gamma(t) = \pi \circ e^{t\vec{H}}\lambda$ be an ample equiregual curve, with initial covector $\lambda \in T^*_{x_0}M$. Recall that, according to the definition of $g_\lambda(t)$, it holds

$$g_\lambda(t) = \log|\mu(P_t)|, \tag{5.22}$$

where P_t is the parallelotope whose edges are the projections $\{X_{ai}(t)\}_{ai\in D}$ of the horizontal part of the canonical frame $X_{ai} = \pi_* \circ e_*^{t\vec{H}} F_{ai}(t) \in T_{\gamma(t)}M$, namely

$$P_t = \bigwedge_{ai\in D} X_{ai}(t).$$

By Lemma 5.17 we can write the adapted frame $\{X_{ai}\}_{ai\in D}$ in terms of the smooth linear maps $\mathcal{L}_\mathsf{T}$, and we obtain the following formula for the parallelotope

$$P_t = \bigwedge_{i=1}^{m} \bigwedge_{a_i=1}^{d_i} X_{a_i i}(t) = \bigwedge_{i=1}^{m} \bigwedge_{a_i=1}^{d_i} \mathcal{L}_\mathsf{T}^{i-1}(X_{a_i 1}(t)). \tag{5.23}$$

Consider the flag $\{\mathcal{F}^i_{\gamma(t)}\}_{i=1}^m$ and define the following sequence of subspaces of $T_{\gamma(t)}M$: we denote by $V_1 = \mathcal{F}^1_{\gamma(t)}$ the first layer of the flag. By the scalar product on $T_{\gamma(t)}M$ it is well-defined the space $(\mathcal{F}^1_{\gamma(t)})^\perp$, perpendicular to the first layer. Let $V_2 := \mathcal{F}^2_{\gamma(t)} \cap (\mathcal{F}^1_{\gamma(t)})^\perp$. This subspace has dimension $\dim \mathcal{F}^2_{\gamma(t)} - \dim \mathcal{F}^1_{\gamma(t)}$. Going on in this way, for $1 < i \leq m$ let $V_i := \mathcal{F}^i_{\gamma(t)} \cap (\mathcal{F}^{i-1}_{\gamma(t)})^\perp$. It has dimension $\dim \mathcal{F}^i_{\gamma(t)} - \dim \mathcal{F}^{i-1}_{\gamma(t)}$. Therefore there exists a natural isomorphism between $\mathcal{F}^i_{\gamma(t)}/\mathcal{F}^{i-1}_{\gamma(t)}$ and V_i, such that every $Y \in \mathcal{F}^i_{\gamma(t)}/\mathcal{F}^{i-1}_{\gamma(t)}$ is associated with the equivalent element of its class that lies in V_i. In conclusion, for the computation of $g_\lambda(t)$ in (5.22), it is equivalent to substitute the elements $\mathcal{L}^{i-1}_{\mathsf{T}}(X_{a_i1}(t))$ of the parallelotope in (5.23) with the corresponding equivalent element in V_i.

Now recall the map $\mathcal{L}^{i-1}_{\mathsf{T}} : \mathcal{D}_{\gamma(t)} \to \mathcal{F}^i_{\gamma(t)}/\mathcal{F}^{i-1}_{\gamma(t)}$. For every $i = 1, \dots, m$ they descend to an isomorphism $\mathcal{L}^{i-1}_{\mathsf{T}} : \mathcal{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}} \to \mathcal{F}^i_{\gamma(t)}/\mathcal{F}^{i-1}_{\gamma(t)} \simeq V_i$. Then, thanks to the inner product structure, we can consider the map $(\mathcal{L}^{i-1}_{\mathsf{T}})^* \circ \mathcal{L}^{i-1}_{\mathsf{T}} : \mathcal{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}} \to \mathcal{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}}$ obtained by composing $\mathcal{L}^{i-1}_{\mathsf{T}}$ with its adjoint $(\mathcal{L}^{i-1}_{\mathsf{T}})^*$. This composition is a symmetric invertible operator and we define the smooth families of symmetric operators

$$M_i(t) \doteq (\mathcal{L}^{i-1}_{\mathsf{T}})^* \circ \mathcal{L}^{i-1}_{\mathsf{T}} : \mathcal{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}} \to \mathcal{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}}, \qquad i = 1, \dots, m.$$

Recall in particular that for every $v_1, v_2 \in \mathcal{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}}$ it holds the identity

$$\langle (\mathcal{L}^{i-1}_{\mathsf{T}})^* \circ \mathcal{L}^{i-1}_{\mathsf{T}} v_1, v_2 \rangle_{\mathcal{D}_{\gamma(t)}} = \langle \mathcal{L}^{i-1}_{\mathsf{T}} v_1, \mathcal{L}^{i-1}_{\mathsf{T}} v_2 \rangle_{V_i}.$$

By the expression of the parallelotope P_t with elements of the subspaces V_i and the definition of μ as the dual of an orthonormal basis of $T_{\gamma(t)}M$, we have

$$|\mu(P_t)| = |\mu \left(\bigwedge_{i=1}^m \bigwedge_{a_i=1}^{d_i} \mathcal{L}^{i-1}_{\mathsf{T}}(X_{a_i1}(t)) \right)| = \sqrt{\prod_{i=1}^m \det M_i(t)}.$$

Clearly this formula does not depend on the chosen extension $Y_{k+1}, \dots, Y_n$ of the orthonormal basis of $\mathcal{D}$, since the only important point is that the volume μ evaluated at this basis is 1.

Finally to determine $\rho(\lambda)$, recall that it is $\frac{d}{dt}\big|_{t=0} g_\lambda(t) = \frac{d}{dt}\big|_{t=0} \log |\mu(P_t)|$. Then a simple computation shows that

$$\rho(\lambda) = \frac{1}{2} \sum_{i=1}^m \operatorname{tr}\left(M_i(0)^{-1} \dot{M}_i(0) \right).$$

We stress once more, that this last formula is expressed uniquely in terms of the volume μ along the curve and the symbol of $\gamma(t)$.

5.7 Sub-Riemannian manifolds

In this section we consider a sub-Riemannian manifold and we investigate the further properties of the invariant ρ.

The problem of finding the geodesics in a sub-Riemannian manifold, corresponds locally to the optimal control problem

$$\begin{cases} \dot{x} = \sum_{i=1}^{k} u_i(t) f_i(t) \\ J_T(u) = \frac{1}{2}\int_0^T |u(\tau)|^2 d\tau \rightsquigarrow \min, \end{cases}$$

that is an affine control problem, where the drift field is zero. Moreover the fields $f_1, \dots, f_k$ are orthonormal and span a completely non-holonomic distribution. The Hamoltonian function is fiber-wise quadratic in λ and in coordinates it is

$$H(\lambda) = \frac{1}{2}\sum_{i=1}^{k} \langle \lambda, f_i \rangle^2.$$

5.7.1 Homogeneity properties

For all $c > 0$, let $H_c := H^{-1}(c/2)$ be the Hamiltonian level set. In particular H_1 is the unit cotangent bundle: the set of initial covectors associated with unit-speed geodesics. Since the Hamiltonian function is fiber-wise quadratic, we have the following property for any $c > 0$

$$e^{t\vec{H}}(c\lambda) = c e^{ct\vec{H}}(\lambda). \tag{5.24}$$

Let $P_c : T^*M \to T^*M$ be the dilation along the fibers $P_c(\lambda) = c\lambda$. Indeed $\alpha \mapsto P_{e^\alpha}$ is a one-parameter group of diffeomorphisms. Its generator is the *Euler vector field* $\mathfrak{e} \in \Gamma(\mathcal{V})$, and is characterized by $P_c = e^{(\ln c)\mathfrak{e}}$. We can rewrite (5.24) as the following commutation rule for the flows of $\vec{H}$ and $\mathfrak{e}$:

$$e^{t\vec{H}} \circ P_c = P_c \circ e^{ct\vec{H}}.$$

Observe that P_c maps H_1 diffeomorphically on H_c. Let $\lambda \in H_1$ be associated with an ample, equiregular geodesic with Young diagram D. Clearly also the geodesic associated with $\lambda^c := c\lambda \in H_c$ is ample and equiregular, with the same Young diagram. This corresponds to a reparametrization of the same curve: in fact $\lambda^c(t) = e^{t\vec{H}}(c\lambda) = c(\lambda(ct))$, hence $\gamma^c(t) = \pi(\lambda^c(t)) = \gamma(ct)$. The canonical frame associated to $\lambda^c(t)$ can be recovered by the one associated to $\lambda(t)$ as shown in the following Proposition. Its proof can be found in [10].

Proposition 5.24. *Let $\lambda \in H_1$ and $\{E_{ai}, F_{ai}\}_{ai\in D}$ be the associated canonical frame along the extremal $\lambda(t)$. Let $c > 0$ and define, for $ai \in D$*

$$E^c_{ai}(t) := \frac{1}{c^i}(d_{\lambda(ct)}P_c)E_{ai}(ct), \qquad F^c_{ai}(t) := c^{i-1}(d_{\lambda(ct)}P_c)F_{ai}(ct).$$

The moving frame $\{E^c_{ai}(t), F^c_{ai}(t)\}_{ai\in D} \in T_{\lambda^c(t)}(T^*M)$ is a canonical frame associated with the initial covector $\lambda^c = c\lambda \in H_c$, with matrix

$$R^{\lambda^c}_{ai,bj}(t) = c^{i+j} R^{\lambda}_{ai,bj}(ct).$$

By this Proposition, it follows an homogeneity property of the function ρ.

Corollary 5.25. *For every $c > 0$ it holds*

$$e^{g_{c\lambda}(t)} = c^{\mathcal{Q}-n} e^{g_\lambda(ct)},$$

where $\mathcal{Q}$ and n are respectively the homogeneous and topological dimension of the sub-Riemannian manifold. In particular

$$\rho(c\lambda) = c\rho(\lambda) \qquad \forall \lambda \mathcal{A}.$$

Proof. Let $X_{ai}^c(t)$ and $X_{ai}(ct)$ be the basis of $T_{\gamma^c(t)}M = T_{\gamma(ct)}M$ induced by the canonical frame. Then by Proposition 5.24 it holds the identity $X_{ai}^c(t) = c^{i-1}X_{ai}(ct)$. Therefore by the definition of g_λ and $g_{c\lambda}$ we have

$$\begin{aligned} e^{g_{c\lambda}(t)} =& |\mu_{\gamma(ct)}(X_{11}^c(t), \dots, X_{kn_k}^c(t))| \\ =& \prod_{i=1}^{m} \prod_{j=1}^{d_i} c^{i-1} |\mu_{\gamma(ct)}(X_{11}(ct), \dots, X_{kn_k}(ct))| \\ =& c^{\mathcal{Q}-n} e^{g_\lambda(ct)}. \end{aligned}$$

In particular $g_{c\lambda}(t) = g_\lambda(ct) + (\mathcal{Q} - n)\log(c)$ and with a derivation in $t = 0$ we obtain the rescaling property

$$\rho(c\lambda) = c\rho(\lambda) \qquad \forall c > 0.$$

□

5.7.2 Contact manifolds

In this section we focus on the special case of a contact sub-Riemannian manifold. For this type of manifolds, given a geodesic $\gamma(t) = \pi e^{t\vec{H}}\lambda$, it is possible to compute explicitly the value of the associated smooth function $g_\lambda(t)$ and the constant c_0 of Theorem 5.1.

Definition 5.26. *A sub-Riemannian manifold $(M, \mathcal{D}, g)$ of odd dimension $2n+1$ is* contact *if there exists a non degenerate 1-form $\omega \in \Lambda^1(M)$, such that $\mathcal{D}_x = \ker \omega_x$ for every $x \in M$ and $d\omega|_{\mathcal{D}}$ is non degenerate. In this case $\mathcal{D}$ is called* contact distribution.

Since $d\omega|_{\mathcal{D}}$ is not degenerate, the distribution is equiregular of step 2.

Notice moreover that, given a sub-Riemannian contact manifold, the contact form ω is not unique, indeed also $\alpha\omega$ is a contact form for any $0 \neq \alpha \in \mathbb{R}$. Once a contact form ω is fixed we can associate the *Reeb vector field*, X_0, which is the unique vector field such that $\omega(X_0) = 1$ and $d\omega(X_0, \cdot) = 0$. Since the Reeb vector field X_0 is transversal to $\mathcal{D}$, we normalize ω so that $||X_0||_{\mathcal{D}^2/\mathcal{D}} = 1$.

The contact form ω induces a fiber-wise linear map $J : \mathcal{D} \to \mathcal{D}$, defined by

$$\langle JX, Y\rangle = d\omega(X, Y) \qquad \forall X, Y \in \mathcal{D}.$$

Observe that the restriction J_q of J to the fibers of $\mathcal{D}$ is a linear skew-simmetric operator on $(\mathcal{D}_q, \langle\cdot,\cdot\rangle_q)$.

Let $X_1, \dots, X_{2n}$ be a local orthonormal frame of $\mathcal{D}$, then $X_1, \dots, X_{2n}, X_0$ is a local adapted frame to the flag of the distribution. Let $\nu^1, \dots, \nu^{2n}, \nu^0$ be the associated dual frame. The Popp volume μ on M is then the volume obtained by wedging the forms ν^i, i.e.,

$$\mu = \nu^1 \wedge \ldots \wedge \nu^{2n} \wedge \nu^0. \tag{5.25}$$

We compute now the value of the function $g_\lambda(t)$ with respect to the Popp's volume and a given geodesic $\gamma = \pi e^{t\vec{H}}\lambda$. Recall the definition of g_λ in Eq. (5.19). It can be computed as

$$g_\lambda(t) = \log|\mu(P_t)|, \tag{5.26}$$

where P_t is the parallelotope, whose edges are given by the projections on $T_{\gamma(t)}M$ of the fields $F_{ai}(t)$ of a canonical basis along $\lambda(t)$.

Let T be an horizontal extension of the tangent vector field $\dot\gamma(t)$ and consider the map $\mathcal{L}_\mathsf{T} : \mathcal{D}_{\gamma(t)} \to \mathcal{D}^2_{\gamma(t)}/\mathcal{D}_{\gamma(t)}$. Since the manifold is contact, this map is surjective. Its kernel is a subspace of $\mathcal{D}_{\gamma(t)}$ of dimension $2n-1$. So let $X_1, \ldots, X_{2n}$ be an orthonormal basis of $\mathcal{D}_{\gamma(t)}$ such that $X_1 \in (\ker \mathcal{L}_\mathsf{T})^\perp$. Then the other vector fields $X_2, \ldots, X_{2n} \in \ker \mathcal{L}_\mathsf{T}$. Since they are an adapted frame to the flag of the distribution, there exists an orthogonal map that transforms the first $2n$ vectors projections of the canonical basis, in this basis. Then the definition (5.26) of $g_\lambda(t)$ does not change if we take the first $2n$ edges of the parallelotope equal to $X_1, \ldots, X_{2n}$. Moreover, by Lemma 5.17, the last projected vector $X_{ai} = X_{1,2}$ can be written as

$$X_{1,2}(t) = -\mathcal{L}_\mathsf{T} X_1(t) \mathrm{mod}\mathcal{D}.$$

Notice that since X_1 is the only vector of the orthonormal basis, which is not in the kernel of $\mathcal{L}_\mathsf{T}$, this basis is indeed adapted to the Young diagram of γ. Since the Popp's volume can be written as in (5.25) we find that the volume of the parallelotope is equal to the length of the component of $\mathcal{L}_\mathsf{T} X_1(t)$ with respect to X_0, i.e.,

$$|\mu(P_t)| = |\langle [\mathsf{T}, X_1(t)], X_0\rangle_{\gamma(t)}|.$$

This quantity can be written equivalently in terms of the map J. Indeed

$$\begin{aligned}|\mu(P_t)| &= |\langle [\mathsf{T}, X_1], X_0\rangle_{\gamma(t)}| = |\omega_{\gamma(t)}([\mathsf{T}, X_1])| = |d\omega_{\gamma(t)}(\mathsf{T}, X_1)| \\ &= |\langle J_{\gamma(t)}\mathsf{T}, X_1\rangle_{\gamma(t)}|.\end{aligned}$$

Since $\langle J\mathsf{T}, Y\rangle = -\omega(\mathcal{L}_\mathsf{T} Y)$ for every horizontal field Y, then $\ker \mathcal{L}_\mathsf{T} = J\mathsf{T}^\perp$. This implies that $J\mathsf{T}$ is a multiple of X_1, i.e., $J\mathsf{T} = ||J\mathsf{T}||X_1$. Then we simplify the formula for $|\mu(P_t)|$ as

$$|\mu(P_t)| = |\langle J_{\gamma(t)}\mathsf{T}, X_1\rangle_{\gamma(t)}| = ||J\mathsf{T}_{\gamma(t)}||.$$

In particular, if the manifold has dimension 3, then $\ker \mathcal{L}_\mathsf{T}$ has dimension 1 and $\mathsf{T} = ||\mathsf{T}||X_2$. Moreover, if we denote by c_{ij}^k the structure constants such that $[X_i, X_j] = \sum_{k=0}^2 c_{ij}^k X_k$, then the normalization of ω implies $c_{12}^0 = -1$ and

$$\begin{aligned}|\mu(P_t)| &= |\langle [\mathsf{T}, X_1], X_0\rangle| = ||\mathsf{T}_{\gamma(t)}||\,|\langle [X_2, X_1], X_0\rangle| \\ &= ||\mathsf{T}_{\gamma(t)}|| \qquad \text{if } 2n+1 = 3.\end{aligned}$$

The norm of the tangent vector $\dot g(t)$ is constant. This implies that $g(t)$ itself is a constant function and $\rho(\lambda(t)) = 0$ for every t.

For the leading constant c_0 of Theorem 5.1, the computation of its exact value is an easy consequence of formula (5.29). Since the Young diagram is made of $2n$ rows of length 1, except the first one of length 2, the leading constant is $c_0 = \frac{1}{12}$.

5.8 Proof of Theorem 5.1

In this section we prove the following version of Theorem 5.1. Together with the discussion about the function ρ given in the previous sections, the statement follows.

Proposition 5.27. *Let $\gamma(t) = \pi(e^{t\vec{H}}\lambda)$ be an ample equiregular geodesic and let $\omega_{\gamma(t)}$ be the n-form defined in* (5.18). *Given a volume μ on M, define implicitly the smooth function $g : [0,T] \to \mathbb{R}$ by $\mu_{\gamma(t)} = e^{g(t)}\omega_{\gamma(t)}$. Then we have the following Taylor expansion*

$$\left\langle \left(\pi \circ e^{t\vec{H}}\right)^* \mu, E(0)\right\rangle\Big|_\lambda = c_0 t^{\mathcal{N}} e^{g(t)} \left(1 - t^2 \frac{\mathrm{tr}\mathcal{R}_\lambda}{6} + o(t^2)\right) \tag{5.27}$$

where E is the n-dimensional row vector introduced in Remark 5.12 and c_0 is a constant depending only on the structure of the Young diagram. Its value can be found in (5.29).

Proof. The left hand side of the equation can be computed as

$$\left\langle \left(\pi e^{t\vec{H}}\right)^* \mu, E(0)\right\rangle\Big|_\lambda = \left\langle e^{g(t)}\omega, \left(\pi e^{t\vec{H}}\right)_* E(0)\right\rangle\Big|_{\gamma(t)}.$$

For every $ai \in D$, the field $e^{t\vec{H}}_* E_{ai}(0)$ is a Jacobi field, so in coordinates with respect to the canonical frame we can write

$$e^{t\vec{H}}_* E(0) = E(t)M(t) + F(t)N(t)$$

for $n \times n$ matrices M and N, that satisfy the Jacobi equations (5.14). More explicitly we have the system

$$\begin{cases} \dot{N}_{ai,bj} = N_{ai-1,bj} & \text{if } i \neq 1 \\ \dot{N}_{a1,bj} = M_{a1,bj} & \\ \dot{M}_{ai,bj} = -R(t)_{ai,ch}N_{ch,bj} - M_{ai+1,bj} & \text{if } i \neq n_a \\ \dot{M}_{an_a,bj} = -R(t)_{an_a,ch}N_{ch,bj}. & \end{cases} \tag{5.28}$$

Moreover $M(0) = \mathrm{Id}$ and $N(0) = 0$. Clearly the left hand side of (5.27) is equal to $e^{g(t)} \det N(t)$.

In the following we find the Taylor expansion of the matrix $N(t)$ in 0.

Let us begin with the case of a Young diagram made of a single row. For clarity we will avoid the index a in the notation for N and M. Fix integers $1 \leq i, j \leq n$. The elements N_{ij} are successive integrals of M_{1j}. So it is enough to find the asymptotics expansion of M_{1j}. Notice that

$$\begin{aligned} M_{1j}(0) &= \delta_{1j} \\ \dot{M}_{1j} &= -R_{1h}N_{hj} - M_{2j}(1 - \delta_{1n}) \\ M^{(2)}_{1j} &= -\dot{R}_{1h}N_{hj} - \sum_{h\neq 1} R_{1h}N_{h-1j} - R_{11}M_{1j} + (1-\delta_{1n})\left(R_{2h}N_{hj} + M_{3j}(1-\delta_{2n})\right) \end{aligned}$$

In these equations the only non-vanishing component at $t = 0$ is $M_{jj}(0) = 1$, that can be obtained only by deriving the elements M_{ij} with $i < j$. So, in the expansion of $M_{1j}(t)$, the element M_{jj} is obtained for the first time with the $(j-1)$-th derivative. It appears for the second time multiplied by $R_{11}(0)$ at the $(j+1)$-th derivative, since in $M^{(2)}_{1j}$ we have again a component depending on M_{1j}, while all the other components need more derivatives to generate a non vanishing term. We can conclude that the asymptotics of M_{1j} at $t=0$ is

$$M_{1j}(t) = (-1)^{j-1}\frac{t^{j-1}}{(j-1)!} - (-1)^{j-1}R_{11}\frac{t^{j+1}}{(j+1)!} + o(t^{j+1}).$$

Since M_{1j} is the i-th derivative of N_{ij} we have also the expansion for N:

$$N_{ij}(t) = (-1)^{j-1}\frac{t^{i+j-1}}{(i+j-1)!} - (-1)^{j-1}R_{11}\frac{t^{i+j+1}}{(i+j+1)!} + o(t^{i+j+1}).$$

Let us now consider a general distribution of dimension $k>1$. Now we have to study the whole system in (5.28). Fix indeces $ai, bj \in D$. Again, to find N it's enough to determine the expansion of $M_{a1,bj}$, since this is the i-th derivative of $N_{ai,bj}$. To find this expansion, notice that

$$\begin{aligned}
M_{a1,bj}(0) &= \delta_{ab}\delta_{1j} \\
\dot{M}_{a1,bj} &= -R_{a1,ch}N_{ch,bj} - M_{a2,bj}(1-\delta_{1n_a}) \\
M^{(2)}_{a1,bj} &= -\dot{R}_{a1,ch}N_{ch,bj} - \sum_{h\neq 1} R_{a1,ch}N_{ch-1,bj} - R_{a1,c1}M_{c1,bj} \\
&\quad + (1-\delta_{1n_a})\left(R_{a2,ch}N_{ch,bj} + M_{a3,bj}(1-\delta_{2n_a})\right)
\end{aligned}$$

If $a=b$ the argument is similar to the one with $k=1$, but this time with every derivative we generate also terms like $M_{ch,aj}$, that for $c\neq a$ need even more derivatives to give a term different from zero. So we have an expansion as before:

$$\begin{aligned}
M_{a1,aj}(t) &= (-1)^{j-1}\frac{t^{j-1}}{(j-1)!} - (-1)^{j-1}R_{a1,a1}\frac{t^{j+1}}{(j+1)!} + o(t^{j+1}), \\
N_{ai,aj}(t) &= (-1)^{j-1}\frac{t^{i+j-1}}{(i+j-1)!} - (-1)^{j-1}R_{a1,a1}\frac{t^{i+j+1}}{(i+j+1)!} + o(t^{i+j+1}).
\end{aligned}$$

On the other hand, if $a\neq b$, then the first term different from zero of $M_{a1,bj}$ appears at the $j-1+2$ derivative, multiplied by $R_{a1,b1}$, since we need first to generate the element $M_{b1,bj}$, that appears only at the second derivative of $M_{a1,bj}$. Therefore the Taylor expansions of $M_{ai,bj}$ and of a generic element of the matrix N can be derived as

$$\begin{aligned}
M_{a1,bj}(t) &= \delta_{ab}(-1)^{j-1}\frac{t^{j-1}}{(j-1)!} - (-1)^{j-1}R_{a1,b1}\frac{t^{j+1}}{(j+1)!} + o(t^{j+1}), \\
N_{ai,bj}(t) &= \tilde{N}_{ai,bj}t^{i+j-1} - G_{ai,bj}t^{i+j+1} + o(t^{i+j+1}).
\end{aligned}$$

where the matrices $\tilde{N}$ and G are

$$\tilde{N}_{ai,bj} := (-1)^{j-1}\frac{\delta_{ab}}{(i+j-1)!} \qquad \text{and} \qquad G_{ai,bj} := (-1)^{j-1}\frac{R_{a1,b1}}{(i+j+1)!}.$$

Let us come back to equation (5.27). To find the asymptotics of the left hand side, we need only to determine the asymptotic of $\det N(t)$. Let $I_{\sqrt{t}}$ be a diagonal matrix, whose

jj-th element is equal to $\sqrt{t}^{2i-1}$, for $k_{i-1} < j \le i$. Then the Taylor expansion of N can be written as $N(t) = I_{\sqrt{t}}\left(\tilde{N} - t^2 G + O(t^3)\right) I_{\sqrt{t}}$ and its determinant is

$$\det N(t) = \det \tilde{N}\, t^{\mathcal{N}} \left(1 - \operatorname{tr}\left(\tilde{N}^{-1}G\right) t^2 + o(t^2)\right),$$

where $\mathcal{N}$ is the geodesic dimension given in Definition 5.6. Notice that since the matrix $\tilde{N}$ is block-wise diagonal, to find the trace of $\tilde{N}^{-1}G$ we just need the elements of G with $a = b$. Recall Eq. (5.17) that relates the curvature operator $\mathcal{R}_\lambda$ with the elements of the matrix $R_{a1,b1}$. In particular for the diagonal elements it holds

$$\mathcal{R}_{aa} = 3\frac{n_a}{4n_a^2 - 1} R_{a1,a1}(0).$$

So to find the coefficient of order t of the Lemma, it will be enough to show that

$$\begin{aligned}\operatorname{tr}\left(\tilde{N}^{-1}G\right) &= \sum_{a=1}^{k}\left(\sum_{i,j=1}^{n_a}[\tilde{N}]^{-1}_{ai,aj}\frac{(-1)^{j-1}}{(i+j+1)!}\right) R_{a1,a1}(0)\\ &= \sum_{a=1}^{k}\frac{1}{2}\frac{n_a}{4n_a^2-1}R_{a1,a1}(0).\end{aligned}$$

Then the proof will be completed by the following Lemma. □

Lemma 5.28. *The determinat of $\tilde{N}$ is*

$$c_0 := \det \tilde{N} = \prod_{a=1}^{k} \frac{\prod_{j=1}^{n_a-1} j!}{\prod_{j=n_a}^{2n_a-1} j!}. \tag{5.29}$$

Moreover, let $\hat{N}$ and $\hat{G}$ be $n \times n$ matrices, whose elements are $\hat{N}_{ij} = \frac{(-1)^{j-1}}{(i+j-1)!}$ and $\hat{G}_{ij} = \frac{(-1)^{j-1}}{(i+j+1)!}$. Then

$$\operatorname{tr}\left(\hat{N}^{-1}\hat{G}\right) = \frac{1}{2}\frac{n}{4n^2-1}. \tag{5.30}$$

The proof of this Lemma requires only computational discussions and is given in Appendix 5.9.

5.9 Appendix: Proof of Lemma 5.28

In this Section we provide the proof of Lemma 5.28, which will conclude the proof of the main Theorem.

We have to show the value of the leading constant c_0 and the trace of $\hat{N}^{-1}\hat{G}$ for $\hat{N}$ and $\hat{G}$ defined in Lemma 5.28.

The matrix $\hat{N}$ has already been studied in [3], Section 7.3 and Appendix G, and its inverse can be expressed as a product of two matrices $\hat{N}^{-1}_{ij} = \left(\hat{S}^{-1}\hat{A}^{-1}\right)_{ij}$, where

$$\begin{aligned}\hat{A}^{-1}_{ij} &:= \frac{(-1)^{i-j}}{(i-j)!} \qquad i \ge j,\\ \hat{S}^{-1}_{ij} &:= \frac{1}{i+j-1}\binom{n+i-1}{i-1}\binom{n+j-1}{j-1}\frac{(n!)^2}{(n-i)!(n-j)!}.\end{aligned}$$

Therefore the inverse of $\hat{N}$ is

$$\hat{N}^{-1}_{ij} = \sum_{h=j}^{n} \frac{(-1)^{h-j}}{(i+h-1)(h-j)!} \binom{n+i-1}{i-1} \binom{n+h-1}{h-1} \frac{(n!)^2}{(n-i)!(n-h)!}. \tag{5.31}$$

The trace in (5.30) is given by the following sum

$$\operatorname{tr}\left(\hat{N}^{-1}\hat{G}\right) = \sum_{i,j=1}^{n} \hat{N}^{-1}_{ij}\hat{G}_{ji} =$$

$$\sum_{i,j=1}^{n} \sum_{h=j}^{n} \frac{(-1)^{h-j}}{(i+h-1)(h-j)!} \binom{n+i-1}{i-1} \binom{n+h-1}{h-1} \frac{(n!)^2}{(n-i)!(n-h)!} \frac{(-1)^i}{(i+j+1)!}.$$

Notice that for $i = 1, \ldots, n-2$ the element $\hat{C}_{ji} = \hat{N}_{j(i+2)}$ therefore this sum reduces to the sum of the components with $i = n-1$ and $i = n$. In particular we prove our claim if we show that for $i = n-1$ it holds

$$\sum_{j=1}^{n} \sum_{k=j}^{n} \frac{(-1)^{k-j+n}}{(n+k-2)(k-j)!} \binom{2n-2}{n-2} \binom{n+k-1}{k-1} \frac{(n!)^2}{(n-k)!(n+j)!} = -\frac{n-1}{4(2n-1)} \tag{5.32}$$

while for $i = n$

$$\sum_{j=1}^{n} \sum_{k=j}^{n} \frac{(-1)^{k-j+n+1}}{(n+k-1)(k-j)!} \binom{2n-1}{n-1} \binom{n+k-1}{k-1} \frac{(n!)^2}{(n-k)!(n+j+1)!} = \frac{n+1}{4(2n+1)}. \tag{5.33}$$

Let us study equation (5.32). By reordering the factors, equation (5.32) is equivalent to show that

$$\sum_{j=1}^{n} \sum_{k=j}^{n} \frac{(-1)^{n+k+j}}{n+k-2} \binom{2n}{n+k} \binom{n+k}{n+j} \binom{n+k-1}{k-1} =$$

$$= \sum_{k=1}^{n} \left[\sum_{j=1}^{k} (-1)^j \binom{n+k}{n+j} \right] \frac{(-1)^{n+k}}{n+k-2} \binom{2n}{n+k} \binom{n+k-1}{k-1} = -\frac{1}{2}$$

We can solve the sum in the square brackets by a change of index $j' = n+j$ and by using the general identity $0 = (-1+1)^N = \sum_{j=0}^{N} (-1)^j \binom{N}{j}$, then

$$\sum_{j=1}^{k} (-1)^j \binom{n+k}{n+j} = (-1)^n \sum_{j=n+1}^{n+k} (-1)^j \binom{n+k}{j} = -(-1)^n \sum_{j=0}^{n} (-1)^j \binom{n+k}{j}$$
$$= -\binom{n+k-1}{n},$$

where the last equality is justified by the following identity, that can be easily proved by induction on n:

$$\sum_{j=0}^{n} (-1)^j \binom{n+k}{j} = (-1)^n \binom{n+k-1}{n}.$$

To prove equation (5.32) it remains to show that

$$\sum_{k=1}^{n} \frac{(-1)^{n+k}}{n+k-2} \binom{2n}{n-k} \binom{n+k-1}{k-1}^2 = \frac{1}{2}. \tag{5.34}$$

Performing the same kind of transformations, we find that equation (5.33) is equivalent to

$$\sum_{k=1}^{n} \frac{(-1)^{n+k}}{(n+k)(n+k-1)} \binom{2n+1}{n-k} \binom{n+k}{k-1}^2 = \frac{1}{2}. \tag{5.35}$$

Let us study Eq. (5.34). Schechter studied in [45] the $n \times n$ matrices of the form

$$H_{ij} := \frac{1}{a_i - b_j} \tag{5.36}$$

where $a_1, \dots, a_n, b_1, \dots, b_n$ are $2n$ distinct reals. In particular he performs formulas for the inverse of H and for the sum of the coefficients on the i-th row, namely

$$(H^{-1})_{ij} = \frac{1}{b_i - a_j} \frac{\prod_{k=1}^{n} (b_i - a_k)(a_j - b_k)}{\prod_{k \neq j} (a_j - a_k) \prod_{l \neq i} (b_i - b_l)}$$

$$\sum_{j=1}^{n} (H^{-1})_{ij} = -\frac{\prod_{k=1}^{n} (b_i - a_k)}{\prod_{k \neq i} (b_i - b_k)} \tag{5.37}$$

We apply these two formulas to two different matrices: the Hilbert matrix and a modification of it.

Let $H_1 := \left[\frac{1}{i+j-1}\right]$ be the Hilbert matrix. It has the form given in (5.36) by choosing $a_i := i$ and $b_j := -j+1$. We compute the coefficients of the $n-1$-th row of H_1^{-1}:

$$\begin{aligned}(H_1^{-1})_{n-1,j} &= \frac{1}{-n+2-j} \frac{\prod_k (-n+2-k)(j+k-1)}{\prod_{k \neq j} (j-k) \prod_{l \neq n-1} (-n+2+l-1)} \\ &= \frac{n(n-1)^2}{2(2n-1)} \, (n+j) \, \frac{(-1)^{n+j+1}}{n+j-2} \binom{2n}{n-j} \binom{n+j-1}{j-1}^2 \\ &= -\frac{n(n-1)^2}{2(2n-1)} \, (n+j) \, \beta_{j,n}\end{aligned} \tag{5.38}$$

where we denote by $\beta_{j,n}$ the coefficients in the sum in (5.34).

The second matrix is $H_2 := \left[\frac{1}{a_i - b_j}\right]$, with $a_i = i$ for $i < n$ and $a_n = -n$, while $b_j = -j+1$. We compute the coefficients of the $n-1$-th row of H_2^{-1}. For $j < n$ we have

$$\begin{aligned}(H_2^{-1})_{n-1,j} &= \frac{1}{-n+2-j} \frac{\prod_{k \neq n} (-n+2-k)(-n+2+n) \prod_k (j+k-1)}{\prod_{k \neq j, k \neq n} (j-k) \, (n+j) \prod_{l \neq n-1} (-n+2+l-1)} \\ &= \frac{n(n-1)}{2(2n-1)} \, (n-j) \, \frac{(-1)^{n+j+1}}{n+j-2} \binom{2n}{n-j} \binom{n+j-1}{j-1}^2 \\ &= -\frac{n(n-1)}{2(2n-1)} \, (n-j) \, \beta_{j,n}.\end{aligned} \tag{5.39}$$

For $j = n$ we get

$$\begin{aligned}(H_2^{-1})_{n-1,n} =& \frac{1}{-n+2+n} \frac{\prod_{k\neq n}(-n+2-k)(-n+2+n)\prod_k(-n+k-1)}{\prod_{k\neq n}(-n-k)\ (n+j)\prod_{l\neq n-1}(-n+2+l-1)} \\ =& \frac{n^2(n-1)}{2(2n-1)}.\end{aligned}$$

Now, our goal is to compute the sum of $\beta_{j,n}$. However, from equations (5.38) and (5.39) we can determinate the sum of $(n+j)\beta_{j,n}$ and $(n-j)\beta_{j,n}$. So by summing together these two quantities we get the desired one.

Let $\alpha_1 := \sum_{j=1}^n (H_1^{-1})_{n-1,j}$ and $\alpha_2 := \sum_{j=1}^n (H_2^{-1})_{n-1,j}$, then we find

$$\sum_{j=1}^n \beta_{jn} = -\frac{1}{2n}\left[\frac{2(2n-1)}{n(n-1)^2}\alpha_1 + \frac{2(2n-1)}{n(n-1)}\left(\alpha_2 - \frac{n^2(n-1)}{2(2n-1)}\right)\right].$$

Now the proof of equation (5.34) is completed once we use formula (5.37) to find $\alpha_1 = -\frac{(2n-2)!}{(n-2)!^2}$ and $\alpha_2 = \frac{2(2n-3)!}{(n-2)!^2}$.

Equation (5.35) can be solved with the same method. More precisely, let $\gamma_{j,n}$ be the coefficients in the sum (5.35), and consider the same matrix H_1, while H_2 is obtained by fixing $a_j = j$ if $j < n$ and $a_n = -n-1$, and $b_j = -j+1$. Then

$$(H_1^{-1})_{nj} = (n+j+1)\,\frac{n(n+1)^2}{2(2n+1)}\gamma_{j,n}$$

while for $j < n$

$$(H_2^{-1})_{nj} = (n-j)\,\frac{n(n+1)^2}{(2n+1)(2n-1)}\gamma_{j,n}$$

and $(H_2^{-1})_{nn} = -\frac{n(n+1)^2}{2(2n-1)}$.

Let us denote by $\eta_1 := \sum_{j=1}^n (H_1^{-1})_{n,j} = \frac{(2n-1)!}{(n-1)!^2}$ and $\eta_2 := \sum_{j=1}^n (H_2^{-1})_{n,j} = -\frac{2(2n-2)!}{(n-1)!^2}$, then the sum in (5.35) is given by

$$\sum_{j=1}^n \gamma_{j,n} = \frac{1}{2n+1}\left[\frac{2(2n+1)}{n(n+1)^2}\eta_1 + \frac{(2n+1)(2n-1)}{n(n+1)^2}\left(\eta_2 + \frac{n(n+1)^2}{2(2n-1)}\right)\right] = \frac{1}{2}.$$

Finally we have completed the proof of the second statement in the Lemma.

Next we find the value of the leading constant $c_0 := \det \tilde{N}$. This is a block matrix, whose only non vanishing blocks are the one on the diagonal. Moreover, every aa-block of the diagonal is the matrix $\hat{N}$ of dimension n_a. So to find the determinant of $\tilde{N}$, it is sufficient to evaluate the determinant of the generic matrix $\hat{N}$ of dimension n. To this end, let us recall Cramer's rule for the evaluation of an inverse matrix $\hat{N}$:

$$\hat{N}_{ij}^{-1} = (-1)^{i+j}\frac{\det \hat{N}_{ji}^0}{\det \hat{N}},$$

where $\hat{N}_{ji}^0$ is the matrix $\hat{N}$ without the j-th row and the i-th column. So we can compute the determinant of the matrix $\hat{N}_{(n)}$ of dimension n through the (n,n)-entry of the matrix

$\hat{N}^{-1}$ and the value of the determinant of the same matrix $\hat{N}_{(n-1)}$ of dimension $n-1$, namely we get the recursive formula:

$$\det \hat{N}_{(n)} = \frac{\det \hat{N}_{(n-1)}}{\hat{N}^{-1}_{nn}} = \det \hat{N}_{(n-1)} \frac{(n-1)!^2}{(2n-2)!(2n-1)!},$$

where the last equality follows from equation (5.31). When $n = 1$ the determinant is equal to 1, then we can explicitly find

$$\det \hat{N}_{(n)} = \frac{\prod_{j=1}^{n-1} j!}{\prod_{j=n}^{2n-1} j!}.$$

Finally the value of the constant c_0 is then the product

$$c_0 = \prod_{a=1}^{k} \det \hat{N}_{(n_a)} = \prod_{a=1}^{k} \frac{\prod_{j=1}^{n_a-1} j!}{\prod_{j=n_a}^{2n_a-1} j!}.$$

This concludes the proof of Lemma 5.28.

Appendix A

Kolmogorov hypoelliptic operator in dimension 2

In Chapter 4 we have shown a good behavior of the small time asymptotics of the heat kernel on the diagonal, for the class of hypoelliptic operators with constant second order part and linear drift field. An interesting question is whether this expansion can be found also in non-linear operators. To achieve more information on this topic, we show the first terms in the asymptotic expansion for a slightly more general operator, that is the Kolmogorov operator in dimension 2, and we compare it with the curvature operator defined in [2] for the associated affine control system.

A.1 Heat equation in dimension 2 with one controlled vector field

Let f_0 and f_1 be two smooth vector fields on the two dimensional Euclidean space $\mathbb{R}^2$ and consider the following generalization of the heat equation in $\mathbb{R}^2$

$$\frac{\partial \varphi}{\partial t} - f_0 \varphi - \frac{1}{2} f_1^2 \varphi \qquad \forall \varphi \in C_0^2(\mathbb{R} \times \mathbb{R}^2). \tag{A.1}$$

We denote by L the operator $f_0 + \frac{1}{2} f_1^2$ and we call f_0 the *drift* field. If we assume that the fields f_0, f_1 are bounded with bounded derivatives of any order and that they satisfy the condition

$$\operatorname{span}\{f_1, [f_0, f_1]\} = \mathbb{R}^2, \tag{A.2}$$

then the operator (A.1) satisfies the Hörmander condition and admits a fundamental solution $p(t, x, y)$.

Let x_0 be an equilibrium point for the drift field. In this section we compute the small time asymptotic expansion of p at x_0.

Let (x_1, x_2) be coordinates on $\mathbb{R}^2$, centered in x_0, such that f_1 is equal to $\frac{\partial}{\partial x_1}$. The drift field f_0 can be written in these coordinates as

$$f_0 = \alpha_1(x_1, x_2) \frac{\partial}{\partial x_1} + \alpha_2(x_1, x_2) \frac{\partial}{\partial x_2},$$

for smooth functions α_1, α_2, which vanish at the origin.

Condition (A.2) implies that the derivative of α_2 at x_0 doesn't vanish. Indeed, since in x_0 the drift vector field is zero, the bracket $[f_0, f_1]$ in x_0 is just

$$[f_0, f_1]|_{x_0} = -\left.\frac{\partial \alpha_1}{\partial x_1}\right|_{x_0} \frac{\partial}{\partial x_1} - \left.\frac{\partial \alpha_2}{\partial x_1}\right|_{x_0} \frac{\partial}{\partial x_2},$$

then condition (A.2) implies that

$$\left.\frac{\partial \alpha_2}{\partial x_1}\right|_{x_0} \neq 0. \tag{A.3}$$

In the same spirit of Chapter 3 we define the following dilation of factor $\epsilon > 0$ around $x_0 = 0$: for $x = (x_1, x_2)$ let

$$\delta_\epsilon(x_1, x_2) := (\epsilon x_1, \epsilon^3 x_2).$$

Notice that x_0 is a fixed point of δ_ϵ. Under the action of the dilations the vector fields f_0 and f_1 are modified as

$$\begin{aligned}
\left(\delta_{\frac{1}{\epsilon}*}\right) f_1(x) =& \frac{1}{\epsilon}\frac{\partial}{\partial x_1} \\
\left(\delta_{\frac{1}{\epsilon}*}\right) f_0(x) =& \frac{1}{\epsilon^2}\left.\frac{\partial \alpha_2}{\partial x_1}\right|_0 x_1 \frac{\partial}{\partial x_2} + \frac{1}{\epsilon}\cdot\frac{1}{2}\left.\frac{\partial^2 \alpha_2}{\partial x_1^2}\right|_0 x_1^2 \frac{\partial}{\partial x_2} \\
&+ \left.\frac{\partial \alpha_1}{\partial x_1}\right|_0 x_1 \frac{\partial}{\partial x_1} + \left.\frac{\partial \alpha_2}{\partial x_2}\right|_0 x_2 \frac{\partial}{\partial x_2} + \frac{1}{6}\left.\frac{\partial^3 \alpha_2}{\partial x_1^3}\right|_0 x_1^3 \frac{\partial}{\partial x_2} + o(1)
\end{aligned}$$

If we perform these dilations to the Kolmogorov operator in (A.1), we can write it in the following series with respect to ϵ:

$$\begin{aligned}
\frac{\partial}{\partial t} - \epsilon^2 \delta_{\frac{1}{\epsilon}*} L =& \frac{\partial}{\partial t} - \left(\frac{1}{2}\frac{\partial^2}{\partial x_1^2} + \frac{\partial \alpha_2}{\partial x_1} x_1 \frac{\partial}{\partial x_2}\right) - \epsilon\left(\frac{1}{2}\frac{\partial^2 \alpha_2}{\partial x_1^2} x_1^2 \frac{\partial}{\partial x_2}\right) \\
&- \epsilon^2\left(\frac{\partial \alpha_1}{\partial x_1} x_1 \frac{\partial}{\partial x_1} + \frac{\partial \alpha_2}{\partial x_2} x_2 \frac{\partial}{\partial x_2} + \frac{1}{6}\frac{\partial^3 \alpha_2}{\partial x_1^3} x_1^3 \frac{\partial}{\partial x_2}\right) + o(\epsilon^2)\mathcal{Z} \qquad \text{(A.4)} \\
=& \frac{\partial}{\partial t} - L_0 - \epsilon \mathcal{X} - \epsilon^2 \mathcal{Y} + o(\epsilon^2)\mathcal{Z}
\end{aligned}$$

where $\frac{\partial}{\partial t} - L_0$ is the principal part of the operator, that is obtained by taking $\epsilon = 0$, $\mathcal{X}$ and $\mathcal{Y}$ denote the operators in the backets at order ϵ and ϵ^2 respectively, while $\mathcal{Z}$ is a remainder term. As shown in Proposition 3.4 the fundamental solution, $q^\epsilon(s, x, y)$, of this dilated operator can be found as a function of p, namely

$$q^\epsilon(s, x, y) = \epsilon^4 p(\epsilon^2 s, \delta_\epsilon x, \delta_\epsilon y).$$

Notice that if we find the asymptotic expansion of q^ϵ at the diagonal for ϵ that goes to zero, then we obtain the desired small time asymptotics of p. Indeed let us fix in the previous equation $s = 1$, $x = y = 0$ and let ϵ go to zero as $\sqrt{t}$. Then

$$p(t, x_0, x_0) = p(\epsilon^2, \delta_\epsilon x_0, \delta_\epsilon x_0)\big|_{\epsilon=\sqrt{t}} = \frac{1}{t^2} q^{\sqrt{t}}(1, x_0, x_0).$$

Since we have written the dilated operator as a perturbation of a principal part, the asymptotic of q^ϵ with respect to ϵ can be found using Duhamel's formula (3.16), by means of the fundamental solution of L_0. The principal opeartor $\frac{\partial}{\partial t} - L_0$ is linear and hypoelliptic, thanks to condition (A.3), with drift matrix A and a constant matrix B of the second order term equal to

$$A := \begin{pmatrix} 0 & 0 \\ \frac{\partial \alpha_2}{\partial x_1}\Big|_0 & 0 \end{pmatrix} \qquad B := \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Its fundamental solution was computed in Section 2.4 and it is the Gaussian density

$$q_0(t,x,y) = \frac{e^{-\frac{1}{2}(y-e^{At}x)^* D_t^{-1}(y-e^{At}x)}}{2\pi\sqrt{\det D_t}}$$

with covariant matrix $D_t = e^{tA}\int_0^t e^{-\tau A}BB^* e^{-\tau A^*} d\tau e^{tA^*}$.

An iteration of Duhamel's formula for three times, leads to an approximation of q^ϵ in terms of q_0 and the perturbative operators $\mathcal{X}, \mathcal{Y}$, namely

$$q^\epsilon = q_0 + \epsilon q_0 * \mathcal{X} q_0 + \epsilon^2 (q_0 * \mathcal{X} q_0 * \mathcal{X} q_0 + q_0 * \mathcal{Y} q_0) + o(\epsilon^2),$$

where the remainder term is a small o of ϵ^2 as explained in the proof of Theorem 3.20. Here we recall that $a * b$ denotes a convolution in time and space variables of two functions $a(t,x,y)$ and $b(t,x,y)$, more precisely

$$a * b(t,x,y) := \int_0^t \int_{\mathbb{R}^n} a(s,x,z) b(t-s,z,y) dz.$$

To arrive to the asymptotic expansion of $p(t,x_0,x_0)$ we have to find the necessary convolutions of q_0 with its derivatives at the point $(1,0,0)$, namely

$$\begin{aligned} p(t,x_0,x_0) =& \frac{1}{t^2}\Big[q_0(1,0,0) + \sqrt{t} q_0 * \mathcal{X} q_0(1,0,0) \\ &+ t\left(q_0 * \mathcal{X} q_0 * \mathcal{X} q_0 + * \mathcal{Y} q_0\right)(1,0,0) + o(t)\Big]. \end{aligned} \tag{A.5}$$

For the leading term $q_0(1,0,0)$ we just have to evaluate the determinant of D_t in $t=1$, i.e.,

$$\det \begin{pmatrix} t & a\frac{t^2}{2} \\ a\frac{t^2}{2} & a^2\frac{t^2}{3} \end{pmatrix}\Bigg|_{t=0} = \frac{a^2}{12},$$

where $a := \frac{\partial \alpha_2}{\partial x_1}\Big|_0$. For the convolutions, one has to make a careful study of the integrals, and take advantage of the classical moments of an n-dimensional Gaussian function. These computations lead to the following results:

- $q_0 * \mathcal{X} q_0(t,0,0) = q_0 * x_1^2 \frac{\partial q_0}{\partial x_2}(t,0,0) = 0;$
- $q_0 * x_1^2 \frac{\partial q_0}{\partial x_2} * x_1^2 \frac{\partial q_0}{\partial x_2}(t,0,0) = \frac{1}{2\pi\sqrt{\det D_t}} \frac{9}{70} \left(\frac{\partial \alpha_2}{\partial x_1}\Big|_0\right)^{-2} t;$
- $q_0 * x_1 \frac{\partial q_0}{\partial x_1}(t,0,0) = q_0 * x_2 \frac{\partial q_0}{\partial x_2}(t,0,0) = -\frac{1}{2\pi\sqrt{\det D_t}} \frac{t}{2};$
- $q_0 * x_1^3 \frac{\partial q_0}{\partial x_2}(t,0,0) = -\frac{1}{2\pi\sqrt{\det D_t}} \frac{3}{14} \left(\frac{\partial \alpha_2}{\partial x_1}\Big|_0\right)^{-1} t.$

We have then proved the following Theorem:

Theorem A.1. *Let $p(t,x,y)$ be the fundamental solution of the operator in* (A.1) *and let x_0 be an equilibrium point of the drift field f_0. Then for t going to 0 we have the following asymptotic expansion of p:*

$$p(t,0,0) = \frac{\sqrt{12}}{2\pi\left|\frac{\partial\alpha_2}{\partial x_1}\Big|_0\right|}\frac{1}{t^2}\left[1+t\left(-\frac{\operatorname{div} f_0}{2}+\frac{9}{280}\left(\frac{\partial\alpha_2}{\partial x_1}\Big|_0\right)^{-2}\left(\frac{\partial^2\alpha_2}{\partial x_1^2}\Big|_0\right)^2 -\frac{1}{28}\left(\frac{\partial\alpha_2}{\partial x_1}\Big|_0\right)^{-1}\frac{\partial^3\alpha_2}{\partial x_1^3}\Big|_0\right)+o(t)\right]. \tag{A.6}$$

A.2 Relation with geodesic curvature

In this section we want to investigate the geometry behind the asymptotic expansion of the heat kernel for the Kolmogorov operator (A.1). Indeed, let

$$p(t,x_0,x_0) = \frac{c_0}{t^2}\left(1+\sum_{i=1}^{m} a_i(x_0)t^i + O(t^{m+1})\right).$$

In the Riemannian case and in the model class of linear operators studied in Chapter 4, we have seen that the coefficients $a_i(x_0)$ depend on geometric invariants. Let

$$\dot{x} = f_0 + u(t) f_1 \qquad J_u(t) = \frac{1}{2}\int_0^t u_1(\tau)^2 d\tau$$

be the control problem associated to (A.1), with cost function $J_u(t)$ and control $u\in\mathbb{R}$. Here we compute the geodesic curvature $\mathcal{R}_0 = \frac{2}{5}R_{11}$ defined in Chapter 5, for the fixed geodesic $\gamma(t)=x_0$, and we compare it with the coefficients of the asymptotic expansion. We find that terms depending on $\mathcal{R}_0$ in the asymtptotics can appear only in $a_i(x_0)$ for $i\geq 2$. In other words, the first coefficient $a_1(x_0)$ does not derive from $\mathcal{R}_0$.

Let $\lambda(t)$ be the extremal in T^*M such that $\gamma(t)=\pi(\lambda(t))$ and fix a canonical basis $\{E_i(t),F_i(t)\}_{i=1}^2$ of $T_{\lambda(t)}(T^*M)$ (see Chapter 5). Recall by Eq. (5.11) that $R_{11}=\sigma(\dot{F}_1,F_1)$, so if we determine the canonical basis, we find the curvature operator.

In the coordinates $(x_1,x_2)\in\mathbb{R}^2$ adopted in the previous section, let $h_i(\lambda)=\langle\lambda,\frac{\partial}{\partial x_i}\rangle$ for $i=1,2$ be coordinates on T^*M and let $\frac{\partial}{\partial h_i}$ be the associated coordinate vector fields in T^*_xM. Then the Hamiltonian vector field is

$$\vec{h} = \sum_{i=1}^{2}(\alpha_i\vec{h}_i + h_i\vec{\alpha}_i) + h_1\vec{h}_1$$

where

$$\vec{h}_i = \frac{\partial}{\partial x_i} \qquad \vec{\alpha}_i = -\sum_{j=1}^{2}\frac{\partial\alpha_i}{\partial x_j}\frac{\partial}{\partial h_j} \qquad i=1,2.$$

Moreover, by the hamiltonian equations, it follows immediately that the extremal $\lambda(t)$ associated to $\gamma(t)=x_0$ is identically zero, since the drift field is zero in x_0 and the curve doesn't move. So $h_1(\lambda(t))=h_2(\lambda(t))=0$.

The vector fields $E_1(t), E_2(t)$ lie in $T_0(T^*_{x_0}M)$ for every t, so there exist coefficients $v_1(t), v_2(t)$ such that $E_2(t) = v_1(t)\frac{\partial}{\partial h_1} + v_2(t)\frac{\partial}{\partial h_2}$. By the canonical relations (5.11), E_1 is the derivative of E_2, so we find

$$E_1(t) = [\vec{h}, E_2(t)] = \left[\alpha_i \vec{h}_i + h_i \vec{\alpha}_i + h_1 \vec{h}_1, v_j(t)\frac{\partial}{\partial h_j}\right] = v_j \frac{\partial \alpha_i}{\partial x_j}\frac{\partial}{\partial h_i} - v_1 \frac{\partial}{\partial x_1},$$

where we use Einstein summation convention on repeated indices. Since $\pi_*(E_1) = 0$ we find that $v_1(t) = 0$ for every t. Then

$$E_2(t) = v_2(t)\frac{\partial}{\partial h_2}, \qquad E_1(t) = v_2(t)\frac{\partial \alpha_2}{\partial x_i}\frac{\partial}{\partial h_i}.$$

The value of v_2 is found by the normalization condition $1 = \sigma(E_1, F_1) = \sigma(\dot{E}_1, E_1)$. In particular

$$\begin{aligned}\dot{E}_1 = -F_1 &= [\vec{h}, v_2(t)\frac{\partial \alpha_2}{\partial x_i}\frac{\partial}{\partial h_i}] \\ &= \alpha_i v_2 \frac{\partial^2 \alpha_2}{\partial x_i \partial x_j}\frac{\partial}{\partial h_j} + h_1 v_2 \frac{\partial^2 \alpha_2}{\partial x_1 \partial x_j}\frac{\partial}{\partial h_j} + v_2 \frac{\partial \alpha_2}{\partial x_j}\frac{\partial \alpha_j}{\partial x_i}\frac{\partial}{\partial h_i} - v_2 \frac{\partial \alpha_2}{\partial x_1}\frac{\partial}{\partial x_1}.\end{aligned}$$

By evaluating the normalization condition at $\lambda(t) = 0$, we find

$$v_2(t) = \left| \left.\frac{\partial \alpha_2}{\partial x_1}\right|_{x_0} \right|^{-1}.$$

In the same way as in the previous computations, one can determine the derivative of F_1 and find, by the relation $R_{11} = \sigma(\dot{F}_1, F_1)$, that

$$\mathcal{R}_0 = \frac{2}{5}R_{11} = -\frac{2}{5}\left(\left(\frac{\partial \alpha_1}{\partial x_1}\right)^2 + 2\frac{\partial \alpha_1}{\partial x_2}\frac{\partial \alpha_2}{\partial x_1} + \left(\frac{\partial \alpha_2}{\partial x_2}\right)^2\right).$$

Remark A.2. *The coefficients that determine the curvature operator can appear in the asymptotic expansion of the heat kernel* (A.6) *only at order t^2 or greater. Indeed, let us consider again Eq.* (A.4) *and* (A.5). *If we improve the approximation of $\epsilon^2 \delta_{\frac{1}{\epsilon}*}L$ up to order 4, we see that the coefficient $\frac{\partial \alpha_1}{\partial x_2}\frac{\partial \alpha_2}{\partial x_1}$ in $\mathcal{R}_0$ derives from the part of order ϵ^4 of the dilated operator, while the other coefficients of the curvature appear from a multiplication of $\mathcal{Y}$ with itself. Therefore in the asymptotics* (A.5), *coefficients depending on $\mathcal{R}_0$ can be found only at order $\sqrt{t}^4 = t^2$ or greater.*

Bibliography

[1] A. A. Agrachev, D. Barilari, and E. Paoli. Variation of volume and curvature in sub-riemannian geometry. *(in preparation)*.

[2] A. A. Agrachev, D. Barilari, and L. Rizzi. The curvature: a variational approach. *To appear on Memoires AMS*.

[3] A. A. Agrachev, D. Barilari, and L. Rizzi. The curvature: a variational approach. *arXiv:1306.5318 [math.DG]*.

[4] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[5] Andrei Agrachev, Davide Barilari, and Ugo Boscain. Introduction to Riemannian and sub-Riemannian geometry. *Lecture Notes*, 2011. http://people.sissa.it/agrachev/agrachev_files/notes.html.

[6] D. Barilari. Trace heat kernel asymptotics in 3d contact sub-riemannian geometry. *Journal of Mathematical Sciences*, 195(3):391–411, 2013.

[7] D. Barilari, U. Boscain, G. Charlot, and R. W. Neel. On the heat diffusion for generic Riemannian and sub-Riemannian structures. *ArXiv e-prints*, October 2013.

[8] D. Barilari, U. Boscain, and R. W. Neel. Small time heat kernel asymptotics at the sub-Riemannian cut locus. *Journal of Differential Geometry*, 92(3):373–416, 2012.

[9] D. Barilari and E. Paoli. Curvature terms in small time heat kernel expansion for a model class of hypoelliptic Hörmander operators. *preprint on ArXiv*, October 2015.

[10] D. Barilari and L. Rizzi. On Jacobi fields and canonical connection in sub-Riemannian geometry. *ArXiv e-prints*, June 2015.

[11] F. Baudoin and J. Wang. The subelliptic heat kernel on the CR sphere. *Math. Z.*, 275(1-2):135–150, 2013.

[12] F. Baudoin and J. Wang. The subelliptic heat kernels of the quaternionic Hopf fibration. *Potential Anal.*, 41(3):959–982, 2014.

[13] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique hors du cut-locus. *Ann. Sci. École Norm. Sup. (4)*, 21(3):307–331, 1988.

[14] G. Ben Arous. Noyau de la chaleur hypoelliptique et geometrie sous-riemannienne. In Michel Métivier and Shinzo Watanabe, editors, *Stochastic Analysis*, volume 1322 of *Lecture Notes in Mathematics*, pages 1–16. Springer Berlin Heidelberg, 1988.

[15] G. Ben Arous. Développement asymptotique du noyau de la chaleur hypoelliptique sur la diagonale. *Annales de l'institut Fourier*, 39(1):73–99, 1989.

[16] G. Ben Arous, M. Gradinaru, and M. Ledoux. Hölder norms and the support theorem for diffusions. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 30(3):415–436, 1994.

[17] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale (i). *Probability Theory and Related Fields*, 90:175–202, 1991.

[18] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale (ii). *Probability Theory and Related Fields*, 90:377–402, 1991.

[19] M. Berger, P. Gauduchon, and E. Mazet. *Le spectre d'une variété riemannienne.* Lecture Notes in Mathematics, Vol. 194. Springer-Verlag, Berlin, 1971.

[20] R.M. Bianchini and G. Stefani. Graded approximations and controllability along a trajectory. *SIAM J. Control and Optimization*, 28(4):903–924, July 1990.

[21] J.-M. Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 1984.

[22] M. Bramanti, G. Cupini, E. Lanconelli, and E.Priola. Global l^p estimates for degenerate ornstein-uhlenbeck operators with variable coefficients. *Mathematische Nachrichten*, 286:1087–1101, 2013.

[23] C. Cinti, A. Pascucci, and S. Polidoro. Pointwise estimates for solutions to a class of non-homogeneous kolmogorov equations. *Mathematische Annalen*, 340(2):237–264, 2008.

[24] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *J. Differential Geom.*, 259(6):1577–1630, 2010.

[25] A. Friedman. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, 1964.

[26] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths. Theory and Applications.*, volume 120 of *Cambridge Studies of Advanced Mathematics.* Cambridge University Press, 2010.

[27] P. K. Friz and M. Hairer. *A Course on Rough Paths. With an Introduction to Regularity Structures.* Universitext. Springer International Publishing, 2014.

[28] M. Hairer. On malliavin's proof of Hörmander's theorem. *Bull. Sci. Math.*, 135(6-7):650 – 666, 2011.

[29] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[30] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland Kodansha, 1989.

[31] E. Lanconelli, A. Pascucci, and S. Polidoro. Linear and nonlinear ultraparabolic equations of kolmogorov type arising in diffusion theory and in finance. nonlinear problems in mathematical physics and related topics. *Kluwer Academic/Plenum Publisher*, 2:243–265, 2002.

[32] E. Lanconelli and S. Polidoro. On a class of hypoelliptic evolution operators. *Rend. Sem. Mat. Univ. Politec. Torino*, 52(1):29–63, 1994. Partial differential equations, II (Turin, 1993).

[33] R. Léandre. Majoration en temps petit de la densité d'une diffusion dégénérée. *Probab. Theory Related Fields*, 74(2):289–294, 1987.

[34] R. Léandre. Minoration en temps petit de la densité d'une diffusion dégénérée. *J. Funct. Anal.*, 74(2):399–414, 1987.

[35] P. Malliavin. Stochastic calculus of variation and hypoelliptic operators. *Proc. Inter. Symp. SDE*, pages 195 – 263, 1978.

[36] H. P. McKean, Jr. and I. M. Singer. Curvature and the eigenvalues of the laplacian. *J. Differential Geom.*, 1(1-2):43–69, 1967.

[37] A. Millet and M. Sanz-Solé. A simple proof of the support theorem for diffusion processes. *Séminaire de probabilités de Strasbourg*, 28:36–48, 1994.

[38] S. A. Molčanov. Diffusion processes and Riemannian geometry. *Uspehi Mat. Nauk*, 30(1(181)):3–59, 1975.

[39] Richard Montgomery. Abnormal minimizers. *SIAM Journal on Control and Optimization*, 32(6):1605–1620, 1994.

[40] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[41] R. Neel and D. Stroock. Analysis of the cut locus via the heat kernel. In *Surveys in differential geometry. Vol. IX*, Surv. Differ. Geom., IX, pages 337–349. Int. Press, Somerville, MA, 2004.

[42] B. Øksendal. *Stochastic Differential Equations, An Introduction with Applications.* Universitext. Springer, 2003.

[43] E. Paoli. Small time asymptotics on the diagonal for Hörmander's type hypoelliptic operators. *ArXiv e-prints*, February 2015.

[44] S. Rosenberg. *The Laplacian on a Rimennian Manifold*, volume 31 of *London Mathematical Society Student Texts.* Cambridge University Press, 1997.

[45] Samuel Schechter. On the inversion of certain matrices. *Mathematical Tables and Other Aids to Computation*, 13(66):pp. 73–77, 1959.

[46] N. K. Stanton and D. S. Tartakoff. The heat equation for the $\bar{\partial}_b$-Laplacian. *Comm. Partial Differential Equations*, 9(7):597–686, 1984.

[47] D. W. Stroock and S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 3:333–359, 1972.

[48] S. R. S. Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20:431–455, 1967.

[49] I. Zelenko and C. Li. Differential geometry of curves in lagrange grassmannians with given young diagram. *Differential Geometry and its Applications*, 27(6):723 – 742, 2009.

Ringraziamenti

Dopo questi quattro anni di dottorato, ci sono molte persone che vorrei ringraziare per questa bella parte della mia vita:

Il mio professore Andrei Agrachev, che mi ha introdotto e guidato nella matematica, proponendomi problemi interessanti, con passione ed entusiasmo per la ricerca. È sempre stata una qualità che gli ho ammirato.

Davide Barilari, che mi ha accompagnato e incoraggiato. Con lui ho avuto molte discussioni, approfondimenti, domande e risposte, che hanno reso dinamico e avvincente il nostro lavoro.

Tra gli altri matematici il più grande ringraziamento va ai miei compagni di corso: Lucia in primis, compagna anche di ufficio e persona profonda, che mi ha invitato spesso a fare un passo oltre alla matematica, discutendo di temi importanti; e poi Stefano, Gabriele, Luca, Riccardo, Aleks, con Francesca, Silvia, Jessica e Matteo, che hanno reso la Sissa un ambiente familiare e con cui è stato sempre bello poter condividere tanti momenti. Grazie per essere qui con me a festeggiare ora! Ringrazio inoltre Carolina compagna di studio e di casa a Parigi, con cui l'amicizia ha potuto nascere e crescere proprio in questi anni, ti ringrazio per le tante discussioni e confronti avuti insieme.

Tra i non matematici ma tanto amici, ringrazio Francesca e Petr, che mi hanno fatto compagnia e spesso coccolato negli anni di convivenza insieme a Trieste. È stato un bel periodo, in cui conoscersi e condividere tanto insieme.

Ringrazio tutti i miei amici della Sissa, che in questi ultimi mesi mi hanno "adottata": penso in particolare a Ilaria, Marco, Chiara, Ivan, Pavel, ... con cui ho condiviso tanti pranzi, caffè e risate.

Un grande, importante grazie va alla mia famiglia, che mi ha sempre fatto sentire il suo affetto e mi ha sostenuto e incoraggiato nella scelta del dottorato e negli anni a seguire, dandomi forza e coraggio, accompagnandomi ogni giorno. Con loro penso anche alla compagnia di Marina, Giuliano, Stefano, il nonno Nino, e aggiungo anche Elisa e Francesco, che sono un'amicizia splendida, che mi fa crescere.

Infine, ringrazio Marco, che rende felici e piene tutte le mie giornate, e che mi accoglie ogni giorno così come sono, nei momenti di gioia come in quelli più inquieti. Ti ringrazio per avermi lasciata libera anche in queste scelte che hanno comportato la lontananza, facendomi crescere e aspettandomi.