The curvature: a variational approach

A. Agrachev, D. Barilari and L. Rizzi

November 21, 2013

Abstract

The curvature discussed in this paper is a rather far going generalisation of the Riemannian sectional curvature. We define it for a wide class of optimal control problems: a unified framework including geometric structures such as Riemannian, sub-Riemannian, Finsler and sub-Finsler structures; a special attention is paid to the sub-Riemannian (or Carnot–Carathéodory) metric spaces. Our construction of curvature is direct and naive, and it is similar to the original approach of Riemann. Surprisingly, it works in a very general setting and, in particular, for *all* sub-Riemannian spaces.

Contents

1 Introduction **3**
1.1 Structure of the paper . . . 7
1.2 Statements of the main theorems . . . 7

I Statements of the results **9**

2 General setting **9**
2.1 Affine control systems . . . 9
2.1.1 State-feedback equivalence . . . 11
2.2 End-point map . . . 12
2.3 Lagrange multipliers rule . . . 13
2.4 Pontryagin Maximum Principle . . . 14
2.5 Regularity of the value function . . . 15

3 Flag and growth vector of an admissible curve **16**
3.1 Growth vector of an admissible curve . . . 16
3.2 Linearised control system and growth vector . . . 18
3.2.1 Linearisation of a control system in $\mathbb{R}^n$ . . . 18
3.2.2 Linearisation of a control system in the general setting . . . 19
3.2.3 Growth vector and controllability . . . 20
3.3 State-feedback invariance of the flag of an admissible curve . . . 21
3.4 An alternative definition . . . 22

4 Geodesic cost and its asymptotics **24**
4.1 Geodesic cost . . . 24
4.2 Hamiltonian inner product . . . 25
4.3 Asymptotics of the geodesic cost function and curvature . . . 27
4.3.1 Spectrum of $\mathcal{I}_\lambda$ for equiregular geodesics . . . 28

4.4 Examples . . . 29
4.4.1 Linear-quadratic control problem . . . 29
4.4.2 Riemannian geometry . . . 30
4.4.3 Finsler geometry . . . 32

5 Sub-Riemannian geometry **32**
5.1 Basic definitions . . . 32
5.1.1 Nilpotent approximation and privileged coordinates . . . 34
5.1.2 Approximating trajectories . . . 36
5.2 Ample geodesics and regularity of the squared distance . . . 36
5.3 Reparametrization and homogeneity of the curvature operator . . . 38
5.4 Asymptotics of the sub-Laplacian of the geodesic cost . . . 38
5.5 Equiregular distributions . . . 41
5.5.1 Popp's volume . . . 41
5.5.2 Slow growth distributions . . . 42
5.6 Geodesic dimension and sub-Riemannian homotheties . . . 44
5.7 Heisenberg group . . . 46
5.7.1 Distance in the Heisenberg group . . . 47
5.7.2 Asymptotic expansion of the distance . . . 48
5.7.3 Second differential of the geodesic cost . . . 49
5.7.4 Sub-Laplacian of the geodesic cost . . . 50

II Technical tools and proofs **51**

6 Jacobi curves **51**
6.1 Curves in the Lagrange Grassmannian . . . 51
6.1.1 Ample, equiregular, monotone curves . . . 52
6.1.2 The Young diagram of an equiregular Jacobi curve . . . 53
6.2 The Jacobi curve and the second differential of the geodesic cost . . . 53
6.2.1 Second differential at a non critical point . . . 53
6.2.2 Second differential of the geodesic cost . . . 54
6.3 The Jacobi curve and the Hamiltonian inner product . . . 56
6.4 Proof of Theorem A . . . 57
6.4.1 Coordinate presentation of the Jacobi curve . . . 57
6.5 Proof of Theorem D . . . 61

7 Asymptotics of the Jacobi curve: equiregular case **62**
7.1 The canonical frame . . . 63
7.1.1 A remark on the notation . . . 63
7.2 Main result . . . 64
7.2.1 Restriction . . . 65
7.2.2 A remark on the coefficients . . . 66
7.2.3 Examples . . . 66
7.3 Proof of Theorem 7.2 . . . 67
7.3.1 Rank 1 curve with vanishing $R(t)$ . . . 67
7.3.2 General rank 1 curve . . . 68
7.3.3 General rank k curve . . . 69
7.3.4 Proof of Corollary 7.3 . . . 70
7.4 Proof of Theorem B . . . 71

8 Sub-Laplacian and Jacobi curves **72**
8.1 Coordinate lift of a local frame 72
8.2 Sub-Laplacian of the geodesic cost 73
8.2.1 Computation of the linear term 76

III Appendix **78**

A Proof of Proposition 7.5 **78**

B Proof of Lemma 7.6 **79**

C Proof of Theorem 2.19 **81**

D Proof of Proposition 5.14 **84**

E Proof of Proposition 5.42 **84**

1 Introduction

The curvature discussed in this paper is a rather far going generalization of the Riemannian sectional curvature. We define it for a wide class of optimal control problems: a unified framework including geometric structures such as Riemannian, sub-Riemannian, Finsler and sub-Finsler structures; a special attention is paid to the sub-Riemannian (or Carnot–Carathéodory) metric spaces. Our construction of the curvature is direct and naive, and it is similar to the original approach of Riemann. Surprisingly, it works in a very general setting and, in particular, for *all* sub-Riemannian spaces.

Interesting metric spaces often appear as limits of families of Riemannian metrics. We first try to explain the nature of our curvature by describing the case of a contact sub-Riemannian structure in terms of such a family and then we move to the general construction.

Let M be an odd-dimensional Riemannian manifold endowed with a contact vector distribution $\mathscr{D} \subset TM$. Given $x_0, x_1 \in M$, the contact sub-Riemannian distance $\mathsf{d}(x_0, x_1)$ is the infimum of the lengths of Legendrian curves connecting x_0 and x_1. Recall that Legendrian curves are just integral curves of the distribution $\mathscr{D}$. The metric d is easily realized as the limit of a family of Riemannian metrics d^ε as $\varepsilon \to 0$. To define d^ε we start from the original Riemannian structure on M, keep fixed the length of vectors from $\mathscr{D}$ and multiply by $\frac{1}{\varepsilon}$ the length of the orthogonal to $\mathscr{D}$ tangent vectors to M. It is easy to see that $\mathsf{d}^\varepsilon \to \mathsf{d}$ uniformly on compacts in $M \times M$ as $\varepsilon \to 0$.

The distance converges, what about the curvature? Let ω be a contact differential form that annihilates $\mathscr{D}$, i.e. $\mathscr{D} = \omega^\perp$. Given $v_1, v_2 \in T_xM,\ v_1 \wedge v_2 \neq 0$, we denote by $K^\varepsilon(v_1 \wedge v_2)$ the sectional curvature for the metric d^ε and section $\mathrm{span}\{v_1, v_2\}$. It is not hard to show that $K^\varepsilon(v_1 \wedge v_2) \to -\infty$ if $v_1, v_2 \in \mathscr{D}$ and $d\omega(v_1, v_2) \neq 0$. Moreover, $\mathrm{Ric}^\varepsilon(v) \to -\infty$ as $\varepsilon \to 0$ for any nonzero vector $v \in \mathscr{D}$, where Ric^ε is the Ricci curvature for the metric d^ε. On the other hand, the distance between x and the conjugate locus of x tends to 0 as $\varepsilon \to 0$ and $K^\varepsilon(v_1 \wedge v_2)$ tends to $+\infty$ for some $v_1, v_2 \in T_xM$, as well as $\mathrm{Ric}^\varepsilon(v)$ for some $v \in T_xM$.

What about the geodesics? For any $\varepsilon > 0$ and any $v \in T_xM$ there is a unique geodesic of the Riemannian metric d^ε that starts from x with velocity v. On the other hand, the velocities of all geodesics of the limit metric d belong to $\mathscr{D}$ and for any nonzero vector $v \in \mathscr{D}$ there exists a one-parametric family of geodesics whose initial velocity is equal to v. Too bad up to now, and here is the first encouraging fact: the family of geodesic flows converges if we re-write it as a family of flows on the cotangent bundle.

Indeed, any Riemannian structure on M induces a self-adjoint isomorphism $G : TM \to T^*M$, where $\langle Gv, v\rangle$ is the square of the length of the vector $v \in TM$, and $\langle\cdot,\cdot\rangle$ denotes the standard pairing between tangent and cotangent vectors. The geodesic flow, treated as flow on T^*M is a Hamiltonian flow associated with the Hamiltonian function $H : T^*M \to \mathbb{R}$, where $H(\lambda) = \langle\lambda, G^{-1}\lambda\rangle,\ \lambda \in T^*M$. Let $(\lambda(t), \gamma(t))$ be a trajectory of the Hamiltonian flow, with $\lambda(t) \in T^*_{\gamma(t)}M$. The square of the Riemannian distance from x_0 is a smooth function on a neighbourhood of x_0 in M and the differential of this function at $\gamma(t)$ is equal to $2t\lambda(t)$ for any small $t \geq 0$. Let H^ε be the Hamiltonian corresponding to the metric d^ε. It is easy to see that H^ε converges with all derivatives to a Hamiltonian H^0. Moreover, geodesics of the limit sub-Riemannian metric are just projections to M of the trajectories of the Hamiltonian flow on T^*M associated to H^0.

We can recover the Riemannian curvature from the asymptotic expansion of the square of the distance from x_0 along a geodesic: this is essentially what Riemann did. Then we can write a similar expansion for the square of the limit sub-Riemannian distance to get an idea of the curvature in this case. Note that the metrics d^ε converge to d with all derivatives in any point of $M \times M$, where d is smooth. The metrics d^ε are not smooth at the diagonal but their squares are smooth. The point is that no power of d is smooth at the diagonal! Nevertheless, the desired asymptotic expansion can be controlled.

Fix a point $x_0 \in M$ and $\lambda_0 \in T^*_{x_0}M$ such that $\langle\lambda_0, \mathscr{D}\rangle \neq 0$. Let $(\lambda^\varepsilon(t), \gamma^\varepsilon(t))$, for $\varepsilon \geq 0$, be the trajectory of the Hamiltonian flow associated to the Hamiltonian H^ε and initial condition (λ_0, x_0). We set:

$$c_t^\varepsilon(x) \doteq -\frac{1}{2t}(\mathsf{d}^\varepsilon)^2(x, \gamma^\varepsilon(t)) \text{ if } \varepsilon > 0, \qquad c_t^0(x) \doteq -\frac{1}{2t}\mathsf{d}^2(x, \gamma^0(t)).$$

There exists an interval $(0, \delta)$ such that the functions c_t^ε are smooth at x_0 for all $t \in (0, \delta)$ and all $\varepsilon \geq 0$. Moreover, $d_{x_0}c_t^\varepsilon = \lambda_0$. Let $\dot{c}_t^\varepsilon = \frac{\partial}{\partial t}c_t^\varepsilon$, then $d_{x_0}\dot{c}_t^\varepsilon = 0$. In other words, x_0 is a critical point of the function $\dot{c}_t^\varepsilon$ and its Hessian $d^2_{x_0}\dot{c}_t^\varepsilon$ is a well-defined quadratic form on $T_{x_0}M$. Recall that $\varepsilon = 0$ is available, but t must be positive. We are going to study the asymptotics of the family of quadratic forms $d^2_{x_0}\dot{c}_t^\varepsilon$ as $t \to 0$ for fixed ε. This asymptotic is a little bit different for $\varepsilon > 0$ and $\varepsilon = 0$. The difference reflects the structural difference of the Riemannian and sub-Riemannian metrics and emphasises the role of the curvature.

Given $v, w \in T_xM,\ \varepsilon > 0$, we denote $\langle v|w\rangle_\varepsilon = \langle G^\varepsilon v, w\rangle$ the inner product generating d^ε. Recall that $\langle v|v\rangle_\varepsilon$ does not depend on ε if $v \in \mathscr{D}$ and $\langle v|v\rangle_\varepsilon \to \infty$ $(\varepsilon \to 0)$ if $v \notin \mathscr{D}$; we will write $|v|^2 \doteq \langle v|v\rangle_\varepsilon$ in the first case. For fixed $\varepsilon > 0$, we have:

$$d^2_{x_0}\dot{c}_t^\varepsilon(v) = \frac{1}{t^2}\langle v|v\rangle_\varepsilon + \frac{1}{3}\langle R^\varepsilon(\dot{\gamma}^\varepsilon, v)\dot{\gamma}^\varepsilon)|v\rangle_\varepsilon + O(t), \qquad v \in T_{x_0}M,$$

where $\dot{\gamma}^\varepsilon = \dot{\gamma}^\varepsilon(0)$ and R^ε is the Riemannian curvature tensor of the metric d^ε. For $\varepsilon = 0$, only vectors $v \in \mathscr{D}$ have a finite length and the above expansion is modified as follows:

$$d^2_{x_0}\dot{c}_t^0(v) = \frac{1}{t^2}\mathcal{I}_\gamma(v) + \frac{1}{3}\mathcal{R}_\gamma(v) + O(t), \qquad v \in \mathscr{D} \cap T_{x_0}M,$$

where $\mathcal{I}_\gamma(v) \geq |v|^2$ and $\mathcal{R}_\gamma$ is the *sub-Riemannian curvature* at x_0 along the geodesic $\gamma = \gamma^0$. Both $\mathcal{I}_\gamma$ and $\mathcal{R}_\gamma$ are quadratic forms on $\mathscr{D}_{x_0} \doteq \mathscr{D} \cap T_{x_0}M$. The principal "structural" term $\mathcal{I}_\gamma$ has the following properties:

$$\max\{\mathcal{I}_\gamma(v) |\, v \in \mathscr{D}_{x_0},\ |v|^2 = 1\} = 4,$$
$$\mathcal{I}_\gamma(v) = |v|^2 \text{ if and only if } d\omega(v, \dot{\gamma}(0)) = 0.$$

In other words, the symmetric operator on $\mathscr{D}_{x_0}$ associated with the quadratic form $\mathcal{I}_\gamma$ has eigenvalue 1 of multiplicity $\dim \mathscr{D}_{x_0} - 1$ and eigenvalue 4 of multiplicity 1. The trace of this operator, which, in this case, does not depend on γ, equals $\dim \mathscr{D}_{x_0} + 3$. This trace has a simple geometric interpretation, it is equal to the *geodesic dimension* of the sub-Riemannian space.

The geodesic dimension is defined as follows. Let $\Omega \subset M$ be a bounded and measurable subset of positive volume and let $\Omega_{x_0,t}$, for $0 \le t \le 1$, be a family of subsets obtained from Ω by the homothety of Ω with respect to a fixed point x_0 along the shortest geodesics connecting x_0 with the points of Ω, so that $\Omega_{x_0,0} = \{x_0\}$, $\Omega_{x_0,1} = \Omega$. The volume of $\Omega_{x_0,t}$ has order $t^{\mathcal{N}_{x_0}}$, where $\mathcal{N}_{x_0}$ is the geodesic dimension at x_0 (see Section 5.6 for details).

Note that the topological dimension of our contact sub-Riemannian space is $\dim \mathscr{D}_{x_0} + 1$ and the Hausdorff dimension is $\dim \mathscr{D}_{x_0} + 2$. All three dimensions are obviously equal for Riemannian or Finsler manifolds. The structure of the term $\mathcal{I}_\gamma$ and comparison of the asymptotic expansions of $d^2_{x_0} \dot{c}^\varepsilon_t$ for $\varepsilon > 0$ and $\varepsilon = 0$ explains why sectional curvature goes to $-\infty$ for certain sections.

The curvature operator which we define can be computed in terms of the symplectic invariants of the so-called Jacobi curve, namely a curve in the Lagrange Grassmannian related with the linearisation of the Hamiltonian flow. These symplectic invariants can be computed, in principle, via an algorithm which is, however, quite hard to implement. Explicit computations of the contact sub-Riemannian curvature will appear in a forthcoming paper. The current paper deals with the general setting. A precise construction in full generality is presented in the forthcoming sections but, since the paper is long, we find it worth to briefly describe the main ideas in the introduction.

Let M be a smooth manifold, $\mathscr{D} \subset TM$ be a vector distribution (not necessarily contact), f_0 be a vector field on M and $L : TM \to M$ be a Tonelli Lagrangian (i.e. $L|_{T_xM}$ has a superlinear growth and its Hessian is positive definite for any $x \in M$). *Admissible paths* on M are curves whose velocities belong to the "affine distribution" $f_0 + \mathscr{D}$. Let $\mathcal{A}_t$ be the space of admissible paths defined on the segment $[0,t]$ and $N_t = \{(\gamma(0), \gamma(t)) : \gamma \in \mathcal{A}_t\} \subset M \times M$. The optimal cost (or action) function $S_t : N_t \to \mathbb{R}$ is defined as follows:

$$S_t(x,y) = \inf \left\{ \int_0^t L(\dot\gamma(\tau))\, d\tau : \gamma \in \mathcal{A}_t,\ \gamma(0) = x,\ \gamma(t) = y \right\}.$$

The space $\mathcal{A}_t$ equipped with the $W^{1,\infty}$-topology is a smooth Banach manifold; the functional $J_t : \gamma \mapsto \int_0^t L(\dot\gamma(\tau))\, d\tau$ and the evaluation maps $F_\tau : \gamma \mapsto \gamma(\tau)$ are smooth on $\mathcal{A}_t$.

The optimal cost $S_t(x,y)$ is the solution of the conditional minimum problem for the functional J_t under conditions $F_0(\gamma) = x$, $F_t(\gamma) = y$. The Lagrange multipliers rule for this problem reads:

$$d_\gamma J_t = \lambda_t D_\gamma F_t - \lambda_0 D_\gamma F_0. \tag{1}$$

Here λ_t and λ_0 are "Lagrange multipliers", $\lambda_t \in T^*_{\gamma(t)}M$, $\lambda_0 \in T^*_{\gamma(0)}M$. We have:

$$D_\gamma F_t : T_\gamma \mathcal{A}_t \to T_{\gamma(t)}M, \qquad \lambda_t : T_{\gamma(t)}M \to \mathbb{R},$$

and the composition $\lambda_t D_\gamma F_t$ is a linear functional on $T_\gamma \mathcal{A}_t$. Moreover, Eq. (1) implies that

$$d_\gamma J_\tau = \lambda_\tau D_\gamma F_\tau - \lambda_0 D_\gamma F_0, \tag{2}$$

for some $\lambda_\tau \in T^*_{\gamma(\tau)}M$ and any $\tau \in [0,t]$. The curve $\tau \mapsto \lambda_\tau$ is a trajectory of the Hamiltonian system associated to the Hamiltonian $H : T^*M \to \mathbb{R}$ defined by

$$H(\lambda) = \max_{v \in f_0(x) + \mathscr{D}_x} \left(\langle \lambda, v \rangle - L(v) \right), \qquad \lambda \in T^*_xM,\ x \in M.$$

Moreover, any trajectory of this Hamiltonian system satisfies relation (2), where γ is the projection of the trajectory to M. Trajectories of the Hamiltonian system are called *normal extremals* and their projections to M are called *normal extremal trajectories*.

We recover the sub-Riemannian setting when $f_0 = 0$, $L(v) = \frac{1}{2}\langle Gv, v\rangle$. In this case, the optimal cost is related with the sub-Riemannian distance $S_t(x,y) = \frac{1}{2t}\mathsf{d}^2(x,y)$, and normal extremal trajectories are normal sub-Riemannian geodesics.

Let γ be an admissible path; the germ of γ at the point $x_0 = \gamma(0)$ defines a flag in $T_{x_0}M$ $\{0\} = \mathscr{F}^0_\gamma \subset \mathscr{F}^1_\gamma \subset \mathscr{F}^2_\gamma \subset \ldots \subset T_{x_0}M$ in the following way. Let V be a section of the vector distribution $\mathscr{D}$ such that $\dot\gamma(t) = f_0(\gamma(t)) + V(\gamma(t))$, $t \geq 0$, and $P_{0,t}$ be the local flow on M generated by the vector field $f_0 + V$; then $\gamma(t) = P_{0,t}(\gamma(0))$. We set:

$$\mathscr{F}^i_\gamma = \operatorname{span}\left\{ \left.\frac{d^j}{dt^j}\right|_{t=0} (P_{0,t})^{-1}_* \mathscr{D}_{\gamma(t)} : j = 0, \ldots, i-1 \right\}.$$

The flag $\mathscr{F}^i_\gamma$ depends only on the germs of $f_0 + \mathscr{D}$ and γ at the initial point x_0.

A normal extremal trajectory γ is called *ample* if $\mathscr{F}^m_\gamma = T_{x_0}M$ for some $m > 0$. If γ is ample, then $J_t(\gamma) = S_t(x_0, \gamma(t))$ for all sufficiently small $t > 0$ and S_t is a smooth function in a neighbourhood of $(\gamma(0), \gamma(t))$. Moreover, $\frac{\partial S_t}{\partial y}\big|_{y=\gamma(t)} = \lambda_t$, $\frac{\partial S_t}{\partial x}\big|_{x=\gamma(0)} = -\lambda_0$, where λ_t is the normal extremal whose projection is γ.

We set $c_t(x) \doteq -S_t(x, \gamma(t))$; then $d_{x_0}c_t = \lambda_0$ for any $t > 0$ and x_0 is a critical point of the function $\dot c_t$. The Hessian of this function $d^2_{x_0}\dot c_t$ is a well-defined quadratic form on $T_{x_0}M$. We are going to write an asymptotic expansion of $d^2_{x_0}\dot c_t|_{\mathscr{D}_{x_0}}$ as $t \to 0$ (see Theorem A):

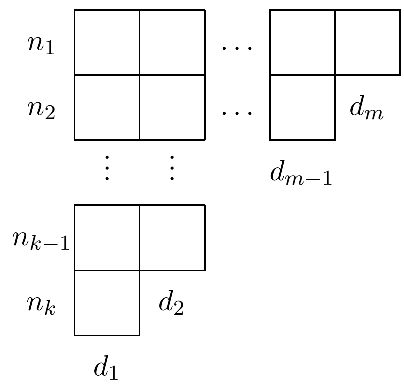
$$d^2_{x_0}\dot c_t(v) = \frac{1}{t^2}\mathcal{I}_\gamma(v) + \frac{1}{3}\mathcal{R}_\gamma(v) + O(t), \qquad \forall v \in \mathscr{D}_{x_0}.$$

Now we introduce a natural Euclidean structure on $T_{x_0}M$. Recall that $L|_{T_{x_0}M}$ is a strictly convex function, and $d^2_w(L|_{T_{x_0}M})$ is a positive definite quadratic form on T_xM, $\forall\, w \in T_xM$. If we set $|v|^2_\gamma = d^2_{\dot\gamma(0)}(L|_{T_{x_0}M})(v)$, $v \in T_{x_0}M$ we have the inequality

$$\mathcal{I}_\gamma(v) \geq |v|^2_\gamma, \qquad \forall\, v \in \mathscr{D}_{x_0}.$$

The inequality $\mathcal{I}_\gamma(v) \geq |v|^2_\gamma$ means that the eigenvalues of the symmetric operator on $\mathscr{D}_{x_0}$ associated with the quadratic form $\mathcal{I}_\gamma$ are greater or equal than 1. The quadratic form $\mathcal{R}_\gamma$ is the *curvature* of our constrained variational problem along the extremal trajectory γ.

A mild regularity assumption allows to explicitly compute the eigenvalues of $\mathcal{I}_\gamma$. We set $\gamma_\varepsilon(t) = \gamma(\varepsilon + t)$ and assume that $\dim \mathscr{F}^i_{\gamma_\varepsilon} = \dim \mathscr{F}^i_\gamma$ for all sufficiently small $\varepsilon \geq 0$ and all i. Then $d_i = \dim \mathscr{F}^i_\gamma - \dim \mathscr{F}^{i-1}_\gamma$, for $i \geq 1$ is a non-increasing sequence of natural numbers with $d_1 = \dim \mathscr{D}_{x_0} = k$. We draw a Young tableau with d_i blocks in the i-th column and we define $n_1, \ldots, n_k$ as the lengths of its rows (that may depend on γ).



The eigenvalues of the symmetric operator $\mathcal{I}_\gamma$ are $n_1^2, \dots, n_k^2$ (see Theorem B). Some of these numbers may be equal (in the case of multiple eigenvalues) and are all equal to 1 in the Riemannian case. In the sub-Riemannian setting, the trace of $\mathcal{I}_\gamma$ is

$$\operatorname{tr} \mathcal{I}_\gamma = n_1^2 + \cdots + n_k^2 = \sum_{i=1}^{m} (2i-1) d_i,$$

when computed for generic normal sub-Riemannian geodesic, is equal to the geodesic dimension of the space (see Theorem D).

The construction of the curvature presented here was preceded by a rather long research line (see [1,5,8,10,22,36]). For what concerns the alternative approaches to this topic, in recent years, several efforts have been made to introduce a notion of curvature to non-Riemannian situations, such as sub-Riemannian manifolds and, more in general, metric measure spaces. Motivated by the lack of classical Riemannian tools (such as the Levi-Civita connection and the theory of Jacobi fields) different approaches have been explored in order to extend some classical results in geometric analysis to such structures. In particular, to this extent, different notions of generalized Ricci curvature bound have been introduced.

For instance, one can see [13,14] and references therein for a heat equation approach to the generalization of the curvature-dimension inequality and [11,24,33,34] and references therein for an optimal transport approach to the generalization of Ricci curvature.

1.1 Structure of the paper

In Sections 2–4 we give a detailed exposition of the main constructions in a more general and flexible setting than in this introduction. Section 5 is devoted to the specification to the case of sub-Riemannian spaces and to some further results: an estimate of the Young tableau in terms of the nilpotent approximation (Lemma 5.15), an asymptotic expansion of the sub-Laplacian applied to the square of the distance (Theorem C), the computation of the geodesic dimension (Theorem D).

Before entering into details of the proofs, we end Section 5 by repeating our construction for one of the simplest sub-Riemannian structures: the Heisenberg group. In particular, we recover by a direct computation the results of Theorems A, B and C.

The proofs of the main results are concentrated in Sections 6–8 where we introduce the main technical tools: Jacobi curves, their symplectic invariants and Li–Zelenko structural equations.

1.2 Statements of the main theorems

The main results, namely Theorems A, B, C and D, are spread in Part I of the paper. For convenience of the reader we collect them here, without any pretence at completeness. To be consistent with the original statements, in this section we express the dependence of the operators and the scalar product on γ through the associated initial covector λ.

Let $\gamma : [0,T] \to M$ be an ample geodesic with initial covector $\lambda \in T^*_{x_0}M$, and let $\mathcal{Q}_\lambda(t)$ be the symmetric operator associated with the second derivative $d^2_{x_0}\dot{c}_t$ via the scalar product $\langle \cdot | \cdot \rangle_\lambda$, defined for sufficiently small $t > 0$.

Theorem A (Section 4.3)**.** *The map $t \mapsto t^2 \mathcal{Q}_\lambda(t)$ can be extended to a smooth family of operators on $\mathscr{D}_{x_0}$ for small $t \geq 0$, symmetric with respect to $\langle \cdot | \cdot \rangle_\lambda$. Moreover,*

$$\mathcal{I}_\lambda \doteq \lim_{t \to 0^+} t^2 \mathcal{Q}_\lambda(t) \geq \mathbb{I} > 0,$$

as operators on $(\mathscr{D}_{x_0}, \langle\cdot|\cdot\rangle_\lambda)$. *Finally*

$$\left.\frac{d}{dt}\right|_{t=0} t^2 \mathcal{Q}_\lambda(t) = 0.$$

Then, let $\lambda \in T^*_{x_0}M$ be the initial covector associated with an ample geodesic. The *curvature* is the symmetric operator $\mathcal{R}_\lambda : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ defined by

$$\mathcal{R}_\lambda \doteq \frac{3}{2} \left.\frac{d^2}{dt^2}\right|_{t=0} t^2 \mathcal{Q}_\lambda(t).$$

Moreover, the *Ricci curvature* at $\lambda \in T^*_{x_0}M$ is defined by $\mathrm{Ric}(\lambda) \doteq \operatorname{tr} \mathcal{R}_\lambda$. In particular, we have the following Laurent expansion for the family of symmetric operators $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$

$$\mathcal{Q}_\lambda(t) = \frac{1}{t^2}\mathcal{I}_\lambda + \frac{1}{3}\mathcal{R}_\lambda + O(t), \qquad t > 0.$$

Theorem B (Section 4.3.1)**.** *Let* $\gamma : [0,T] \to M$ *be an ample and equiregular geodesic. Then the symmetric operator* $\mathcal{I}_\lambda : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ *satisfies*

(i) $\operatorname{spec} \mathcal{I}_\lambda = \{n_1^2, \dots, n_k^2\}$,

(ii) $\operatorname{tr} \mathcal{I}_\lambda = n_1^2 + \ldots + n_k^2$.

Let M be a sub-Riemannian manifold and let Δ_μ be the sub-Laplacian associated with a smooth volume μ. The next result is an explicit expression for the asymptotics of the sub-Laplacian of the squared distance from a geodesic, computed at the initial point x_0 of the geodesic γ. Let $\mathfrak{f}_t \doteq \frac{1}{2}\mathsf{d}^2(\,\cdot\,, \gamma(t))$.

Theorem C (Section 5.4)**.** *Let* γ *be an equiregular geodesic with initial covector* $\lambda \in T^*_{x_0}M$. *Assume also that* $\dim \mathscr{D}$ *is constant in a neighbourhood of* x_0. *Then there exists a smooth* n*-form* ω *defined along* γ*, such that for any volume form* μ *on* M, $\mu_{\gamma(t)} = e^{g(t)}\omega_{\gamma(t)}$, *we have*

$$\Delta_\mu \mathfrak{f}_t|_{x_0} = \operatorname{tr} \mathcal{I}_\lambda - \dot{g}(0)t - \frac{1}{3}\mathrm{Ric}(\lambda)t^2 + O(t^3).$$

Let $x_0 \in M$, and let $\Sigma_{x_0} \subset M$ be the set of smooth points for the function $x \mapsto \mathsf{d}^2(x_0, x)$. Let $\Omega_{x_0,t}$ be the homothety of a set $\Omega \subset \Sigma_{x_0}$ with respect to x_0 along the geodesics connecting x_0 with the points of Ω.

Theorem D (Section 5.6)**.** *Let* μ *be a smooth volume. For any bounded, measurable set* $\Omega \subset \Sigma_{x_0}$, *with* $0 < \mu(\Omega) < +\infty$ *we have*

$$\mu(\Omega_{x_0,t}) \sim t^{\mathcal{N}_{x_0}}, \qquad \text{for } t \to 0.$$

where $\mathcal{N}_{x_0}$ *is the geodesic dimension at the point* x_0.

Part I
Statements of the results

2 General setting

In this section we introduce a general framework that allows to treat smooth control system on a manifold in a coordinate free way, i.e. invariant under state and feedback transformations. For the sake of simplicity, we will restrict our definition to the case of nonlinear affine control systems, although the construction of this section can be extended to any smooth control system (see [1]).

2.1 Affine control systems

Definition 2.1. Let M be a connected smooth n-dimensional manifold. An *affine control system* on M is a pair $(\mathbb{U}, f)$ where:

(i) $\mathbb{U}$ is a smooth rank k vector bundle with base M and fiber $\mathbb{U}_x$ i.e., for every $x \in M$, $\mathbb{U}_x$ is a k-dimensional vector space,

(ii) $f : \mathbb{U} \to TM$ is a smooth affine morphism of vector bundles, i.e. the diagram (3) is commutative and f is *affine* on fibers.

$$\begin{array}{ccc} \mathbb{U} & \xrightarrow{f} & TM \\ & \searrow^{\pi_{\mathbb{U}}} & \downarrow \pi \\ & & M \end{array} \tag{3}$$

The maps $\pi_{\mathbb{U}}$ and π are the canonical projections of the vector bundles $\mathbb{U}$ and TM, respectively.

We denote points in $\mathbb{U}$ as pairs (x, u), where $x \in M$ and $u \in \mathbb{U}_x$ is an element of the fiber. According to this notation, the image of the point (x, u) through f is $f(x, u)$ or $f_u(x)$ and we prefer the second one when we want to emphasize f_u as a vector on T_xM. Finally, let $L^\infty([0, T], \mathbb{U})$ be the set of measurable, essentially bounded functions $u : [0, T] \to \mathbb{U}$.

Definition 2.2. A Lipschitz curve $\gamma : [0, T] \to M$ is said to be *admissible* for the control system if there exists a *control* $u \in L^\infty([0, T], \mathbb{U})$ such that $\pi_{\mathbb{U}} \circ u = \gamma$ and

$$\dot{\gamma}(t) = f(\gamma(t), u(t)), \qquad \text{for a.e. } t \in [0, T].$$

The pair (γ, u) of an admissible curve γ and its control u is called *admissible pair*.

We denote by $\overline{f} : \mathbb{U} \to TM$ the linear bundle morphism induced by f. In other words we write $f(x, u) = f_0(x) + \overline{f}(x, u)$, where $f_0(x) \doteq f(x, 0)$ is the image of the zero section. In terms of a local frame for $\mathbb{U}$, $\overline{f}(x, u) = \sum_{i=1}^k u_i f_i(x)$.

Definition 2.3. The *distribution* $\mathscr{D} \subset TM$ is the family of subspaces

$$\mathscr{D} = \{\mathscr{D}_x\}_{x \in M}, \qquad \text{where} \qquad \mathscr{D}_x \doteq \overline{f}(\mathbb{U}_x) \subset T_xM.$$

The family of *horizontal vector fields* $\overline{\mathscr{D}} \subset \operatorname{Vec}(M)$ is

$$\overline{\mathscr{D}} = \operatorname{span}\left\{\overline{f} \circ \sigma,\ \sigma : M \to \mathbb{U} \text{ is a smooth section of } \mathbb{U}\right\}.$$

Observe that, if the rank of $\overline{f}$ is not constant, $\mathscr{D}$ is not a sub-bundle of TM. Therefore the dimension of $\mathscr{D}_x$, in general, depends on $x \in M$.

Given a smooth function $L : \mathbb{U} \to \mathbb{R}$, called a *Lagrangian*, the *cost functional at time* T, called $J_T : L^\infty([0,T], \mathbb{U}) \to \mathbb{R}$, is defined by

$$J_T(u) \doteq \int_0^T L(\gamma(t), u(t))dt,$$

where $\gamma(t) = \pi(u(t))$. We are interested in the problem of minimizing the cost among all admissible pairs (γ, u) that join two fixed points $x_0, x_1 \in M$ in time T. This corresponds to the optimal control problem

$$\begin{aligned} &\dot{x} = f(x,u) = f_0(x) + \sum_{i=1}^k u_i f_i(x), \qquad x \in M, \\ &x(0) = x_0,\ x(T) = x_1, \qquad J_T(u) \to \min, \end{aligned} \tag{4}$$

where we have chosen some local trivialization of $\mathbb{U}$.

Definition 2.4. Let $M' \subset M$ be an open subset with compact closure. For $x_0, x_1 \in M'$ and $T > 0$, we define the *value function*

$$S_T(x_0, x_1) \doteq \inf\{J_T(u) \,|\, (\gamma, u) \text{ admissible pair}, \gamma(0) = x_0,\, \gamma(T) = x_1,\, \gamma \subset M'\}.$$

The value function depends on the choice of a relatively compact subset $M' \subset M$. This choice, which is purely technical, is related with Theorem 2.19, concerning the regularity properties of S. We stress that all the objects defined in this paper by using the value function do not depend on the choice of M'.

Assumptions. In what follows we make the following general assumptions:

(A1) The affine control system is *bracket generating*, namely

$$\mathrm{Lie}_x \left\{ (\mathrm{ad}\, f_0)^i\, \overline{\mathscr{D}} \,|\, i \in \mathbb{N} \right\} = T_x M, \qquad \forall\, x \in M, \tag{5}$$

where $(\mathrm{ad}\, X)Y = [X, Y]$ is the Lie bracket of two vector fields and $\mathrm{Lie}_x \mathcal{F}$ denotes the Lie algebra generated by a family of vector fields $\mathcal{F}$, computed at the point x. Observe that the vector field f_0 is not included in the generators of the Lie algebra (5).

(A2) The function $L : \mathbb{U} \to \mathbb{R}$ is a *Tonelli Lagrangian*, i.e. it satisfies

(A2.a) The Hessian of $L|_{\mathbb{U}_x}$ is positive definite for all $x \in M$. In particular, $L|_{\mathbb{U}_x}$ is strictly convex.

(A2.b) L has superlinear growth, i.e. $L(x,u)/|u| \to +\infty$ when $|u| \to +\infty$.

Assumptions (A1) and (A2) are necessary conditions in order to have a nontrivial set of strictly normal minimizer and allow us to introduce a well defined smooth Hamiltonian (see Section 3).

2.1.1 State-feedback equivalence

All our considerations will be local. Hence, up to restricting our attention to a trivializable neighbourhood of M, we can assume that $\mathbb{U} \simeq M \times \mathbb{R}^k$. By choosing a basis of $\mathbb{R}^k$, we can write $f(x,u) = f_0(x) + \sum_{i=1}^k u_i f_i(x)$. Then, a Lipschitz curve $\gamma : [0,T] \to M$ is admissible if there exists a measurable, essentially bounded control $u : [0,T] \to \mathbb{R}^k$ such that

$$\dot{\gamma}(t) = f_0(\gamma(t)) + \sum_{i=1}^k u_i(t) f_i(\gamma(t)), \qquad \text{for a.e. } t \in [0,T].$$

We use the notation $u \in L^\infty([0,T],\mathbb{R}^k)$ to denote a measurable, essentially bounded control with values in $\mathbb{R}^k$. By choosing another (local) trivialization of $\mathbb{U}$, or another basis of $\mathbb{R}^k$, we obtain a different *presentation* of the same affine control system. Besides, by acting on the underlying manifold M via diffeomorphisms, we obtain equivalent affine control system starting from a given one. The following definition formalizes the concept of equivalent control systems.

Definition 2.5. Let $(\mathbb{U}, f)$ and $(\mathbb{U}', f')$ be two affine control systems on the same manifold M. A *state-feedback transformation* is a pair (ϕ, ψ), where $\phi : M \to M$ is a diffeomorphism and $\psi : \mathbb{U} \to \mathbb{U}'$ an invertible affine bundle map, such that the following diagram is commutative.

$$\begin{array}{ccc} \mathbb{U} & \xrightarrow{f} & TM \\ \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \phi_*} \\ \mathbb{U}' & \xrightarrow[f']{} & TM \end{array} \tag{6}$$

In other words, $\phi_* f(x,u) = f'(\phi(x), \psi(x,u))$ for every $(x,u) \in \mathbb{U}$. In this case $(\mathbb{U}, f)$ and $(\mathbb{U}', f')$ are said *state-feedback equivalent.*

Notice that, if $(\mathbb{U}, f)$ and $(\mathbb{U}', f')$ are state-feedback equivalent, then $\operatorname{rank} \mathbb{U} = \operatorname{rank} \mathbb{U}'$. Moreover, different presentations of the same control systems are indeed feedback equivalent (i.e. related by a state-feedback transformation with $\phi = \mathbb{I}$). Definition 2.5 corresponds to the classical notion of point-dependent reparametrization of the controls. The next lemma states that a state-feedback transformation preserves admissible curves.

Lemma 2.6. *Let $\gamma_{x_0,u}$ be the admissible curve starting from x_0 and associated with u. Then $\phi(\gamma_{x_0,u}(t)) = \gamma_{\phi(x_0),v}(t)$ where $v(t) = \psi(x(t), u(t))$.*

Proof. Denote $x(t) = \gamma_{x_0,u}(t)$ and set $y(t) \doteq \phi(x(t))$. Then, by definition, $\dot{x}(t) = f(x(t), u(t))$ and $x(0) = x_0$. Hence $y(0) = \phi(x_0)$ and

$$\dot{y}(t) = \phi_* f(x(t), u(t)) = f'(\phi(x(t)), \psi(x(t), u(t))) = f'(y(t), v(t)).$$

□

Remark 2.7. Notice that every state-feedback transformation (ϕ, ψ) can be written as a composition of a pure state one, i.e. with $\psi = \mathbb{I}$, and a pure feedback one, i.e. with $\phi = \mathbb{I}$. For later convenience, let us discuss how two feedback equivalent systems are related. Consider a presentation of an affine control system

$$\dot{x} = f(x,u) = f_0(x) + \sum_{i=1}^k u_i f_i(x).$$

By the commutativity of diagram (6), a feedback transformation writes

$$\begin{cases} u' = \psi(x,u) \\ x' = \phi(x) \end{cases} \qquad u_i' = \psi_i(x,u) = \psi_{i,0}(x) + \sum_{j=1}^{k} \psi_{i,j}(x) u_j, \qquad i = 1, \dots, k,$$

where $\psi_{i,0}$ and $\psi_{i,j}$ denote, respectively, the affine and the linear part of the i-th component of ψ. In particular, for a pure feedback transformation, the original system is equivalent to

$$\dot{x} = f'(x,u') = f_0'(x) + \sum_{i=1}^{k} u_i' f_i'(x),$$

where $f_0(x) \doteq f_0'(x) + \sum_{i=1}^{k} \psi_{i,0}(x) f_i'(x)$ and $f_i(x) \doteq \sum_{j=1}^{k} \psi_{j,i}(x) f_j'(x)$.

We conclude this section recalling some well known facts about non-autonomous flows. By Caratheodory Theorem, for every control $u \in L^\infty([0,T], \mathbb{R}^k)$ and every initial condition $x_0 \in M$, there exists a unique Lipschitz solution to the Cauchy problem

$$\begin{cases} \dot{\gamma}(t) = f_0(\gamma(t)) + \sum_{i=1}^{k} u_i(t) f_i(\gamma(t)), \\ \gamma(0) = x_0, \end{cases} \tag{7}$$

defined for small time (see, e.g. [9,28]). We denote such a solution by $\gamma_{x_0,u}$ (or simply γ_u when the base point x_0 is fixed). Moreover, for a fixed control $u \in L^\infty([0,T], \mathbb{R}^k)$, it is well defined the family of diffeomorphisms $P_{0,t} : M \to M$, given by $P_{0,t}(x) \doteq \gamma_{x,u}(t)$, which is Lipschitz with respect to t. Analogously one can define the flow $P_{s,t} : M \to M$, by solving the Cauchy problem with initial condition given at time s. Notice that $P_{t,t} = \mathbb{I}$ for all $t \in \mathbb{R}$ and $P_{t_1,t_2} \circ P_{t_0,t_1} = P_{t_0,t_2}$, whenever they are defined. In particular $(P_{t_1,t_2})^{-1} = P_{t_2,t_1}$.

2.2 End-point map

In this section, for convenience, we assume to fix some (local) presentation of the affine control system, hence $L^\infty([0,T], \mathbb{U}) \simeq L^\infty([0,T], \mathbb{R}^k)$. For a more intrinsic approach see [1, Sec. 1].

Definition 2.8. Fix a point $x_0 \in M$ and $T > 0$. The *end-point map at time* T of the system (7) is the map

$$E_{x_0,T} : \mathcal{U} \to M, \qquad u \mapsto \gamma_{x_0,u}(T),$$

where $\mathcal{U} \subset L^\infty([0,T], \mathbb{R}^k)$ is the open subset of controls such that the solution $t \mapsto \gamma_{x_0,u}(t)$ of the Cauchy problem (7) is defined on the whole interval $[0,T]$.

The end-point map is smooth. Moreover, its Fréchet differential is computed by the following well-known formula (see, e.g. [9]).

Proposition 2.9. *The differential of $E_{x_0,T}$ at $u \in \mathcal{U}$, i.e. $D_u E_{x_0,T} : L^\infty([0,T], \mathbb{R}^k) \to T_x M$, where $x = \gamma_u(T)$, is*

$$D_u E_{x_0,T}(v) = \int_0^T (P_{s,T})_* \overline{f}_{v(s)}(\gamma_u(s)) ds, \qquad \forall\, v \in L^\infty([0,T], \mathbb{R}^k). \tag{8}$$

In other words the differential $D_u E_{x_0,T}$ applied to the control v computes the integral mean of the linear part $\overline{f}_{v(t)}$ of the vector field $f_{v(t)}$ along the trajectory defined by u, by pushing it forward to the final point of the trajectory through the flow $P_{s,T}$ (see Fig. 1).

More explicitly, $f(x,u) = f_0(x) + \sum_{i=1}^{k} u_i f_i(x)$, and Eq. (8) is rewritten as follows

$$D_u E_{x_0,T}(v) = \int_0^T \sum_{i=1}^{k} v_i(s) (P_{s,T})_* f_i(\gamma_u(s)) ds, \qquad \forall\, v \in L^\infty([0,T], \mathbb{R}^k).$$

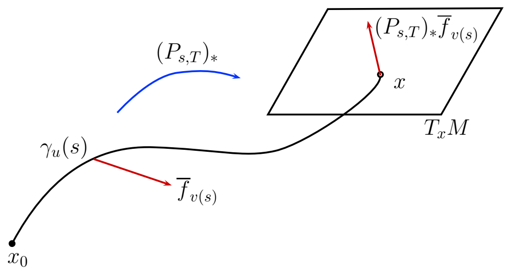


Figure 1: Differential of the end-point map.

2.3 Lagrange multipliers rule

Fix $x_0, x \in M$. The problem of finding the infimum of the cost J_T for all admissible curves connecting the endpoints x_0 and x, respectively, in time T, can be naturally reformulated via the end-point map as a constrained extremal problem

$$S_T(x_0, x) = \inf\{J_T(u) \mid E_{x_0,T}(u) = x\} = \inf_{E^{-1}_{x_0,T}(x)} J_T. \tag{9}$$

Definition 2.10. We say that $u \in \mathcal{U}$ is an *optimal control* if it is a solution of Eq. (9).

Remark 2.11. When f is not injective, a curve γ may be associated with multiple controls. Nevertheless, among all the possible controls u associated with the same admissible curve, there exists a unique control u^* which, for a.e. $t \in [0, T]$, minimizes the Lagrangian function. Then, since we are interested in optimal controls, we assume that any admissible curve γ is always associated with the control u^* which minimizes the Lagrangian, and in this way we have a one-to-one correspondence between admissible curves and controls. With this observation, we say that the admissible curve γ is an *optimal trajectory* (or *minimizer*) if the associated control u^* is optimal according to Definition 2.10.

Notice that, in general, $D_uE_{x_0,T}$ is not surjective and the set $E^{-1}_{x_0,T}(x) \subset M$ is not a smooth submanifold. The Lagrange multipliers rule provides a necessary condition to be satisfied by a control u which is a constrained critical point for (9).

Proposition 2.12. *Let $u \in \mathcal{U}$ be an optimal control, with $x = E_{x_0,T}(u)$. Then (at least) one of the two following statements holds true*

(i) $\exists\, \lambda_T \in T^*_xM$ *s.t.* $\lambda_T\, D_uE_{x_0,T} = d_uJ_T$,

(ii) $\exists\, \lambda_T \in T^*_xM$, $\lambda_T \neq 0$, *s.t.* $\lambda_T\, D_uE_{x_0,T} = 0$,

where $\lambda_T\, D_uE_{x_0,T}$ denotes the composition of linear maps

$$\begin{array}{ccc} L^\infty([0,T], \mathbb{R}^k) & \xrightarrow{D_uE_{x_0,T}} & T_xM \\ & \searrow^{d_uJ_T} & \downarrow \lambda_T \\ & & \mathbb{R} \end{array}$$

Definition 2.13. A control u, satisfying the necessary conditions for optimality of Proposition 2.12, is called *normal extremal* in case (i), while it is called *abnormal extremal* in case (ii). We use the same terminology to classify the associated extremal trajectory γ_u.

Notice that a single control $u \in \mathcal{U}$ can be associated with two different covectors (or *Lagrange multipliers*) such that both (i) and (ii) are satisfied. In other words, an optimal trajectory may be simultaneously normal and abnormal. We now introduce a key definition for what follows.

Definition 2.14. A normal extremal trajectory $\gamma : [0,T] \to M$ is called *strictly normal* if it is not abnormal. Moreover, if for all $s \in [0,T]$ the restriction $\gamma|_{[0,s]}$ is also strictly normal, then γ is called *strongly normal*.

Remark 2.15. A trajectory is abnormal if and only if the differential $D_u E_{x_0,T}$ is not surjective. By linearity of the integral, it is easy to show from Eq. (8) that this is equivalent to the relation

$$\operatorname{span}\{(P_{s,T})_* \mathscr{D}_{\gamma(s)}, s \in [0,T]\} \neq T_{\gamma(T)}M.$$

In particular γ is strongly normal if and only if a short segment $\gamma|_{[0,\varepsilon]}$ is strongly normal, for some $\varepsilon \leq T$.

2.4 Pontryagin Maximum Principle

In this section we recall a weak version of the Pontryagin Maximum Principle (PMP) for the optimal control problem, which rewrites the necessary conditions satisfied by normal optimal solutions in the Hamiltonian formalism. In particular it states that every normal optimal trajectory of problem (4) is the projection of a solution of a fixed Hamiltonian system defined on T^*M.

Let us denote by $\pi : T^*M \to M$ the canonical projection of the cotangent bundle, and by $\langle \lambda, v \rangle$ the pairing between a cotangent vector $\lambda \in T^*_xM$ and a vector $v \in T_xM$. The Liouville 1-form $\varsigma \in \Lambda^1(T^*M)$ is defined as follows: $\varsigma_\lambda = \lambda \circ \pi_*$, for every $\lambda \in T^*M$. The canonical symplectic structure on T^*M is defined by the non degenerate closed 2-form $\sigma = d\varsigma$. In canonical coordinates $(p,x) \in T^*M$ one has

$$\varsigma = \sum_{i=1}^n p_i dx_i, \qquad \sigma = \sum_{i=1}^n dp_i \wedge dx_i.$$

We denote by $\vec{h}$ the Hamiltonian vector field associated with a function $h \in C^\infty(T^*M)$. Namely, $d_\lambda h = \sigma(\cdot, \vec{h}(\lambda))$ for every $\lambda \in T^*M$ and the coordinates expression of $\vec{h}$ is

$$\vec{h} = \sum_{i=1}^n \frac{\partial h}{\partial p_i}\frac{\partial}{\partial x_i} - \frac{\partial h}{\partial x_i}\frac{\partial}{\partial p_i}.$$

Let us introduce the smooth control-dependent Hamiltonian on T^*M:

$$\mathcal{H}(\lambda, u) = \langle \lambda, f(x,u) \rangle - L(x,u), \qquad \lambda \in T^*M,\ x = \pi(\lambda).$$

Assumption (A2) guarantees that, for each $\lambda \in T^*M$, the restriction $u \mapsto \mathcal{H}(\lambda, u)$ to the fibers of $\mathbb{U}$ has a unique maximum $\bar{u}(\lambda)$. Moreover, the fiber-wise strong convexity of the Lagrangian and an easy application of the implicit function theorem prove that the map $\lambda \mapsto \bar{u}(\lambda)$ is smooth. Therefore, it is well defined the *maximized Hamiltonian* (or simply, *Hamiltonian*) $H : T^*M \to \mathbb{R}$

$$H(\lambda) \doteq \max_{v \in U_x} \mathcal{H}(\lambda, v) = \mathcal{H}(\lambda, \bar{u}(\lambda)), \qquad \lambda \in T^*M, x = \pi(\lambda).$$

Remark 2.16. When $f(x,u) = f_0(x) + \sum_{i=1}^k u_i f_i(x)$ is written in a local frame, then $\bar{u} = \bar{u}(\lambda)$ is characterized as the solution of the system

$$\frac{\partial \mathcal{H}}{\partial u_i}(\lambda, u) = \langle \lambda, f_i(x) \rangle - \frac{\partial L}{\partial u_i}(x,u) = 0, \qquad i = 1, \dots, k. \tag{10}$$

Theorem 2.17 (PMP, [9, 28])**.** *The admissible curve $\gamma : [0,T] \to M$ is a normal extremal trajectory if and only if there exists a Lipschitz lift $\lambda : [0,T] \to T^*M$, such that $\gamma(t) = \pi(\lambda(t))$ and*

$$\dot{\lambda}(t) = \vec{H}(\lambda(t)), \qquad t \in [0,T].$$

In particular, γ and λ are smooth. Moreover, the associated control can be recovered from the lift as $u(t) = \bar{u}(\lambda(t))$, and the final covector $\lambda_T = \lambda(T)$ is the normal Lagrange multiplier associated with u, namely $\lambda_T \, D_u E_{x_0,T} = d_u J_T$.

Thus, every normal extremal trajectory $\gamma : [0,T] \to M$ can be written as $\gamma(t) = \pi \circ e^{t\vec{H}}(\lambda_0)$, for some initial covector $\lambda_0 \in T^*M$ (although it may be non unique). This observation motivates the next definition. For simplicity, and without loss of generality, we assume that $\vec{H}$ is complete.

Definition 2.18. Fix $x_0 \in M$. The *exponential map* with base point x_0 is the map $\mathcal{E}_{x_0} : \mathbb{R}^+ \times T^*_{x_0}M \to M$, defined by $\mathcal{E}_{x_0}(t, \lambda_0) = \pi \circ e^{t\vec{H}}(\lambda_0)$.

When the first argument is fixed, we employ the notation $\mathcal{E}_{x_0,t} : T^*_{x_0}M \to M$ to denote the exponential map with base point x_0 and time t, namely $\mathcal{E}_{x_0,t}(\lambda) = \mathcal{E}_{x_0}(t,\lambda)$. Indeed, the exponential map is smooth.

From now on, we call *geodesic* any trajectory that satisfies the normal necessary conditions for optimality. In other words, geodesics are admissible curves associated with a normal Lagrange multiplier or, equivalently, projections of integral curves of the Hamiltonian flow.

2.5 Regularity of the value function

The next well known regularity property of the value function is crucial for the forthcoming sections (see Definition 2.4).

Theorem 2.19. *Let $\gamma : [0,T] \to M'$ be a strongly normal trajectory. Then there exist $\varepsilon > 0$ and an open neighbourhood $U \subset (0,\varepsilon) \times M' \times M'$ such that:*

(i) $(t, \gamma(0), \gamma(t)) \in U$ *for all* $t \in (0,\varepsilon)$,

(ii) For any $(t,x,y) \in U$ there exists a unique (normal) minimizer of the cost functional J_t, among all the admissible curves that connect x with y in time t, contained in M',

(iii) The value function $(t,x,y) \mapsto S_t(x,y)$ is smooth on U.

According to Definition 2.4, the function S, and henceforth U, depend on the choice of a relatively compact $M' \subset M$. For different relatively compacts, the correspondent value functions S agree on the intersection of the associated domains U: they define the same germ.

The proof of this result can be found in Appendix C. We end this section with a useful lemma about the differential of the value function at a smooth point.

Lemma 2.20. *Let $x_0, x \in M$ and $T > 0$. Assume that the function $x \mapsto S_T(x_0, x)$ is smooth at x and there exists an optimal trajectory $\gamma : [0,T] \to M$ joining x_0 to x. Then*

(i) γ is the unique minimizer of the cost functional J_T, among all the admissible curves that connect x_0 with x in time T, and it is strictly normal,

(ii) $d_x S_T(x_0, \cdot) = \lambda_T$, where λ_T is the final covector of the normal lift of γ.

Proof. Under the above assumptions the function

$$v \mapsto J_T(v) - S_T(x_0, E_{x_0,T}(v)), \qquad v \in L^\infty([0,T], \mathbb{R}^k),$$

is smooth and non negative. For every optimal trajectory γ, associated with the control u, that connects x_0 with x in time T, one has

$$0 = d_u(J_T(\cdot) - S_T(x_0, E_{x_0,T}(\cdot)) = d_u J_T - d_x S_T(x_0, \cdot) \circ D_u E_{x_0,T}.$$

Thus, γ is a normal extremal trajectory, with Lagrange multiplier $\lambda_T = d_x S_T(x_0, \cdot)$. By Theorem 2.17, we can recover γ by the formula $\gamma(t) = \pi \circ e^{(t-T)\vec{H}}(\lambda_T)$. Then, γ is the unique minimizer of J_T connecting its endpoints.

Next we show that γ is not abnormal. For y in a neighbourhood of x, consider the map

$$\Theta : y \mapsto e^{-T\vec{H}}(d_y S_T(x_0, \cdot)).$$

The map Θ, by construction, is a smooth right inverse for the exponential map at time T. This implies that x is a regular value for the exponential map and, a fortiori, u is a regular point for the end-point map at time T. ☐

3 Flag and growth vector of an admissible curve

For each smooth admissible curve, we introduce a family of subspaces, which is related with a micro-local characterization of the control system along the trajectory itself.

3.1 Growth vector of an admissible curve

Let $\gamma : [0,T] \to M$ be an admissible, smooth curve such that $\gamma(0) = x_0$, associated with a smooth control u. Let $P_{0,t}$ denote the flow defined by u. We define the family of subspaces of $T_{x_0}M$

$$\mathscr{F}_\gamma(t) \doteq (P_{0,t})_*^{-1} \mathscr{D}_{\gamma(t)}. \tag{11}$$

In other words, the family $\mathscr{F}_\gamma(t)$ is obtained by collecting the distributions along the trajectory at the initial point, by using the flow $P_{0,t}$ (see Fig. 2).

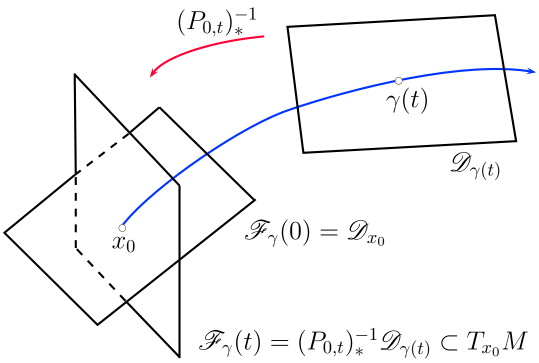


Figure 2: The family of subspaces $\mathscr{F}_\gamma(t)$.

Given a family of subspaces in a linear space it is natural to consider the associated flag.

Definition 3.1. The *flag of the admissible curve* γ is the sequence of subspaces

$$\mathscr{F}^i_\gamma(t) \doteq \operatorname{span}\left\{ \frac{d^j}{dt^j} v(t) \,\middle|\, v(t) \in \mathscr{F}_\gamma(t) \text{ smooth}, \, j \leq i-1 \right\} \subset T_{x_0}M, \qquad i \geq 1.$$

Notice that, by definition, this is a filtration of $T_{x_0}M$, i.e. $\mathscr{F}^i_\gamma(t) \subset \mathscr{F}^{i+1}_\gamma(t)$, for all $i \geq 1$.

Definition 3.2. Let $k_i(t) \doteq \dim \mathscr{F}^i_\gamma(t)$. The *growth vector of the admissible curve* γ is the sequence of integers $\mathcal{G}_\gamma(t) = \{k_1(t), k_2(t), \dots\}$.

An admissible curve is *ample at* t if there exists an integer $m = m(t)$ such that $\mathscr{F}^{m(t)}_\gamma(t) = T_{x_0}M$. We call the minimal $m(t)$ such that the curve is ample the *step at* t *of the admissible curve.* An admissible curve is called *equiregular at* t if its growth vector is locally constant at t. Finally, an admissible curve is *ample* (resp. *equiregular*) if it is ample (resp. equiregular) at each $t \in [0, T]$.

Remark 3.3. One can analogously introduce the family of subspaces (and the relevant filtration) at any base point $\gamma(s)$, for every $s \in [0, T]$, by defining the shifted curve $\gamma_s(t) \doteq \gamma(s+t)$. Then $\mathscr{F}_{\gamma_s}(t) \doteq (P_{s,s+t})^{-1}_* \mathscr{D}_{\gamma(s+t)}$. Notice that the relation $\mathscr{F}_{\gamma_s}(t) = (P_{0,s})_* \mathscr{F}_\gamma(s+t)$ implies that the growth vector of the original curve at t can be equivalently computed via the growth vector at time 0 of the curve γ_t, i.e. $k_i(t) = \dim \mathscr{F}^i_{\gamma_t}(0)$, and $\mathcal{G}_\gamma(t) = \mathcal{G}_{\gamma_t}(0)$.

Let us stress that the the family of subspaces (11) depends on the choice of the local frame (via the map $P_{0,t}$). However, we will prove that the flag of an admissible curve at $t = 0$ and its growth vector (for all t) are invariant by state-feedback transformation and, in particular, independent on the particular presentation of the system (see Section 3.3).

Remark 3.4. The following properties of the growth vector of an *ample* admissible curve highlight the analogy with the "classical" growth vector of the distribution.

(i) The functions $t \mapsto k_i(t)$, for $i = 1, \dots, m(t)$, are lower semicontinuous. In particular, being integer valued functions, this implies that the set of points t such that the growth vector is locally constant is open and dense on $[0, T]$.

(ii) The function $t \mapsto m(t)$ is upper semicontinuous. As a consequence, the step of an admissible curve is bounded on $[0, T]$.

(iii) If the admissible curve is equiregular at t, then $k_1(t) < \ldots < k_m(t)$ is a strictly increasing sequence. Let $i < m$. If $k_i(t) = k_{i+1}(t)$ for all t in a open neighbourhood then, using a local frame, it is easy to see that this implies $k_i(t) = k_{i+1}(t) = \ldots = k_m(t)$ contradicting the fact that the admissible curve is ample at t.

(iv) Assume that an admissible curve is equiregular with step m. The derivation of sections of $\mathscr{F}_\gamma(t)$ induces a well defined map on the quotients

$$\mathscr{F}^i_\gamma(t)/\mathscr{F}^{i-1}_\gamma(t) \longrightarrow \mathscr{F}^{i+1}_\gamma(t)/\mathscr{F}^i_\gamma(t), \qquad \forall\, t \in [0, T].$$

In this case, the maps defined above are surjective and the quotients $\mathscr{F}^i_\gamma/\mathscr{F}^{i-1}_\gamma$ have constant dimensions $d_i \doteq k_{i+1} - k_i = \dim \mathscr{F}^i_\gamma - \dim \mathscr{F}^{i-1}_\gamma$ for $i = 1, \dots, m$. Therefore the sequence $d_1 \geq \ldots \geq d_m$ is decreasing, namely $\dim \mathscr{F}^i_\gamma - \dim \mathscr{F}^{i-1}_\gamma \geq \dim \mathscr{F}^{i+1}_\gamma - \dim \mathscr{F}^i_\gamma$.

Next, we show how the family $\mathscr{F}_\gamma(t)$ can be conveniently employed to characterize strictly and strongly normal geodesics.

Proposition 3.5. *Let* $\gamma : [0, T] \to M$ *be a geodesic. Then* *(i)* γ *is strictly normal if and only if* $\mathrm{span}\{\mathscr{F}_\gamma(s), s \in [0,T]\} = T_{x_0}M$*,*

(ii) γ *is strongly normal if and only if* $\mathrm{span}\{\mathscr{F}_\gamma(s), s \in [0,t]\} = T_{x_0}M$ *for all* $0 < t \leq T$*,*

(iii) If γ *is ample at* $t = 0$*, then it is strongly normal.*

Proof. Recall that a geodesic $\gamma : [0,T] \to M$ is abnormal on $[0,T]$ if and only if the differential $D_u E_{x_0,T}$ is not surjective, which implies (see Remark 2.15)

$$\mathrm{span}\{(P_{s,T})_* \mathscr{D}_{\gamma(s)}, s \in [0,T]\} \neq T_{\gamma(T)}M.$$

By applying the inverse flow $(P_{0,T})_*^{-1} : T_{\gamma(T)}M \to T_{\gamma(0)}M$, we obtain

$$\mathrm{span}\{\mathscr{F}_\gamma(s), s \in [0,T]\} \neq T_{x_0}M.$$

This proves (i). In particular, this implies that a geodesic is strongly normal if and only if

$$\mathrm{span}\{\mathscr{F}_\gamma(s), s \in [0,t]\} = T_{x_0}M, \qquad \forall\, 0 < t \leq T,$$

which proves (ii). We now prove (iii). We argue by contradiction. If the geodesic is not strongly normal, there exists some $\lambda \in T^*_{x_0}M$ such that $\langle \lambda, \mathscr{F}_\gamma(t)\rangle = 0$, for all $0 < t \leq T$. Then, by taking derivatives at $t = 0$, we obtain that $\langle \lambda, \mathscr{F}^i_\gamma(0)\rangle = 0$, for all $i \geq 0$, which is impossible since the curve is ample at $t = 0$ by hypotesis. $\square$

Remark 3.6. Ample geodesics play a crucial role in our approach to curvature, as we explain in Section 4. By Proposition 3.5, these geodesics are strongly normal. One may wonder whether the generic covector $\lambda_0 \in T^*_{x_0}M$ corresponds to a strongly normal (or even ample) geodesic. The answer to this question is trivial when there are no abnormal trajectories (e.g. in Riemannian geometry), but the matter is quite delicate in general. For this reason, in order to define the curvature of an affine control system, we assume in the following that the set of ample geodesics is non empty. Eventually, we address the problem of existence of ample geodesics for linear quadratic control systems and sub-Riemannian geometry. In these cases, we will prove that a generic normal geodesic is ample.

3.2 Linearised control system and growth vector

It is well known that the differential of the end-point map at a point $u \in \mathcal{U}$ is related with the linearisation of the control system along the associated trajectory. The goal of this section is to discuss the relation between the controllability of the linearised system and the ampleness of the geodesic.

3.2.1 Linearisation of a control system in $\mathbb{R}^n$

We start with some general considerations. Consider the nonlinear control system in $\mathbb{R}^n$

$$\dot{x} = f(x,u), \qquad x \in \mathbb{R}^n,\ u \in \mathbb{R}^k,$$

where $f : \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n$ is smooth. Fix $x_0 \in \mathbb{R}^n$, and consider the end-point map $E_{x_0,t} : \mathcal{U} \to \mathbb{R}^n$ for $t \geq 0$. Consider a smooth solution $x_u(t)$, associated with the control $u(t)$, such that $x_u(0) = x_0$. The differential of the end-point map $D_u E_{x_0,t} : L^\infty([0,T],\mathbb{R}^k) \to \mathbb{R}^n$ at u is related with the end-point map of the *linearised system* at the pair $(x_u(t), u(t))$. More precisely, for every

$v \in L^\infty([0,T],\mathbb{R}^k)$ the trajectory $y(t) \doteq D_u E_{x_0,t}(v) \in \mathbb{R}^n$ is the solution of the non-autonomous linear system

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t), \\ y(0) = 0, \end{cases} \tag{12}$$

where $A(t) \doteq \dfrac{\partial f}{\partial x}(x_u(t), u(t))$ and $B(t) \doteq \dfrac{\partial f}{\partial u}(x_u(t), u(t))$ are smooth families of $n \times n$ and $n \times k$ matrices, respectively. We have the formula

$$y(t) = D_u E_{x_0,t}(v) = M(t) \int_0^t M(s)^{-1} B(s) v(s) ds,$$

where $M(t)$ is the solution of the matrix Cauchy problem $\dot{M}(t) = A(t)M(t)$, with $M(0) = \mathbb{I}$. Indeed the solution $M(t)$ is defined on the whole interval $[0,T]$, and it is invertible therein.

Definition 3.7. The linear control system (12) is *controllable* in time $T > 0$ if, for any $y \in \mathbb{R}^n$, there exists $v \in L^\infty([0,T],\mathbb{R}^k)$ such that the associated solution $y_v(t)$ satisfies $y_v(T) = y$.

Let us recall the following classical controllability condition for a linear non-autonomous system, which is the non-autonomous generalization of the Kalman condition (see e.g. [17]). For a set $\{M_i\}$ of $n \times k$ matrices, we denote with $\operatorname{span}\{M_i\}$ the vector space generated by the columns of the matrices in $\{M_i\}$.

Proposition 3.8. *Consider the control system* (12)*, with* $A(t), B(t)$ *smooth, and define*

$$B_1(t) \doteq B(t), \qquad B_{i+1}(t) \doteq A(t)B_i(t) - \dot{B}_i(t). \tag{13}$$

Assume that there exist $t \in [0,T]$ *and* $m > 0$ *such that* $\operatorname{span}\{B_1(t), B_2(t), \ldots, B_m(t)\} = \mathbb{R}^n$. *Then the system* (12) *is controllable in time* T.

Remark 3.9. Notice that, using $M(t)$ as a time-dependent change of variable, the new curve $\zeta(t) \doteq M(t)^{-1} y(t) \in \mathbb{R}^n$ satisfies

$$\begin{cases} \dot{\zeta}(t) = M(t)^{-1} B(t) v(t), \\ \zeta(0) = 0. \end{cases} \tag{14}$$

If the controllability condition of Proposition (3.8) is satisfied for the pair $(A(t), B(t))$, then it is satisfied also for the pair $(0, C(t))$, with $C(t) = M(t)^{-1}B(t)$, as a consequence of the identity $C^{(i)}(t) = (-1)^i M(t)^{-1} B_{i+1}(t)$. Therefore, the controllability conditions for the control systems (12) and (14) are equivalent. Moreover, both systems are controllable if and only if one of them is controllable.

3.2.2 Linearisation of a control system in the general setting

Let us go back to the general setting. Let γ be a smooth admissible trajectory associated with the control u such that $\gamma(0) = x_0$. We are interested in the linearisation of the affine control system at γ. Consider the image of a fixed control $v \in L^\infty([0,T],\mathbb{R}^k)$ through the differential of the end-point map $E_{x_0,t}$, for every $t \geq 0$:

$$D_u E_{x_0,t} : L^\infty([0,T],\mathbb{R}^k) \to T_{\gamma(t)}M, \qquad \gamma(t) = E_{x_0,t}(u).$$

In this case, for each $t \geq 0$, the image of v belongs to a different tangent space. In order to obtain a well defined differential equation, we collect the family of vectors in a single vector space through the composition with the push forward $(P_{0,t})_*^{-1} : T_{\gamma(t)}M \to T_{x_0}M$:

$$(P_{0,t})_*^{-1} \circ D_u E_{x_0,t} : L^\infty([0,T],\mathbb{R}^k) \to T_{x_0}M.$$

Using formula (8) one easily finds

$$(P_{0,t})_*^{-1} \circ D_u E_{x_0,t}(v) = \int_0^t (P_{0,s})_*^{-1} \overline{f}_{v(s)}(\gamma(s))ds.$$

Denoting $\zeta(t) \doteq (P_{0,t})_*^{-1} \circ D_u E_{x_0,t}(v) \in T_{x_0}M$ one has that, in a local frame, this curve satisfies

$$\dot{\zeta}(t) = (P_{0,t})_*^{-1} \overline{f}_{v(t)}(\gamma(t)) = \sum_{i=1}^k v_i(t)(P_{0,t})_*^{-1} f_i(\gamma(t)).$$

Therefore, $\zeta(t)$ is a solution of the control system

$$\begin{cases} \dot{\zeta}(t) = C(t)v(t), \\ \zeta(0) = 0, \end{cases} \tag{15}$$

where the $n \times k$ matrix $C(t)$ has columns $C_i(t) \doteq (P_{0,t})_*^{-1} f_i(\gamma(t))$ for $i = 1, \dots, k$. Eq. (15) is the *linearised system along the admissible curve* γ. By hypotesis, γ is smooth. Then the linearised system is also smooth.

Remark 3.10. Notice that the composition of the end-point map with $(P_{0,t})_*^{-1}$ corresponds to the time dependent transformation $M(t)^{-1}$ of Remark 3.9.

3.2.3 Growth vector and controllability

From the definition of growth vector of an admissible curve, it follows that

$$\mathscr{F}_\gamma^i(t) = \text{span}\{C(t), \dot{C}(t) \dots, C^{(i-1)}(t)\}, \qquad i \geq 1.$$

This gives an efficient criterion to compute the geodesic growth vector of the admissible curve γ_u associated with the control u. Define in any local frame $f_1, \dots, f_k$ and any coordinate system in a neighbourhood of γ, the $n \times n$ and $n \times k$ matrices, respectively:

$$A(t) \doteq \frac{\partial f}{\partial x}(\gamma_u(t), u(t)) = \frac{\partial f_0}{\partial x}(\gamma_u(t)) + \sum_{i=1}^k u_i(t) \frac{\partial f_i}{\partial x}(\gamma_u(t)), \tag{16}$$

$$B(t) \doteq \frac{\partial f}{\partial u}(\gamma_u(t), u(t)) = [f_i(\gamma_u(t))]_{i=1,\dots,k}\,. \tag{17}$$

Denoting by $B_j(t)$ the matrices defined as in (13), and recalling Remark 3.9, we have

$$k_i(t) = \dim \mathscr{F}_\gamma^i(t) = \text{rank}\{B_1(t), \dots, B_i(t)\}.$$

Assume now that the admissible curve γ is actually a normal geodesic of the optimal control system. As a consequence of this discussion and Proposition 3.8, we obtain the following characterisation in terms of the controllability of the linearised system.

Proposition 3.11. *Let $\gamma : [0,T] \to M$ be a geodesic. Then*

(i) γ is strictly normal $\Leftrightarrow$ the linearised system is controllable in time T,

(ii) γ is strongly normal $\Leftrightarrow$ the linearised system is controllable in time t, $\forall\, t \in (0,T]$,

(iii) γ is ample at $t = 0$ $\Leftrightarrow$ the controllability condition of Proposition 3.8 is satisfied at $t = 0$.

In particular (iii)$\Rightarrow$(ii)$\Rightarrow$(i). Moreover, the three properties are equivalent in the analytic case.

The equivalence in the analytic case is a classical fact about the controllability of nonautonomous analytic linear systems. See, for example, [17, Sec. 1.3].

3.3 State-feedback invariance of the flag of an admissible curve

In this section we prove that, albeit the family $\mathscr{F}_\gamma(t)$ depends on the choice of the local trivialization, the flag of an admissible curve at $t = 0$ is invariant by state-feedback transformation, hence it does not depend on the presentation. This also implies that the growth vector of the admissible curve is well-defined (for all t). In this section we use the shorthand $\mathscr{F}_\gamma^i = \mathscr{F}_\gamma^i(0)$, when the flag is evaluated at $t = 0$.

Proposition 3.12. *The flag $\mathscr{F}_\gamma^1 \subset \mathscr{F}_\gamma^2 \subset \ldots \subset T_{x_0}M$ is state-feedback invariant. In particular it does not depend on the presentation of the control system.*

The next corollary is a direct consequence of Proposition 3.12 and Remark 3.3.

Corollary 3.13. *The growth vector of an admissible curve $\mathcal{G}_\gamma(t)$ is state-feedback invariant.*

Proof of Proposition 3.12. Recall that every state-feedback transformation is the composition of pure state and a pure feedback one. For pure state transformations the statement is trivial, since it is tantamount to a change of variables on the manifold. Thus, it is enough to prove the proposition for pure feedback ones. Recall that the subspaces $\mathscr{F}_\gamma^i$ are defined, in terms of a given presentation, as

$$\mathscr{F}_\gamma^i = \operatorname{span}\{C(0), \ldots, C^{(i-1)}(0)\}, \qquad i \geq 1,$$

where the columns of the matrices $C(t)$ are given by the vectors $C_i(t) = (P_{0,t})_*^{-1} f_i(\gamma(t))$. A pure feedback transformation corresponds to a change of presentation. Thus, let

$$\dot{x} = f(x, u) = f_0(x) + \sum_{i=1}^k u_i f_i(x), \qquad \dot{x} = f'(x, u') = f_0'(x) + \sum_{i=1}^k u_i' f_i'(x),$$

related by the pure feedback transformation $u_i' = \psi_i(x, u) = \psi_{i,0}(x) + \sum_{j=1}^k \psi_{i,j}(x) u_j$. In particular (see also Remark 2.7)

$$f_0(x) = f_0'(x) + \sum_{i=1}^k \psi_{i,0}(x) f_i'(x), \qquad f_i(x) = \sum_{j=1}^k \psi_{j,i}(x) f_j'(x). \tag{18}$$

Denote by $A(t), A'(t)$ and $B(t), B'(t)$ the matrices (16) and (17) associated with the two presentations, in some set of coordinates. According to Remark 3.9, $C(t) = M(t)^{-1} B(t)$, where $M(t)$ is the solution of $\dot{M}(t) = A(t) M(t)$, with $M(0) = \mathbb{I}$, and analogous formulae for the "primed" counterparts. In particular, since $C^{(i)}(t) = (-1)^i M(t)^{-1} B_{i+1}(t)$ and $M(0) = M(0)' = \mathbb{I}$, we get

$$\mathscr{F}_\gamma^i = \operatorname{span}\{B_1(0), \ldots, B_i(0)\}, \qquad (\mathscr{F}_\gamma^i)' = \operatorname{span}\{B_1'(0), \ldots, B_i'(0)\}, \tag{19}$$

where $B_i(t)$ and $B_i'(t)$ are the matrices defined in Proposition 3.8 for the two systems. Notice that Eq. (19) is true only at $t = 0$. We prove the following property, which implies our claim: there exists an invertible matrix $\Psi(t)$ such that

$$B_{i+1}(t) = B_{i+1}'(t) \Psi(t) \bmod \operatorname{span}\{B_1'(t), \ldots, B_i'(t)\}, \tag{20}$$

where Eq. (20) is meant column-wise. Indeed, from Eq. (18) we obtain the relations

$$A(t) = A'(t) + B'(t) \Phi(t), \qquad B(t) = B'(t) \Psi(t), \tag{21}$$

where $\Psi(t)$ and $\Phi(t)$ are $k \times k$ and $k \times n$ matrices, respectively, with components

$$\Psi(t)_{i\ell} \doteq \psi_{i,\ell}(x(t)), \qquad \Phi(t)_{i\ell} \doteq \frac{\partial \psi_{i,0}}{\partial x_\ell}(x(t)) + \sum_{j=1}^{k} u_j(t) \frac{\partial \psi_{i,j}}{\partial x_\ell}(x(t)).$$

Notice that, by definition of feedback transformation, $\Psi(t)$ is invertible. We prove Eq. (20) by induction. For $i = 0$, it follows from (21). The induction assumption is (we omit t)

$$B_i = B_i' \Psi + \sum_{j=0}^{i-1} B_j' \Theta_j, \qquad \text{for some time dependent } k \times k \text{ matrices } \Theta_j.$$

Let $X \simeq Y$ denote $X = Y \bmod \operatorname{span}\{B_1', \dots, B_i'\}$, column-wise. Then

$$\begin{aligned} B_{i+1} &= AB_i - \dot{B}_i \simeq \\ &\simeq (A'B_i' - \dot{B}_i')\Psi + \sum_{j=0}^{i-1} (A'B_j' - \dot{B}_j')\Theta_j \simeq B_{i+1}'\Psi. \end{aligned}$$

We used that $A = A' \bmod \operatorname{span}\{B'\}$, hence we can replace A by A'. Moreover all the terms with the derivatives of Θ_j belong to $\operatorname{span}\{B_1', \dots, B_i'\}$. □

3.4 An alternative definition

In this section we present an alternative definition for the flag of an admissible curve, at $t = 0$. The idea is that the flag $\mathscr{F}_\gamma = \mathscr{F}_\gamma(0)$ of a smooth, admissible trajectory γ can be obtained by computing the Lie derivatives along the direction of γ of sections of the distribution, namely elements of $\overline{\mathscr{D}}$. In this sense, the flag of an admissible curve carries informations about the germ of the distribution along the given trajectory.

Let $\gamma : [0, T] \to M$ be a smooth admissible trajectory, such that $x_0 = \gamma(0)$. By definition, this means that there exists a smooth map $u : [0, T] \to \mathbb{U}$ such that $\dot\gamma(t) = f(\gamma(t), u(t))$.

Definition 3.14. We say that $\mathsf{T} \in f_0 + \overline{\mathscr{D}}$ is a smooth admissible extension of $\dot\gamma$ if there exists a smooth section $\sigma : M \to \mathbb{U}$ such that $\sigma(\gamma(t)) = u(t)$ and $\mathsf{T} = f \circ \sigma$.

In other words T is a vector field extending $\dot\gamma$ obtained through the bundle map $f : \mathbb{U} \to TM$ from an extension of the control u (seen as a section of $\mathbb{U}$ over the curve γ). Notice that, if $\dot\gamma(t) = f_0(\gamma(t)) + \sum_{i=1}^k u_i(t) f_i(\gamma(t))$, an admissible extension of $\dot\gamma$ is a smooth field of the form $\mathsf{T} = f_0 + \sum_{i=1}^k \alpha_i f_i$, where $\alpha_i \in C^\infty(M)$ are such that $\alpha_i(\gamma(t)) = u_i(t)$ for all $i = 1, \dots, k$.

With abuse of notation, we employ the same symbol $\mathscr{F}_\gamma^i$ for the following alternative definition.

Definition 3.15. The flag of the admissible curve γ is the sequence of subspaces

$$\mathscr{F}_\gamma^i \doteq \operatorname{span}\{\mathcal{L}_{\mathsf{T}}^j(X)|_{x_0} \,|\, X \in \overline{\mathscr{D}},\, j \leq i-1\} \subset T_{x_0}M, \qquad i \geq 1,$$

where $\mathcal{L}_{\mathsf{T}}$ denotes the Lie derivative in the direction of T.

Notice that, by definition, this is a filtration of $T_{x_0}M$, i.e. $\mathscr{F}_\gamma^i \subset \mathscr{F}_\gamma^{i+1}$, for all $i \geq 1$. Moreover, $\mathscr{F}_\gamma^1 = \mathscr{D}_{x_0}$. In the rest of this section, we show that Definition 3.15 is well posed, and is equivalent to the original Definition 3.1 at $t = 0$.

Proposition 3.16. *Definition 3.15 does not depend on the admissible extension of $\dot\gamma$.*

Proof. Let $\mathscr{F}^i_\gamma$ and $\widetilde{\mathscr{F}}^i_\gamma$ the subspaces obtained via Definition 3.15 with two different extensions T and $\widetilde{\mathsf{T}}$ of $\dot\gamma$, respectively. In particular, the field $V \doteq \widetilde{\mathsf{T}} - \mathsf{T} \in \overline{\mathscr{D}}$ vanishes on the support of γ. We prove that $\widetilde{\mathscr{F}}^i_\gamma = \mathscr{F}^i_\gamma$ by induction. For $i = 1$ the statement is trivial. Then, assume $\widetilde{\mathscr{F}}^i_\gamma = \mathscr{F}^i_\gamma$. Since $\widetilde{\mathscr{F}}^{i+1}_\gamma = \widetilde{\mathscr{F}}^i_\gamma + \mathrm{span}\{\mathcal{L}^i_{\widetilde{\mathsf{T}}}(X)|_{x_0} \,|\, X \in \overline{\mathscr{D}}\}$, it sufficient to prove that

$$\mathcal{L}^i_{\widetilde{\mathsf{T}}}(X) = \mathcal{L}^i_{\mathsf{T}}(X) \bmod \mathscr{F}^i_\gamma, \qquad X \in \overline{\mathscr{D}}. \tag{22}$$

Notice that $\mathcal{L}^i_{\widetilde{\mathsf{T}}}(X) = \mathcal{L}^i_{\mathsf{T}}(X) + W$, where $W \in \mathrm{Vec}(M)$ is the sum of terms of the form

$$W = \mathcal{L}^\ell_{\mathsf{T}}([V, Y]), \qquad \text{for some} \quad Y \in \mathrm{Vec}(M), \quad 0 \leq \ell \leq i-1.$$

In terms of a local set of generators $f_1, \dots, f_k$ of $\overline{\mathscr{D}}$, $V = \sum_{i=1}^k v_j f_j$, where the functions v_i vanish identically on the support of γ, namely $v_j(\gamma(t)) = 0$ for $t \in [0, T]$. Then, an application of the binomial formula for derivations leads to

$$W = \sum_{j=1}^k \mathcal{L}^\ell_{\mathsf{T}}(v_j[f_j, Y]) - \mathcal{L}^\ell_{\mathsf{T}}(Y(v_j) f_j) =$$
$$= \sum_{j=1}^k \sum_{h=0}^\ell \binom{\ell}{h} \left(\mathcal{L}^h_{\mathsf{T}}(v_j) \mathcal{L}^{\ell-h}_{\mathsf{T}}([f_i, Y]) - \mathcal{L}^h_{\mathsf{T}}(Y(v_j)) \mathcal{L}^{\ell-h}_{\mathsf{T}}(f_i) \right).$$

Observe that $\mathcal{L}^h_{\mathsf{T}}(v_j)|_{x_0} = \left.\frac{d^h v_j}{dt^h}\right|_{t=0} = 0$, for all $h \geq 0$. Then, if we evaluate W at x_0, we obtain

$$W|_{x_0} = -\sum_{j=1}^k \sum_{h=0}^\ell \binom{\ell}{h} \mathcal{L}^h_{\mathsf{T}}(Y(v_j))|_{x_0} \mathcal{L}^{\ell-h}_{\mathsf{T}}(f_i).$$

Then, since $0 \leq \ell \leq i-1$, and by the induction hypotesis, $W|_{x_0} \in \mathscr{F}^i_\gamma$ and Eq. (22) follows. ☐

Proposition 3.17. *Definition 3.15 is equivalent to Definition 3.1 at $t = 0$.*

Proof. Recall that, according to Definition 3.1, at $t = 0$

$$\mathscr{F}^i_\gamma = \mathscr{F}^i_\gamma(0) = \mathrm{span}\left\{ \left.\frac{d^j}{dt^j}\right|_{t=0} v(t) \,\middle|\, v(t) \in \mathscr{F}_\gamma(t) \text{ smooth}, \ j \leq i-1 \right\} \subset T_{x_0}M, \qquad i \geq 1.$$

where $\mathscr{F}_\gamma(t) = (P_{0,t})^{-1}_* \mathscr{D}_{\gamma(t)}$. By Proposition 3.12, the flag at $t = 0$ is state-feedback invariant. Then, up to a (local) pure feedback transformation, we assume that the fixed smooth admissible trajectory $\gamma : [0, T] \to M$ is associated with a constant control, namely $\dot\gamma(t) = f_0(\gamma(t)) + \sum_{i=1}^k u_i f_i(\gamma(t))$, where $u \in L^\infty([0, T], \mathbb{R}^k)$ is constant. In this case, the flow $P_{0,t} : M \to M$ is actually the flow of the autonomous vector field $\mathsf{T} \doteq f_0 + \sum_{i=1}^k u_i f_i$, that is $P_{0,t} = e^{t\mathsf{T}}$.

Indeed $\mathsf{T} \in f_0 + \overline{\mathscr{D}}$ is an admissible extension of $\dot\gamma$. Moreover, any smooth $v(t) \in \mathscr{F}_\gamma(t)$ is of the form $v(t) = (P_{0,t})^{-1}_* X|_{\gamma(t)}$, where $X \in \overline{\mathscr{D}}$. Then

$$\left.\frac{d^j}{dt^j}\right|_{t=0} v(t) = \left.\frac{d^j}{dt^j}\right|_{t=0} (P_{0,t})^{-1}_* X|_{\gamma(t)} = \left.\frac{d^j}{dt^j}\right|_{t=0} e^{-t\mathsf{T}}_* X|_{\gamma(t)} = \mathcal{L}^j_{\mathsf{T}}(X)|_{x_0},$$

where in the last equality we have employed the definition of Lie derivative. ☐

Remark 3.18. To end this section, observe that, for any equiregular smooth admissible curve $\gamma : [0,T] \to M$, the Lie derivative in the direction of the curve defines surjective linear maps

$$\mathcal{L}_{\mathsf{T}} : \mathscr{F}^i_{\gamma(t)}/\mathscr{F}^{i-1}_{\gamma(t)} \to \mathscr{F}^{i+1}_{\gamma(t)}/\mathscr{F}^i_{\gamma(t)}, \qquad i \geq 1,$$

for any fixed $t \in [0,T]$ as follows. Let $\mathsf{T} \in \mathrm{Vec}(M)$ be any admissible extension of $\dot{\gamma}$. Similarly, for $X \in \mathscr{F}^i_{\gamma(t)}$, consider a smooth extension of X along the curve γ such that $X|_{\gamma(s)} \in \mathscr{F}^i_{\gamma(s)}$ for all $s \in [0,T]$. Then we define

$$\mathcal{L}_{\mathsf{T}}(X) := [T,X]|_{\gamma(t)} \bmod \mathscr{F}^i_{\gamma(t)}, \qquad t \in [0,T].$$

The proof that $\mathcal{L}_{\mathsf{T}}$ does not depend on the choice of the admissible extension T is the same of Proposition 3.16 and for this reason we omit it. The fact that it depends only on the value of $X \bmod \mathscr{F}^{i-1}_{\gamma(t)}$ at the point $\gamma(t)$ is similar, under the equiregularity assumption.

In particular, notice that the maps $\mathcal{L}^i_{\mathsf{T}} : \mathscr{D}_{\gamma(t)} \to \mathscr{F}^{i+1}_{\gamma(t)}/\mathscr{F}^i_{\gamma(t)}$, for $i \geq 1$, are well defined, surjective linear maps from the distribution (see also point (iv) of Remark 3.4).

4 Geodesic cost and its asymptotics

In this section we define the *geodesic cost function* and we state the main result about the existence of its asymptotics (see Theorem A). This paves the way for the definition of curvature of an affine optimal control system (see Theorem B).

4.1 Geodesic cost

Definition 4.1. Let $x_0 \in M$ and consider a strongly normal geodesic $\gamma : [0,T] \to M$ such that $\gamma(0) = x_0$. The *geodesic cost* associated with γ is the family of functions

$$c_t(x) \doteq -S_t(x,\gamma(t)), \qquad x \in M,\ t > 0,$$

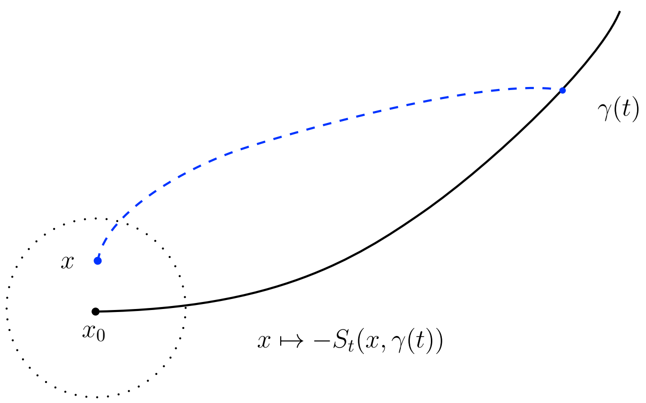


Figure 3: The geodesic cost function.

The geodesic cost function is smooth in a neighbourhood of x_0, and for $t > 0$ sufficiently small. More precisely, Theorem 2.19, applied to the geodesic cost, can be rephrased as follows.

Theorem 4.2. *Let $x_0 \in M$ and $\gamma : [0,T] \to M$ be a strongly normal geodesic such that $\gamma(0) = x_0$. Then there exist $\varepsilon > 0$ and an open set $U \subset (0,\varepsilon) \times M$ such that*

(i) $(t, x_0) \in U$ *for all* $t \in (0,\varepsilon)$*,*

(ii) The geodesic cost function $(t, x) \mapsto c_t(x)$ is smooth on U.

Moreover, for any $(t, x) \in U$, there exists a unique (normal) minimizer of the cost functional J_t, among all the admissible curves that connect x with $\gamma(t)$.

In the following, $\dot{c}_t$ denotes the derivative of the geodesic cost with respect to t.

Proposition 4.3. *Under the assumptions above, $d_{x_0} c_t = \lambda_0$, for all $t \in (0, \varepsilon)$. In particular x_0 is a critical point for the function $\dot{c}_t$ for all $t \in (0, \varepsilon)$.*

Proof. First observe that, in general, if $\gamma(t)$ is an admissible curve for an affine control system, the "reversed" curve $\widetilde{\gamma}(t) \doteq \gamma(T - t)$ is no longer admissible. As a consequence, the value function $(x_0, x_1) \mapsto S_T(x_0, x_1)$ is not symmetric and we cannot directly apply Lemma 2.20 To compute the differential of the value function $x \mapsto -S_t(x, \gamma(t))$ at x_0. Nevertheless, we can still exploit Lemma 2.20, by passing to an associated control problem with reversed dynamic.

Lemma 4.4. *Consider the control system with reversed dynamic*

$$\dot{x} = \widetilde{f}(x, u), \qquad x \in M, \qquad \widetilde{f}(x, u) \doteq -f(x, u),$$
$$J_T(u) \to \min .$$

Let $\widetilde{S}_T$ be the value function of this problem. Then $\widetilde{S}_T(x_0, x_1) = S_T(x_1, x_0)$, for all $x_0, x_1 \in M$.

Proof of Lemma 4.4. It is easy to see that the map $\gamma(t) \mapsto \widetilde{\gamma}(t) \doteq \gamma(T - t)$ defines a one-to-one correspondence between admissible curves for the two problems. Moreover, if γ is associated with the control u, then $\widetilde{\gamma}$ is associated with control $\widetilde{u}(t) \doteq u(T-t)$. Since the cost is invariant by this transformation, one has $\widetilde{S}_T(x_1, x_0) = S_T(x_0, x_1)$. Notice that this transformation preserves normal and abnormal trajectories and minimizers. ☐

The Hamiltonian of the reversed system is $\widetilde{H}(\lambda) = H(-\lambda)$. Let $i : T^*M \to T^*M$ be the fiberwise linear map $\lambda \mapsto -\lambda$. Then, $i_* \vec{\widetilde{H}}(\lambda) = -\vec{H}(-\lambda)$ (i.e. $\vec{\widetilde{H}}$ is i_*-related with $-\vec{H}$). This implies that, if $\lambda(t)$ is the lift of the geodesic $\gamma(t)$ for the original system, then $\widetilde{\lambda}(t) \doteq -\lambda(T-t)$ is the lift of the geodesic $\widetilde{\gamma}(t) = \gamma(T-t)$ for the reversed system. In particular, the final covector of the reversed geodesic $\widetilde{\lambda}_T = \widetilde{\lambda}(T) = -\lambda(0) = -\lambda_0$ is equal to minus the initial covector of the original geodesic. Thus, we can apply Lemma 2.20 and obtain

$$d_{x_0} c_T = -d_{x_0} S_T(\cdot, \gamma(T)) = -d_{x_0}(\widetilde{S}_T(\gamma(T), \cdot)) = -\tilde{\lambda}(T) = \lambda_0.$$

where $\tilde{\gamma} : [0, T] \to M$ is the unique strictly normal minimizer of the cost functional $\widetilde{J}_T = J_T$ of the reversed system such that $\tilde{\gamma}(0) = \gamma(T)$ and $\tilde{\gamma}(T) = x_0$. ☐

4.2 Hamiltonian inner product

In this section we introduce an inner product on the distribution, which depends on a given geodesic. Namely, it is induced by the second derivative of Hamiltonian of the control system at a point $\lambda \in T^*M$, associated with a geodesic.

A non-negative definite quadratic form, defined on the dual of a vector space V^*, induces an inner product on a subspace of V as follows. Recall first that a quadratic form can be defined as a self-adjoint linear map $B : V^* \to V$. B is non-negative definite if, for all $\lambda \in V^*$, $\langle \lambda, B(\lambda) \rangle \geq 0$. Let us define a bilinear map on $\mathrm{Im}\,(B) \subset V$ by the formula

$$\langle w_1 | w_2 \rangle_B \doteq \langle \lambda_1, B(\lambda_2) \rangle, \qquad \text{where } w_i = B(\lambda_i).$$

It is easy to prove that $\langle\cdot|\cdot\rangle_B$ is symmetric and does not depend on the representatives λ_i. Moreover, since B is non-negative definite, $\langle\cdot|\cdot\rangle_B$ is an inner product on $\mathrm{Im}\,(B)$.

Now we go back to the general setting. Fix a point $x \in M$, consider the restriction of the Hamiltonian H to the fiber $H_x \doteq H|_{T_x^*M}$ and denote by $d^2_\lambda H_x$ its second derivative at the point $\lambda \in T_x^*M$. We show that $d^2_\lambda H_x$ is a non-negative quadratic form and, as a self-adjoint linear map $d^2_\lambda H_x : T_x^*M \to T_xM$, its image is exactly the distribution at the base point.

Lemma 4.5. *For every $\lambda \in T_x^*M$, $d^2_\lambda H_x$ is non-negative definite and* $\mathrm{Im}\,(d^2_\lambda H_x) = \mathscr{D}_x$.

Proof. We prove the result by computing an explicit expression for $d^2_\lambda H_x$ in coordinates $\lambda = (p, x)$ on T^*M. Recall that the maximized Hamiltonian H is defined by the identity

$$H(p, x) = \mathcal{H}(p, x, \bar{u}) = \langle p, f_0(x)\rangle + \sum_{i=1}^k \bar{u}_i \langle p, f_i(x)\rangle - L(x, \bar{u}),$$

where $\bar{u} = \bar{u}(p, x)$ is the solution of the maximality condition

$$\langle p, f_i(x)\rangle = \frac{\partial L}{\partial u_i}(x, \bar{u}(p, x)), \qquad i = 1, \ldots, k. \tag{23}$$

By the chain rule, we obtain

$$\frac{\partial H}{\partial p}(p, x) = f_0(x) + \sum_{i=1}^k \bar{u}_i f_i(x) + \underbrace{\frac{\partial \bar{u}_i}{\partial p}\langle p, f_i(x)\rangle - \frac{\partial L}{\partial u_i}\frac{\partial \bar{u}_i}{\partial p}}_{=0}.$$

By differentiating Eq. (23) with respect to p, we get

$$f_i(x) = \sum_{j=1}^k \frac{\partial^2 L}{\partial u_i \partial u_j}\frac{\partial \bar{u}_j}{\partial p}, \qquad i = 1, \ldots, k.$$

Finally, we compute the second derivatives matrix

$$\frac{\partial^2 H}{\partial p^2}(p, x) = \sum_{i=1}^k \frac{\partial \bar{u}_i}{\partial p} f_i^*(x) = \sum_{i,j=1}^k f_i(x)\left(\frac{\partial^2 L}{\partial u_i \partial u_j}\right)^{-1} f_j^*(x). \tag{24}$$

Since the Hessian of L (with respect to u) is positive definite, Eq. (24) implies that $d^2_\lambda H_x$ is non-negative definite and $\mathrm{Im}\, d^2_\lambda H_x \subset \mathscr{D}_x$. Moreover, it is easy to see that rank $\frac{\partial^2 H}{\partial p^2} = \dim \mathscr{D}_x$, therefore $\mathrm{Im}\,(d^2_\lambda H_x) = \mathscr{D}_x$. □

Definition 4.6. For any $\lambda \in T_x^*M$, the *Hamiltonian inner product* (associated with λ) is the inner product $\langle\cdot|\cdot\rangle_\lambda$ induced by $d^2_\lambda H_x$ on $\mathscr{D}_x$.

Remark 4.7. We stress that, for any fixed $x \in M$, the subspace $\mathscr{D}_x \subset T_xM$, where the inner product $\langle\cdot|\cdot\rangle_\lambda$ is defined, does not depend on the choice of the element λ in the fiber T_x^*M. When H_x itself is a quadratic form, $d^2_\lambda H_x = 2H_x$ for every $\lambda \in T_x^*M$. Therefore, the inner product $\langle\cdot|\cdot\rangle_\lambda$ does not depend on the choice of $\lambda \in T_xM$. This is the case, for example, of an optimal control system defined by a sub-Riemannian structure, in which the inner product just defined is precisely the sub-Riemannian one (see Section 5).

4.3 Asymptotics of the geodesic cost function and curvature

Let $f : M \to \mathbb{R}$ be a smooth function defined on a smooth manifold M. Its first differential at a point $x \in M$ is the linear map $d_x f : T_x M \to \mathbb{R}$.

The *second differential* of f, as a symmetric bilinear form, is well defined only at a critical point, i.e. at those points $x \in M$ such that $d_x f = 0$. Indeed, in this case, the map

$$d_x^2 f : T_x M \times T_x M \to \mathbb{R}, \qquad d_x^2 f(v, w) = V(W(f))(x),$$

where V, W are vector fields such that $V(x) = v$ and $W(x) = w$, respectively, is a well defined symmetric bilinear form which does not depend on the choice of the extensions.

The quadratic form associated with the second differential of f at x which, for simplicity, we denote by the same symbol $d_x^2 f : T_x M \to \mathbb{R}$, is

$$d_x^2 f(v) = \left.\frac{d^2}{dt^2}\right|_{t=0} f(\gamma(t)), \qquad \gamma(0) = x, \quad \dot\gamma(0) = v.$$

Now, for $\lambda \in T_{x_0}^* M$, consider the geodesic cost function associated with the strongly normal geodesic $\gamma(t) = \mathcal{E}_{x_0}(t, \lambda)$, starting from x_0. By Proposition 4.3, for every $t \in (0, \varepsilon)$, the function $x \mapsto \dot c_t(x)$ has a critical point at x_0. Hence we can consider the family of quadratic forms defined on the distribution

$$d_{x_0}^2 \dot c_t|_{\mathscr{D}_{x_0}} : \mathscr{D}_{x_0} \to \mathbb{R}, \qquad t \in (0, \varepsilon),$$

obtained by the restriction of the second differential of $\dot c_t$ to the distribution $\mathscr{D}_{x_0}$. Then, using the inner product $\langle \cdot | \cdot \rangle_\lambda$ induced by $d_\lambda^2 H_x$ on $\mathscr{D}_x$ introduced in Section 4.2, we associate with this family of quadratic forms the family of symmetric operators on the distribution $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ defined by the identity

$$d_{x_0}^2 \dot c_t(v) \doteq \langle \mathcal{Q}_\lambda(t) v | v \rangle_\lambda, \qquad t \in (0, \varepsilon),\ v \in \mathscr{D}_{x_0}. \tag{25}$$

The assumption that the geodesic is strongly normal ensures the smoothness of $\mathcal{Q}_\lambda(t)$ for small $t > 0$. If the geodesic is also ample, we have a much stronger statement about the asymptotic behaviour of $\mathcal{Q}_\lambda(t)$ for $t \to 0$.

Theorem A. *Let $\gamma : [0, T] \to M$ be an ample geodesic with initial covector $\lambda \in T_{x_0}^* M$, and let $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ be defined by* (25)*. Then $t \mapsto t^2 \mathcal{Q}_\lambda(t)$ can be extended to a smooth family of operators on $\mathscr{D}_{x_0}$ for small $t \geq 0$, symmetric with respect to $\langle \cdot | \cdot \rangle_\lambda$. Moreover,*

$$\mathcal{I}_\lambda \doteq \lim_{t \to 0^+} t^2 \mathcal{Q}_\lambda(t) \geq \mathbb{I} > 0,$$

as operators on $(\mathscr{D}_{x_0}, \langle \cdot | \cdot \rangle_\lambda)$. Finally

$$\left.\frac{d}{dt}\right|_{t=0} t^2 \mathcal{Q}_\lambda(t) = 0.$$

As a consequence of Theorem A we are allowed to introduce the following definitions.

Definition 4.8. Let $\lambda \in T_{x_0}^* M$ be the initial covector associated with an ample geodesic. The *curvature* is the symmetric operator $\mathcal{R}_\lambda : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ defined by

$$\mathcal{R}_\lambda \doteq \frac{3}{2} \left.\frac{d^2}{dt^2}\right|_{t=0} t^2 \mathcal{Q}_\lambda(t).$$

The *Ricci curvature* at $\lambda \in T_{x_0}^* M$ is defined by $\mathrm{Ric}(\lambda) \doteq \operatorname{tr} \mathcal{R}_\lambda$.

In particular, we have the following Laurent expansion for the family of symmetric operators $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$:

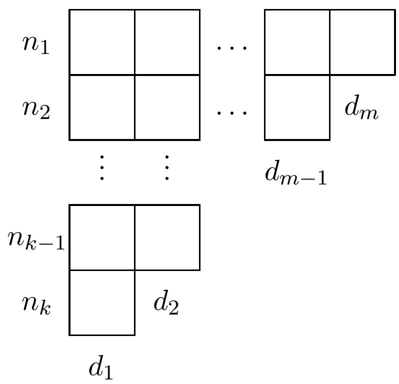
$$\mathcal{Q}_\lambda(t) = \frac{1}{t^2}\mathcal{I}_\lambda + \frac{1}{3}\mathcal{R}_\lambda + O(t), \qquad t > 0. \tag{26}$$

The normalization factor 1/3 appearing in (26) in front of the operator $\mathcal{R}_\lambda$ is necessary for recovering the sectional curvature in the case of a control system defined by a Riemannian structure (see Section 4.4.2). We stress that, by construction, $\mathcal{I}_\lambda$ and $\mathcal{R}_\lambda$ are operators on the distributions, symmetric with respect to the inner product $\langle \cdot | \cdot \rangle_\lambda$.

4.3.1 Spectrum of $\mathcal{I}_\lambda$ for equiregular geodesics

Under the assumption that the geodesic is also equiregular, we can completely characterize the operator $\mathcal{I}_\lambda$, namely compute its spectrum.

Let us consider the growth vector $\mathcal{G}_\gamma = \{k_1, k_2, \dots, k_m\}$ of the geodesic γ which, by the equiregularity assumption, does not depend on t. Let $d_i \doteq \dim \mathscr{F}^i_\gamma - \dim \mathscr{F}^{i-1}_\gamma = k_i - k_{i-1}$, for $i = 1, \dots, m$ (where $k_0 \doteq 0$). Recall that d_i is a decreasing sequence (see (iv) of Remark 3.4). Then we can build a tableau with m columns of length d_i, for $i = 1, \dots, m$, as follows:



$$\sum_{i=1}^m d_i = n = \dim M,$$
$$d_1 = k_1 = k \doteq \dim \mathscr{D}_{x_0}.$$

Finally, for $j = 1, \dots, k$, let n_j be the length of the j-th row of the tableau.

Theorem B. *Let $\gamma : [0, T] \to M$ be an ample and equiregular geodesic with initial covector $\lambda \in T^*_{x_0}M$. Then the symmetric operator $\mathcal{I}_\lambda : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ satisfies*

(i) $\operatorname{spec} \mathcal{I}_\lambda = \{n_1^2, \dots, n_k^2\}$,

(ii) $\operatorname{tr} \mathcal{I}_\lambda = n_1^2 + \dots + n_k^2$.

Remark 4.9. Notice that, although the family $\mathcal{Q}_\lambda(t)$ depends on the cost function, the operator $\mathcal{I}_\lambda$ depends only on the growth vector $\mathcal{G}_\gamma$, which is a state-feedback invariant not related with the cost. This is a consequence of the results of Section 3.3 and Theorem B.

Remark 4.10. By the classical identity $\sum_{i=1}^n (2i-1) = n^2$, we rewrite the trace of $\mathcal{I}_\lambda$ as follows

$$\operatorname{tr} \mathcal{I}_\lambda = \sum_{i=1}^m (2i-1)(\dim \mathscr{F}^i_\gamma - \dim \mathscr{F}^{i-1}_\gamma).$$

Notice that the right hand side of the above equation makes sense also for a non-equiregular (tough still ample) geodesic, where the dimensions are computed at $t = 0$. This number also appears in Section 5, under the name of *geodesic dimension*, in connection with the asymptotics of the volume growth in sub-Riemannian geometry.

The proofs of Theorems A and B are postponed to Section 7, upon the introduction of the required technicals tools.

4.4 Examples

In this section, we discuss two relevant examples, namely an autonomous linear control system on $\mathbb{R}^n$ with quadratic cost and a Riemannian manifold. In the first example it is possible to compute $\mathcal{Q}_\lambda$ and its expansion, by a direct manipulation of the cost geodesic function. In the second example, we recover the sectional curvature of the Riemannian manifold.

4.4.1 Linear-quadratic control problem

Let us consider a classical linear-quadratic control system. Namely $M = \mathbb{R}^n$, $\mathbb{U} = \mathbb{R}^n \times \mathbb{R}^k$ and $f(x,u) = x + Bu$ is linear both in the state and in the control variables. Admissible curves are solutions of

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

where A and B are two $n \times n$ and $n \times k$ matrices, respectively. The cost of an admissible trajectory associated with u is proportional to the square of the L^2-norm of the control

$$J_T(u) = \frac{1}{2}\int_0^T u(t)^* u(t) dt.$$

Since $u : [0,T] \to \mathbb{R}^k$ is measurable and essentially bounded, the trajectory $x(t;x_0)$ associated with u such that $x(0;x_0) = x_0$ is explicitly computed by the Cauchy formula

$$x(t;x_0) = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)ds.$$

In this case, the bracket-generating condition (A1) is the classical Kalman controllability condition, and reads

$$\operatorname{span}\{B, AB, \dots, A^m B\} = \mathbb{R}^n. \tag{27}$$

Since the system is linear, the linearisation along any admissible trajectory coincides with the system itself. Hence it follows that any geodesic is ample and equiregular. In fact, the geodesic growth vector is the same for any geodesic, and is equal to

$$\dim \mathscr{F}^i = \operatorname{rank}\{B, AB, \dots, A^{i-1}B\}, \qquad i \geq 1.$$

A standard computation shows that, under the assumption (27), there are no abnormal trajectories. Let us introduce canonical coordinates $(p,x) \in T^*\mathbb{R}^n \simeq \mathbb{R}^{n*} \times \mathbb{R}^n$. Here, it is convenient to treat $p \in \mathbb{R}^{n*}$ as a row vector, and $x \in \mathbb{R}^n$, $u \in \mathbb{R}^k$ as column vectors. The Hamiltonian of the system for normal extremals is

$$\mathcal{H}(p,x,u) = pAx + pBu - \frac{1}{2}u^* u.$$

The maximality condition gives $\bar{u}(p,x) = B^* p^*$. Then, the maximized Hamiltonian is

$$H(p,x) = pAx + \frac{1}{2}pBB^*p^*.$$

For a normal trajectory with initial covector $\lambda = (p_0, x_0)$, we have $p(t;x_0,p_0) = p_0 e^{-tA}$ and

$$x(t;x_0,p_0) = e^{tA}x_0 + e^{tA}\int_0^t e^{-sA}BB^*e^{-sA^*}ds\, p_0^*. \tag{28}$$

Let us denote by $C(t)$ the controllability matrix

$$C(t) \doteq \int_0^t e^{-sA}BB^*e^{-sA^*}ds.$$

By Eq. (28), we can compute the optimal cost to reach the point $\widetilde{x}(t) = x(t; x_0, p_0)$, starting at point x (close to x_0), in time t, as follows

$$c_t(x) = -S_t(x, \widetilde{x}(t)) = -\frac{1}{2}p_0 C(t) p_0^* + p_0(x - x_0) - \frac{1}{2}(x - x_0)^* C(t)^{-1}(x - x_0).$$

Thus, $d_x^2 \dot{c}_t = -\frac{d}{dt}C(t)^{-1}$, and the family of quadratic forms $\mathcal{Q}_\lambda$, written in terms of the basis defined by the columns of B, is represented by the matrix

$$\mathcal{Q}_\lambda(t) = -B^* \frac{d}{dt} C(t)^{-1} B.$$

The operator $\mathcal{I}_\lambda$ is completely determined by Theorem B. Its eigenvalues coincide with the squares of the Kronecker indices of the control system (see [9]). Moreover, the curvature $\mathcal{R}_\lambda$ is

$$\mathcal{R}_\lambda = -\frac{3}{2}\left.\frac{d^2}{dt^2}\right|_{t=0} \left(t^2 B^* \frac{d}{dt} C(t)^{-1} B\right) = -\frac{3}{2}\left.\frac{d^2}{dt^2}\right|_{t=0} \left(t B^* C(t)^{-1} B\right).$$

We stress that, for this specific case, the operators $\mathcal{I}_\lambda$ and $\mathcal{R}_\lambda$ do not depend neither on the geodesic nor on the initial point since the system is linear (hence it coincides with its linearisation along any geodesic starting at any point).

Remark 4.11. With straightforward but long computations one can generalize these formulae to the case of a quadratic cost with a potential of the form

$$J_T(u) = \frac{1}{2}\int_0^T u(t)^* u(t) + x_u(t)^* Q x_u(t) dt,$$

where Q is a symmetric $n \times n$ matrix, and $x_u(t)$ is the trajectory associated with the control u.

4.4.2 Riemannian geometry

In this example we characterize the operators $\mathcal{Q}_\lambda$ and $\mathcal{I}_\lambda$ for an optimal control system associated with a Riemannian structure. In particular, we show how the operator $\mathcal{R}_\lambda$ is related with the classical sectional curvature.

Let M be an n-dimensional Riemannian manifold. In this case, $\mathbb{U} = TM$, and $f : TM \to TM$ is the identity bundle map. Let $f_1, \dots, f_n$ be a local orthonormal frame for the Riemannian structure. Any Lipschitz curve on M is admissible, and is a solution of

$$\dot{x} = \sum_{i=1}^n u_i f_i(x), \qquad x \in M,\ u \in \mathbb{R}^n.$$

The cost functional, whose extremals are the classical Riemannian geodesics, is

$$J_T(u) = \frac{1}{2}\int_0^T \sum_{i=1}^n u_i(t)^2 dt.$$

Every geodesic is ample and equiregular, and has trivial growth vector $\mathcal{G}_\gamma = \{n\}$ since, for all $x \in M$, $\mathscr{D}_x = T_x M$. Moreover, the Hamiltonian inner product is equal to the Riemannian inner product. As a standard consequence of the Cauchy-Schwartz inequality, and the fact that Riemannian geodesics have constant speed, the value function S_T can be written in terms of the Riemannian distance $\mathsf{d} : M \times M \to \mathbb{R}$ as follows

$$S_T(x, y) = \frac{1}{2T}\mathsf{d}^2(x, y), \qquad x, y \in M.$$

To any initial covector $\lambda \in T^*_{x_0}M$ corresponds, via the Riemannian structure, an initial vector $v \in T^*_{x_0}M$. We call $\gamma_v : [0,T] \to M$ the associated geodesic, such that $\gamma_v(0) = x_0$ and $\dot\gamma_v(0) = v$. Thus, the geodesic cost function associated with γ_v is

$$c_t(x) = -\frac{1}{2t}\mathsf{d}^2(x, \gamma_v(t)).$$

Then, in order to compute the operators $\mathcal{I}_\lambda$ and $\mathcal{R}_\lambda$ we essentially need an asymptotic expansion of the "squared distance from a geodesic".

Let $\gamma_v(t)$, $\gamma_w(s)$ be two arclength parametrized geodesics, with initial vectors $v, w \in T_{x_0}M$, respectively, starting from x_0. Let us define the function $C(t,s) \doteq \frac{1}{2}d^2(\gamma_v(t), \gamma_w(s))$. It is well known that C is smooth at $(0,0)$ (this is not true in more general settings, such as sub-Riemannian geometry).

Lemma 4.12. *The following formula holds true for the Taylor expansion of $C(t,s)$ at $(0,0)$*

$$C(t,s) = \frac{1}{2}\left(t^2 + s^2 - 2\langle v|w\rangle ts\right) - \frac{1}{6}\langle R(v,w)v|w\rangle t^2 s^2 + t^2 s^2 o(|t| + |s|), \tag{29}$$

where $\langle\cdot|\cdot\rangle$ denotes the Riemannian inner product and R is the Riemann tensor.

Proof. Since the geodesics γ_v and γ_w are parametrised by arclength, we have

$$C(t,0) = t^2/2, \qquad C(0,s) = s^2/2, \qquad \forall\, t,s \geq 0. \tag{30}$$

Moreover, by standard computations, we obtain

$$\frac{\partial C}{\partial s}(t,0) = -t\langle v|w\rangle, \qquad \frac{\partial C}{\partial t}(0,s) = -s\langle v|w\rangle, \qquad \forall\, t,s \geq 0. \tag{31}$$

Eqs. (30) and (31) imply that the monomials t^n, st^n, s^n, ts^n with $n \geq 2$ do not appear in the Taylor polynomial. The statement is then reduced to the following identity:

$$-\frac{3}{2}\frac{\partial^4 C}{\partial t^2 \partial s^2}(0,0) = \langle R(v,w)v|w\rangle.$$

This identity appeared for the first time in [23, Th. 8.3], in the context of the Ma-Trudinger-Wang curvature tensor, and also in [35, Eq. 14.1]. For a detailed proof one can see also [18, Prop. 1.5.1]. Essentially, this is the very original definition of curvature introduced by Riemann in his famous *Habilitationsvortrag* (see [29]). ☐

Finally, fix $w \in T_{x_0}M$, and we compute the quadratic form $\langle \mathcal{Q}_\lambda(t)w|w\rangle = d^2_{x_0}\dot c_t(w)$

$$\begin{aligned} d^2_{x_0}\dot c_t(w) &= \left.\frac{\partial^2}{\partial s^2}\right|_{s=0} \frac{\partial}{\partial t}\left(-\frac{1}{2t}d^2(\gamma_v(t), \gamma_w(s))\right) = \frac{\partial}{\partial t}\left(-\frac{1}{t}\frac{\partial^2 C}{\partial s^2}(t,0)\right) = \\ &= \frac{1}{t^2}\frac{\partial^2 C}{\partial s^2}(0,0) + \frac{1}{3}\left(-\frac{3}{2}\frac{\partial^4 C}{\partial t^2 \partial s^2}(0,0)\right) + O(t), \end{aligned}$$

where, in the first equality, we can exchange the order of derivations by the smoothness of $C(t,s)$. By Lemma 4.12 we get $\mathcal{I}_\lambda = \mathbb{I}$ and $\mathcal{R}_\lambda = R(v,\cdot)v$. In particular, $\langle \mathcal{R}_\lambda w|w\rangle = \langle R(v,w)v|w\rangle$ is the Riemannian sectional curvature in the plane generated by $v, w \in T_{x_0}M$.

Remark 4.13. The relation between the operator $\mathcal{R}_\lambda$ and the Riemannian curvature tensor in the Riemannian setting was originally recovered in [5] by using the formalism of Jacobi curves (which we introduce in Section 6).

Remark 4.14. In Section 5, we apply our theory to the sub-Riemannian setting, where an analogue approach, leading to the Taylor expansion of Eq. (29) is not possible, for two major differences between the Riemannian and sub-Riemannian setting. First, geodesics cannot be parametrized by their initial tangent vector. Second, and crucial, for every $x_0 \in M$, the sub-Riemannian squared distance $x \mapsto \mathsf{d}^2(x_0, x)$ is *never* smooth at x_0.

4.4.3 Finsler geometry

The notion of curvature introduced in this paper recovers not only the classical sectional curvature of Riemannian manifolds, but also the notion of *flag curvature* of Finsler manifolds. These structures can be realized as optimal control problems (in the sense of Section 2) by the choice $\mathbb{U} = TM$ and $f : TM \to TM$ equal to the identity bundle map. Moreover the Lagrangian is of the form $L = F^2/2$, where $F \in C^\infty(TM \setminus 0_{TM})$ (0_{TM} is the zero section), is non-negative and positive-homogeneous, i.e. $F(cv) = cF(v)$ for all $v \in TM$ and $c > 0$. Finally L satisfies the Tonelli assumption (A2).

In this setting, it is common to introduce the isomorphism $\tau^* : T^*M \to TM$ (the inverse *Legendre transform*) defined by

$$\tau^*(\lambda) \doteq d_\lambda H_x, \qquad \lambda \in T_x^*M,$$

where H_x is the restriction to the fiber T_x^*M of the Hamiltonian H of the system.

In this case, for all $x \in M$, $\mathscr{D}_x = T_xM$, and the operator $\mathcal{R}_\lambda : T_xM \to T_xM$ can be identified with the Finsler flag curvature operator $R_v^F : T_xM \to T_xM$, where $v = \tau^*(\lambda)$ is the flagpole. A more detailed discussion of Finsler structure and the aforementioned correspondence one can see, for instance, the recent work [27, Example 5.1].

5 Sub-Riemannian geometry

In this section we focus on the sub-Riemannian setting. After a brief introduction, we discuss the existence of ample geodesics, the regularity of the geodesic cost and the homogeneity properties of the family $\mathcal{Q}_\lambda$. Then we state the main result of this section about the sub-Laplacian of the sub-Riemannian distance. Finally, we define the concept of geodesic dimension and we investigate the asymptotic rate of growth of the volume of measurable set under sub-Riemannian geodesic homotheties.

5.1 Basic definitions

Sub-Riemannian structures are particular affine optimal control system, in the sense of Definition 2.1, where the "drift" vector field is zero and the Lagrangian L is induced by an Euclidean structure on the control bundle $\mathbb{U}$. For a general introduction to sub-Riemannian geometry from the control theory viewpoint we refer to [3]. Other classical references are [15, 26].

Definition 5.1. Let M be a connected, smooth n-dimensional manifold. A *sub-Riemannian structure* on M is a pair $(\mathbb{U}, f)$ where:

- (i) $\mathbb{U}$ is a smooth rank k *Euclidean* vector bundle with base M and fiber $\mathbb{U}_x$, i.e. for every $x \in M$, $\mathbb{U}_x$ is a k-dimensional vector space endowed with an inner product.
- (ii) $f : \mathbb{U} \to TM$ is a smooth *linear* morphism of vector bundles, i.e. f is *linear* on fibers and the following diagram is commutative:

$$\begin{array}{ccc} \mathbb{U} & \xrightarrow{f} & TM \\ & \searrow^{\pi_{\mathbb{U}}} & \downarrow \pi \\ & & M \end{array}$$

The maps $\pi_{\mathbb{U}}$ and π are the canonical projections of the vector bundles $\mathbb{U}$ and TM, respectively. Notice that once we have chosen a local trivialization for the vector bundle $\mathbb{U}$, i.e. $\mathbb{U} \simeq M \times \mathbb{R}^k$, we can choose a basis in the fibers and the map f reads $f(x, u) = \sum_{i=1}^k u_i f_i(x)$.

It is always possible to reduce to the case when the control bundle $\mathbb{U}$ is trivial without changing the sub-Riemannian structure defined on it (see [3,30]). In particular it is not restrictive to assume that the vector fields $f_1, \ldots, f_k$ are globally defined.

Remark 5.2. There is no assumption on the rank of the function f. In other words if we consider, in some choice of the trivialization of $\mathbb{U}$, the vector fields $f_1, \ldots, f_k$, they could be linearly dependent at some (or even at every) point. The structure is Riemannian if and only if $\dim \mathscr{D}_x = n$ for all $x \in M$.

Remark 5.3 (On the notation). Throughout this section, to adhere to the standard notation of the sub-Riemannian literature, we use the notation $X_i = f_i$ for the set of (local) vector fields which define the sub-Riemannian structure.

The Euclidean structure on the fibers induces a metric structure on the *distribution* $\mathscr{D}_x = f(\mathbb{U}_x)$ for all $x \in M$ as follows:

$$\|v\|_x^2 \doteq \min \left\{ \|u\|^2 \,\middle|\, v = f(x,u) \right\}, \qquad \forall\, v \in \mathscr{D}_x. \tag{32}$$

It is possible to show that $\|\cdot\|_x$ is a norm on $\mathscr{D}_x$ that satisfies the parallelogram law, i.e. it is actually induced by an inner product $\langle\cdot|\cdot\rangle_x$ on $\mathscr{D}_x$. Notice that the minimum in (32) is always attained since we are minimizing an Euclidean norm in $\mathbb{R}^k$ on an affine subspace.

An admissible trajectory for the sub-Riemannian structure is also called *horizontal*, i.e. a Lipschitz curve $\gamma : [0,T] \to M$ such that

$$\dot\gamma(t) = f(\gamma(t), u(t)), \qquad \text{a.e. } t \in [0,T],$$

for some measurable and essentially bounded map $u : [0,T] \to \mathbb{R}^k$.

Remark 5.4. Given an admissible trajectory it is pointwise defined its *minimal control* $u : [0,T] \to \mathbb{R}^k$ such that $\|\dot\gamma(t)\|^2 = \|u(t)\|^2 = \sum_{i=1}^k u_i^2(t)$ for a.e. $t \in [0,T]$. In what follows, whenever we speak about the control associated with a horizontal trajectory, we implicitly assume to consider its minimal control. This is the sub-Riemannian implementation of Remark 2.11

For every admissible curve γ, it is natural to define its *length* by the formula

$$\ell(\gamma) = \int_0^T \|\dot\gamma(t)\| dt = \int_0^T \left(\sum_{i=1}^k u_i^2(t) \right)^{1/2} dt.$$

Since the length is invariant by reparametrization, we can always assume that $\|\dot\gamma(t)\|$ is constant. The *sub-Riemannian (or Carnot-Carathéodory) distance* between two points $x, y \in M$ is

$$\mathsf{d}(x,y) \doteq \inf\{\ell(\gamma) \,|\, \gamma \text{ horizontal}, \ \gamma(0) = x, \gamma(T) = y\}.$$

It follows from the Cauchy-Schwartz inequality that, if the final time T is fixed, the minima of the length (parametrized with constant speed) coincide with the minima of the energy functional:

$$J_T(\gamma) = \frac{1}{2} \int_0^T \|\dot\gamma(t)\|^2 dt = \frac{1}{2} \int_0^T \sum_{i=1}^k u_i^2(t) dt.$$

Moreover, if γ is a minimizer with constant speed, one has the identity $\ell^2(\gamma) = 2T J_T(\gamma)$.

In particular, the problem of finding the sub-Riemannian geodesics, i.e. curves on M that minimize the distance between two points, coincides with the optimal control problem

$$\begin{aligned} &\dot x = \sum_{i=1}^k u_i X_i(x), \qquad x \in M, \\ &x(0) = x_0,\ x(T) = x_1, \qquad J_T(u) \to \min . \end{aligned}$$

Thus, a sub-Riemannian structure corresponds to an affine optimal control problem (4) where $f_0 = 0$ and the Lagrangian $L(x,u) = \frac{1}{2}\|u\|^2$ is induced by the euclidean structure on $\mathbb{U}$. Moreover the value function at time $T > 0$ is closely related with the sub-Riemannian distance as follows:

$$S_T(x,y) = \frac{1}{2T}\mathsf{d}^2(x,y), \qquad x,y \in M,$$

where we have chosen, in the definition of the value function, $M' = M$, even when the latter is not compact (see Definition 2.4). Indeed, the proof of the regularity of the value function in Appendix C can be adapted by using the fact that small sub-Riemannian balls are compact. The smoothness properties of the sub-Riemannian square distance are discussed in Section 5.2 (see also [2]).

Remark 5.5. The assumption (A1) on the control system in the sub-Riemannian case reads $\mathrm{Lie}_x\overline{\mathscr{D}} = T_xM$, for every $x \in M$. This is the classical *bracket-generating* (or *Hörmander*) condition on the distribution $\mathscr{D}$, which implies the controllability of the system, i.e. $\mathsf{d}(x,y) < \infty$ for all $x,y \in M$. Moreover one can show that d induces on M the original manifold's topology. When (M,d) is complete as a metric space, Filippov Theorem guarantees the existence of minimizers joining x to y, for all $x,y \in M$ (see [3,9]).

The maximality condition (10) of PMP reads $u_i(\lambda) = \langle \lambda, X_i(x)\rangle$, where $x = \pi(\lambda)$. Thus the maximized Hamiltonian is

$$H(\lambda) = \frac{1}{2}\sum_{i=1}^{k}\langle \lambda, X_i(x)\rangle^2, \qquad \lambda \in T^*M.$$

It is easily seen that $H : T^*M \to \mathbb{R}$ is also characterized as the dual of the norm on the distribution

$$H(\lambda) = \frac{1}{2}\|\lambda\|^2, \qquad \|\lambda\| = \sup_{\substack{v\in\mathscr{D}_x \\ \|v\|=1}} \langle \lambda, v\rangle.$$

Since, in this case, H is quadratic on fibers, we obtain immediately the following properties for the exponential map

$$\mathcal{E}_{x_0}(t, s\lambda_0) = \mathcal{E}_{x_0}(ts, \lambda_0), \qquad \lambda_0 \in T^*_{x_0}M, \quad t,s \geq 0,$$

which is tantamount to the fact that the normal geodesic associated with the covector λ_0 is the image of the ray $\{t\lambda_0, t \geq 0\} \subset T^*_{x_0}M$ through the exponential map: $\mathcal{E}_{x_0}(1, t\lambda_0) = \gamma(t)$.

Definition 5.6. Let $\gamma(t) = \pi \circ e^{t\vec{H}}(\lambda_0)$ be a strictly normal geodesic. We say that $\gamma(s)$ is *conjugate* to $\gamma(0)$ along γ if λ_0 is a critical point for $\mathcal{E}_{x_0,s}$, i.e. $D_{\lambda_0}\mathcal{E}_{x_0,s}$ is not surjective.

Remark 5.7. The sub-Riemannian maximized Hamiltonian is a quadratic function on fibers, which implies $d^2_\lambda H_x = 2H_x$, where $H_x = H|_{T^*_xM}$ and $\lambda \in T^*_xM$. In particular $d^2_\lambda H_x$ does not depend on λ and the inner product $\langle\cdot|\cdot\rangle_\lambda$ induced on the distribution $\mathscr{D}_x$ coincides with the sub-Riemannian inner product (see Section 4.2).

5.1.1 Nilpotent approximation and privileged coordinates

In this section we briefly recall the concept of nilpotent approximation. For more details we refer to [6, 7, 15, 19]. See also [25] for equiregular structures. The classical presentation that follows relies on the introduction of a set of privileged coordinates; an intrinsic construction can be found in [3].

Let M be a bracket-generating sub-Riemannian manifold. The *flag* of the distribution at a point $x \in M$ is the sequence of subspaces $\mathscr{D}^0_x \subset \mathscr{D}^1_x \subset \mathscr{D}^2_x \subset \ldots \subset T_xM$ defined by

$$\mathscr{D}^0_x \doteq \{0\}, \qquad \mathscr{D}^1_x \doteq \mathscr{D}_x, \qquad \mathscr{D}^{i+1}_x \doteq \mathscr{D}^i_x + [\mathscr{D}^i, \mathscr{D}]_x,$$

where, with a standard abuse of notation, we understand that $[\mathscr{D}^i, \mathscr{D}]_x$ is the vector space generated by the iterated Lie brackets, up to length $i+1$, of local sections of the distribution, evaluated at x. We denote by $\mathfrak{m} = \mathfrak{m}_x$ the *step of the distribution* at x, i.e. the smallest integer such that $\mathscr{D}^{\mathfrak{m}_x}_x = T_xM$. The sub-Riemannian structure is called *equiregular* if $\dim \mathscr{D}^i_x$ does not depend on $x \in M$, for every $i \geq 1$.

Let O_x be an open neighbourhood of the point $x \in M$. We say that a system of coordinates $\psi : O_x \to \mathbb{R}^n$ is *linearly adapted* to the flag if, in these coordinates, $\psi(x) = 0$ and

$$\psi_*(\mathscr{D}^i_x) = \mathbb{R}^{h_1} \oplus \ldots \oplus \mathbb{R}^{h_i}, \qquad \forall\, i = 1, \ldots, \mathfrak{m}_x,$$

where $h_i = \dim \mathscr{D}^i_x - \dim \mathscr{D}^{i-1}_x$ for $i = 1, \ldots, \mathfrak{m}_x$. Indeed $h_1 + \ldots + h_{\mathfrak{m}_x} = n$.

In these coordinates, $x = (x_1, \ldots, x_{\mathfrak{m}_x})$, where $x_i = (x_i^1, \ldots, x_i^{h_i}) \in \mathbb{R}^{h_i}$, and $T_xM = \mathbb{R}^{h_1} \oplus \ldots \oplus \mathbb{R}^{h_{\mathfrak{m}_x}}$. The space of all differential operators in $\mathbb{R}^n$ with smooth coefficients forms an associative algebra with composition of operators as multiplication. The differential operators with polynomial coefficients form a subalgebra of this algebra with generators $1, x_i^j, \partial_{x_i^j}$, where $i = 1, \ldots, m$; $j = 1, \ldots, k_i$. We define weights of generators as follows: $\nu(1) = 0$, $\nu(x_i^j) = i$, $\nu(\partial_{x_i^j}) = -i$, and the weight of monomials accordingly. Notice that a polynomial differential operator homogeneous with respect to ν (i.e. whose monomials are all of same weight) is homogeneous with respect to dilations $\delta_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ defined by $\delta_\alpha(x_1, \ldots, x_m) = (\alpha x_1, \alpha^2 x_2, \ldots, \alpha^{\mathfrak{m}_x} x_{\mathfrak{m}_x}), \alpha > 0$. In particular for a homogeneous vector field X of weight h it holds $\delta_{\alpha *} X = \alpha^{-h} X$.

Let $X \in \mathrm{Vec}(\mathbb{R}^n)$, and consider its Taylor expansion at the origin as a first order differential operator. Namely, we can write the formal expansion

$$X \approx \sum_{h=-\mathfrak{m}_x}^{\infty} X^{(h)},$$

where $X^{(h)}$ is the homogeneous part of degree h of X (notice that every monomial of a first order differential operator has weight not smaller than $-\mathfrak{m}_x$). Define the filtration of $\mathrm{Vec}(\mathbb{R}^n)$

$$\mathrm{Vec}^{(h)}(\mathbb{R}^n) = \{X \in \mathrm{Vec}(\mathbb{R}^n) : X^{(i)} = 0, \forall\, i < h\}, \qquad h \in \mathbb{Z}.$$

Definition 5.8. A system of coordinates $\psi : O_x \to \mathbb{R}^n$ is called *privileged* for the sub-Riemannian structure if they are linearly adapted and $\psi_* X_i \in \mathrm{Vec}^{(-1)}(\mathbb{R}^n)$ for every $i = 1, \ldots, k$.

The existence of privileged coordinates is proved, e.g. in [6, 15]. Notice, however, that privileged coordinates are not unique. Now we are ready to define the sub-Riemannian tangent space of M at x.

Definition 5.9. Given a set of privileged coordinates, the *nilpotent approximation at x* is the sub-Riemannian structure on $T_xM = \mathbb{R}^n$ defined by the set of vector fields $\widehat{X}_1, \ldots, \widehat{X}_k$, where $\widehat{X}_i \doteq (\psi_* X_i)^{(-1)} \in \mathrm{Vec}(\mathbb{R}^n)$.

The definition is well posed, in the sense that the structures obtained by different sets of privileged coordinates are isometric (see [15, Proposition 5.20]). Then, in what follows we omit the coordinate map in the notation above, identifying $T_xM = \mathbb{R}^n$ and a vector field with its coordinate expression in $\mathbb{R}^n$. The next proposition also justifies the name of the sub-Riemannian tangent space (see [15, Proposition 5.17]).

Proposition 5.10. *The vector fields $\widehat{X}_1, \dots \widehat{X}_k$ generate a nilpotent Lie algebra* $\mathrm{Lie}(\widehat{X}_1, \dots, \widehat{X}_k)$ *of step $\mathfrak{m}_x$. At any point $z \in \mathbb{R}^n$ they satisfy the bracket-generating assumption, namely* $\mathrm{Lie}_z(\widehat{X}_1, \dots, \widehat{X}_k) = \mathbb{R}^n$.

Remark 5.11. The sub-Riemannian distance $\widehat{\mathsf{d}}$ on the nilpotent approximation is homogeneous with respect to dilations δ_α, i.e. $\widehat{\mathsf{d}}(\delta_\alpha(x), \delta_\alpha(y)) = \alpha\, \widehat{\mathsf{d}}(x, y)$.

Definition 5.12. Let $X_1, \dots, X_k$ be a set of vector fields which defines the sub-Riemannian structure on M and fix a system of privileged coordinates at $x \in M$. The ε-approximated system at x is the sub-Riemannian structure induced by the vector fields $X_1^\varepsilon, \dots, X_k^\varepsilon$ defined by

$$X_i^\varepsilon \doteq \varepsilon \delta_{1/\varepsilon *} X_i, \qquad i = 1, \dots, k.$$

The following lemma is a consequence of the definition of ε-approximated system and privileged coordinates.

Lemma 5.13. *$X_i^\varepsilon \to \widehat{X}_i$ in the C^∞ topology of uniform convergence of all derivatives on compact sets in $\mathbb{R}^n$ when $\varepsilon \to 0$, for $i = 1, \dots, k$.*

Therefore, the nilpotent approximation $\widehat{X}$ of a vector field X at a point x is the "principal part" in the expansion when one considers the blown up coordinates near the point x, with rescaled distances.

5.1.2 Approximating trajectories

In this subsection we show, in a system of privileged coordinates $\psi : O_x \to \mathbb{R}^n$, how the normal trajectories of the ε-approximated system converge to corresponding normal trajectories of the nilpotent approximation.

Let $H^\varepsilon : T^*\mathbb{R}^n \to \mathbb{R}$ be the maximized Hamiltonian for the ε-approximated system, and $\mathcal{E}^\varepsilon : T_0^*\mathbb{R}^n \to \mathbb{R}^n$ the corresponding exponential map (starting at 0). We denote by the symbols $\widehat{H}$ and $\widehat{\mathcal{E}}$ the analogous objects for the nilpotent approximation. The ε-approximated normal trajectory $\gamma^\varepsilon(t)$ converges to the corresponding nilpotent trajectory $\widehat{\gamma}(t)$.

Proposition 5.14. *Let $\lambda_0 \in T_0^*\mathbb{R}^n$. Let $\gamma^\varepsilon : [0, T] \to \mathbb{R}^n$ and $\widehat{\gamma} : [0, T] \to \mathbb{R}^n$ be the normal geodesics associated with λ_0 for the ε-approximating system and for the nilpotent system, respectively. Let $u^\varepsilon : [0, T] \to \mathbb{R}^k$ and $\widehat{u} : [0, T] \to \mathbb{R}^k$ be the associated controls. Then there exists a neighbourhood $O_{\lambda_0} \subset T_0^*\mathbb{R}^n$ of λ_0 such that for $\varepsilon \to 0$*

(i) $\mathcal{E}^\varepsilon \to \widehat{\mathcal{E}}$ in the C^∞ topology of uniform convergence of all derivatives on O_{λ_0},

(ii) $\gamma^\varepsilon \to \widehat{\gamma}$ in the C^∞ topology of uniform convergence of all derivatives on $[0, T]$,

(iii) $u^\varepsilon \to \widehat{u}$ in the C^∞ topology of uniform convergence of all derivatives on $[0, T]$.

The proof of Proposition 5.14 is a consequence of a more general statement for the Hamiltonian flow of the approximating systems, which can be found in Appendix D.

5.2 Ample geodesics and regularity of the squared distance

In this section we discuss the existence of ample geodesics in sub-Riemannian geometry and its implications on the regularity of the sub-Riemannian distance.

We start by proving a bound on the growth vector of a normal geodesic in terms of the growth vector on the nilpotent system and the growth vector of the distribution. Let us write the control system and its nilpotent approximation at a point $x_0 \in M$ in privileged coordinates:

$$\dot{x} = f(x,u) = \sum_{i=1}^{k} u_i X_i(x), \qquad \dot{x} = \widehat{f}(x,u) = \sum_{i=1}^{k} u_i \widehat{X}_i(x), \qquad x \in \mathbb{R}^n.$$

Fix an initial covector λ_0, and let $\gamma : [0,T] \to M$ be the corresponding normal geodesic. Let $\mathscr{F}_\gamma(t)$ be the family associated with γ. Moreover, let $\widehat{\gamma} : [0,T] \to \mathbb{R}^n$ be the normal geodesic associated with the same initial covector in some set of privileged coordinates. Besides, let $\widehat{\mathscr{F}}_\gamma(t)$ be the family associated with $\widehat{\gamma}$. Analogously we define the steps $m(t)$ and $\widehat{m}(t)$ of γ and $\widehat{\gamma}$, respectively. Recall that $\mathfrak{m}_x$ is the step of the distribution $\mathscr{D}$ at x.

Lemma 5.15. *The following inequalities hold true, for $t \in [0,T]$:*

(i) $\dim \widehat{\mathscr{F}}^i_\gamma(t) \leq \dim \mathscr{F}^i_\gamma(t) \leq \dim \mathscr{D}^i_{\gamma(t)}$,

(ii) $\mathfrak{m}_{\gamma(t)} \leq m(t) \leq \widehat{m}(t)$.

Proof. Claim (ii) follows from (i). The right inequality in (i) is a direct consequence of the alternative definition of the flag of the geodesic given in Section 3.4.

The left inequality in (i) can be proved considering an ε-approximating structure for the nilpotent approximation in privileged coordinates, and applying the criterion of Section 3.2. Such a system writes, in privileged coordinates

$$\dot{x} = f^\varepsilon(x,u) = \sum_{i=1}^{k} u_i X_i^\varepsilon(x), \qquad x \in \mathbb{R}^n.$$

Let, as usual, $\gamma^\varepsilon : [0,T] \to \mathbb{R}^n$ and $u^\varepsilon : [0,T] \to \mathbb{R}^k$ be the normal geodesic and the normal control associated with the fixed covector λ_0 on the ε-approximating structures. Moreover, let $\widehat{\gamma}$, $\widehat{u}$ be the analogous objects on the nilpotent structure. The C^∞ uniform convergence guaranteed by Lemma 5.13 and Proposition 5.14, respectively, imply that, for $\varepsilon \to 0$

$$A^\varepsilon(t) \doteq \frac{\partial f^\varepsilon}{\partial x}(\gamma^\varepsilon(t), u^\varepsilon(t)) \longrightarrow \widehat{A}(t) \doteq \frac{\partial \widehat{f}}{\partial x}(\widehat{\gamma}(t), \widehat{u}(t)),$$
$$B^\varepsilon(t) \doteq \frac{\partial f^\varepsilon}{\partial u}(\gamma^\varepsilon(t), u^\varepsilon(t)) \longrightarrow \widehat{B}(t) \doteq \frac{\partial \widehat{f}}{\partial u}(\widehat{\gamma}(t), \widehat{u}(t)),$$

together with their derivatives, uniformly on $[0,T]$. The matrices $A^\varepsilon(t)$ and $B^\varepsilon(t)$ represent the linearisation of the ε-approximated system or, equivalently, of the original system in a set of ε-dependant coordinates. Besides they converge, together with their derivatives, to the same objects for the nilpotent system. The inequality is then a consequence of the criterion in Section 3.2 and the semicontinuity of the rank. ☐

Next, we provide a characterization for smooth points of the squared distance (which is the value function for the sub-Riemannian optimal control problem). Let $x_0 \in M$, and let $\Sigma_{x_0} \subset M$ be the set of points x such that there exists a unique non-conjugate and non-abnormal minimizer $\gamma : [0,1] \to M$ joining x_0 with x.

Theorem 5.16 (see [2])**.** *Let $x_0 \in M$ and set $\mathfrak{f} \doteq \frac{1}{2}\mathsf{d}^2(x_0,\cdot)$. Then Σ_{x_0} is open, dense and $\mathfrak{f}$ is smooth on Σ_{x_0}. Moreover if $x \in \Sigma_{x_0}$ then $d_x\mathfrak{f} = \lambda(1)$, where $\lambda(t)$ is the normal lift of $\gamma(t)$.*

This result, together with Lemma 5.15, imply that the set of covectors such that the associated geodesic is ample is open and dense in the fiber.

Lemma 5.17. *The set $A_{x_0} \subseteq T^*_{x_0}M$ of covectors such that the corresponding normal geodesic is ample is a nonempty, Zariski open set, therefore dense.*

Proof. The fact that A_{x_0} is an open Zariski subset is a consequence of Remark 6.16, for any polynomial Hamiltonian, since in this case the non ampleness is an algebraic condition on the fibre. We only need to prove that A_{x_0} is nonempty. To this end, by Lemma 5.15, it is sufficient to prove that there is at least one ample geodesic on the nilpotent approximation at x_0. The nilpotent approximation structure is analytic, hence a geodesic is ample if and only if it is strictly normal (more precisely strictly ⇔ strongly ⇔ ample, see Proposition 3.5). The existence of at least one strictly normal geodesic then follows by Theorem 5.16. □

5.3 Reparametrization and homogeneity of the curvature operator

We already explained that a geodesic is not ample on a proper Zariski closed subset of the fibre. This set includes covectors associated to abnormal geodesics, since $\mathscr{D}_x^\perp \subset T^*_xM \setminus A_x$. On the other hand, for $\lambda \in A_x$, the curvature $\mathcal{R}_\lambda$ is well defined. Observe that A_x is invariant by rescaling, i.e. if $\lambda \in A_x$, then for $\alpha \neq 0$, also $\alpha\lambda \in A_x$. Therefore, we have the following:

Proposition 5.18. *The operators $\mathcal{I}_\lambda$ and $\mathcal{R}_\lambda$ are homogeneous of degree* 0 *and* 2 *with respect to λ, respectively. Namely, for $\lambda \in A_x$ and $\alpha > 0$*

$$\mathcal{I}_{\alpha\lambda} = \mathcal{I}_\lambda, \qquad \mathcal{R}_{\alpha\lambda} = \alpha^2 \mathcal{R}_\lambda. \tag{33}$$

Proof. Let c_t^λ be the geodesic cost associated with the covector $\lambda \in T^*_xM$. By homogeneity of the sub-Riemannian Hamiltonian, for $\alpha > 0$ we have

$$c_t^{\alpha\lambda} = \alpha\, c_{\alpha t}^\lambda.$$

In particular, this implies $d_x^2 \dot{c}_t^{\alpha\lambda} = \alpha^2 d_x^2 \dot{c}_{\alpha t}^\lambda$. The same relation is true for the restrictions to the distribution $\mathscr{D}_x$, therefore $\mathcal{Q}_{\alpha\lambda}(t) = \alpha^2 \mathcal{Q}_\lambda(\alpha t)$ as symmetric operators on $\mathscr{D}_x$. Applying Theorem A to both families one obtains

$$\frac{1}{t^2}\mathcal{I}_{\alpha\lambda} + \frac{1}{3}\mathcal{R}_{\alpha\lambda} + O(t) = \alpha^2 \left(\frac{1}{\alpha^2 t^2}\mathcal{I}_\lambda + \frac{1}{3}\mathcal{R}_\lambda + O(\alpha t) \right),$$

which, in particular, implies Eq. (33). □

Notice that the same proof applies also to a general affine optimal control system, such that the Hamiltonian (or, equivalently, the Lagrangian) is homogeneous of degree two.

5.4 Asymptotics of the sub-Laplacian of the geodesic cost

In this section we discuss the asymptotic behaviour of the sub-Laplacian of the sub-Riemannian geodesic cost. On a Riemannian manifold, the Laplace-Beltrami operator is defined as the divergence of the gradient. This definition can be easily generalized to the sub-Riemannian setting. We will denote by $\langle\cdot|\cdot\rangle$ the inner product defined on the distribution.

Definition 5.19. Let $f \in C^\infty(M)$. The *horizontal gradient* of f is the unique horizontal vector field ∇f such that

$$\langle \nabla f | X \rangle = X(f), \qquad \forall X \in \overline{\mathscr{D}}.$$

For $x \in M$, the restriction of the sub-Riemannian Hamiltonian to the fiber $H_x : T_x^*M \to \mathbb{R}$ is a quadratic form. Then, as a consequence of the formula $\langle d_\lambda H_x | X \rangle = \langle \lambda, X \rangle$, we obtain

$$\nabla f = \sum_{i=1}^{k} X_i(f) X_i. \tag{34}$$

We want to stress that Eq. (34) is true in full generality, also when $\dim \mathscr{D}_x$ is not constant or the vectors $X_1, \dots, X_k$ are not independent.

Definition 5.20. Let $\mu \in \Omega^n(M)$ be a volume form, and $X \in \mathrm{Vec}(M)$. The *μ-divergence* of X is the smooth function $\mathrm{div}_\mu(X)$ defined by

$$\mathcal{L}_X \mu \doteq \mathrm{div}_\mu(X)\mu,$$

where, we recall, $\mathcal{L}_X$ is the Lie derivative in the direction of X.

Notice that the definition of divergence does not depend on the orientation of M, namely the sign of μ. The divergence measures the rate at which the volume of a region changes under the integral flow of a field. Indeed, for any compact $\Omega \subset M$ and t sufficiently small, let $e^{tX} : \Omega \to M$ be the flow of $X \in \mathrm{Vec}(M)$, then

$$\frac{d}{dt}\Big|_{t=0} \int_{e^{tX}(\Omega)} \mu = - \int_\Omega \mathrm{div}_\mu(X)\mu.$$

The next proposition is an easy consequence of the definition of μ-divergence and is sometimes employed as an alternative definition of the latter.

Proposition 5.21. *Let $C_0^\infty(M)$ be the space of smooth functions with compact support. For any $f \in C_0^\infty(M)$ and $X \in \mathrm{Vec}(M)$*

$$\int_M f \mathrm{div}_\mu(X)\mu = - \int_M X(f)\mu.$$

With a divergence and a gradient at our disposal, we are ready to define the sub-Laplacian associated with the volume form μ.

Definition 5.22. Let $\mu \in \Omega^n(M)$, $f \in C^\infty(M)$. The *sub-Laplacian* associated with μ is the second order differential operator

$$\Delta_\mu f \doteq \mathrm{div}_\mu (\nabla f),$$

On a Riemannian manifold, when μ is the Riemannian volume, this definition reduces to the Laplace-Beltrami operator. As a consequence of Eq. (34) and the Leibniz rule for the divergence $\mathrm{div}_\mu(fX) = X(f) + f\,\mathrm{div}_\mu(X)$, we can write the sub-Laplacian in terms of the fields $X_1, \dots, X_k$:

$$\mathrm{div}_\mu (\nabla f) = \sum_{i=1}^{k} \mathrm{div}_\mu (X_i(f) X_i) = \sum_{i=1}^{k} X_i(X_i(f)) + \mathrm{div}_\mu(X_i) X_i(f).$$

Then

$$\Delta_\mu = \sum_{i=1}^{k} X_i^2 + \mathrm{div}_\mu(X_i) X_i. \tag{35}$$

Remark 5.23. If we apply Proposition 5.21 to the horizontal gradient ∇g, we obtain

$$\int_M f \Delta g \mu = - \int_M \langle \nabla f | \nabla g \rangle \mu, \qquad \forall f, g \in C_0^\infty(M).$$

Then Δ_μ is symmetric and negative on $C_0^\infty(M)$. It can be proved that it is also essentially self-adjoint (see [32]). Hence it admits a unique self-adjoint extension to $L^2(M, \mu)$.

Observe that the principal symbol of Δ_μ, which is a function on T^*M, does not depend on the choice of μ, and is proportional to the sub-Riemannian Hamiltonian, namely $2H : T^*M \to \mathbb{R}$. The sub-Laplacian depends on the choice of the volume μ according to the following lemma.

Lemma 5.24. *Let $\mu, \mu' \in \Omega^n(M)$ be two volume forms such that $\mu' = e^a\mu$ for some $a \in C^\infty(M)$. Then*

$$\Delta_{\mu'} f = \Delta_\mu f + \langle \nabla a | \nabla f \rangle.$$

Proof. It follows from the Leibniz rule $\mathcal{L}_X(a\mu) = X(a)\mu + a\mathcal{L}_X\mu = (X(\log a) + \mathrm{div}_\mu(X))a\mu$ for every $a \in C^\infty(M)$. □

The sub-Laplacian, computed at critical points, does not depend on the choice of the volume.

Lemma 5.25. *Let $f \in C^\infty(M)$, and let $x \in M$ be a critical point of f. Then, for any choice of the volume μ,*

$$\Delta_\mu f|_x = \sum_{i=1}^k X_i^2(f)|_x.$$

Proof. The proof follows from Eq. (35), and the fact that $X_i(f)|_x = 0$. □

From now on, when computing the sub-Laplacian of a function at a critical point, we employ the notation $\Delta_\mu f|_x = \Delta f|_x$, since it does not depend on the volume.

Lemma 5.26. *Let $f \in C^\infty(M)$, and let $x \in M$ be a critical point of f. Then $\Delta f|_x = \mathrm{tr}\, d_x^2 f|_{\mathscr{D}_x}$.*

Proof. Recall that if x is a critical point of f, then the second differential $d_x^2 f$ is the quadratic form associated with the symmetric bilinear form

$$d_x^2 f : T_xM \times T_xM \to \mathbb{R}, \qquad (X, Y) \mapsto X(Y(f))|_x.$$

The restriction of $d_x^2 f$ to the distribution can be associated, via the inner product, with a symmetric operator defined on $\mathscr{D}_x$, whose trace is computed in terms of $X_1, \dots, X_k$ as follows

$$\mathrm{tr}\, d_x^2 f|_{\mathscr{D}_x} = \sum_{i=1}^k X_i^2(f)|_x, \tag{36}$$

We stress that Eq. (36) holds true for any set of generators, not necessarily linearly independent, of the sub-Riemannian structure $X_1, \dots, X_k$ such that $H(\lambda) = \frac{1}{2}\sum_{i=1}^k \langle \lambda, X_i \rangle^2$. The statement now is a direct consequence of Lemma 5.25. □

Remember that the derivative of the geodesic cost function $\dot{c}_t$ has a critical point at $x_0 = \gamma(0)$. As a direct consequence of Theorem A, B, Lemma 5.26 and the fact that, in the sub-Riemannian case, the Hamiltonian inner product is the sub-Riemannian one (see Remark 4.7), we get the following asymptotic expansion:

Theorem 5.27. *Let c_t be the geodesic cost associated with a geodesic γ such that $\gamma(0) = x_0$. Then*

$$\Delta \dot{c}_t|_{x_0} = \frac{\mathrm{tr}\, \mathcal{I}_\lambda}{t^2} + \frac{1}{3}\mathrm{Ric}(\lambda) + O(t),$$

where $\mathrm{Ric}(\lambda) = \mathrm{tr}\, \mathcal{R}_\lambda$.

The next result is an explicit expression for the asymptotic of the sub-Laplacian of the geodesic cost computed at the initial point x_0 of the geodesic γ. In the sub-Riemannian case, the geodesic cost is essentially the squared distance from the geodesic, i.e. the function

$$\mathfrak{f}_t(\cdot) \doteq -tc_t(\cdot) = \frac{1}{2}\mathsf{d}^2(\,\cdot\,,\gamma(t)), \qquad t \in (0,1].$$

For this reason, we may state the theorem equivalently in terms of $\mathfrak{f}_t$ or the geodesic cost c_t. Remember also that, since x_0 is not a critical point of $\mathfrak{f}_t$, its sub-Laplacian depends on the choice of the volume form μ.

Theorem C. *Let γ be an equiregular geodesic with initial covector $\lambda \in T^*_{x_0}M$. Assume also that $\dim \mathscr{D}$ is constant in a neighbourhood of x_0. Then there exists a smooth n-form ω defined along γ, such that for any volume form μ on M, $\mu_{\gamma(t)} = e^{g(t)}\omega_{\gamma(t)}$, we have*

$$\Delta_\mu \mathfrak{f}_t|_{x_0} = \operatorname{tr}\mathcal{I}_\lambda - \dot{g}(0)t - \frac{1}{3}\operatorname{Ric}(\lambda)t^2 + O(t^3), \tag{37}$$

As a consequence, for any choice of the volume form μ

$$\lim_{t\to 0} \Delta_\mu \mathfrak{f}_t\big|_{x_0} = \operatorname{tr}\mathcal{I}_\lambda,$$

$$\left.\frac{d^2}{dt^2}\right|_{t=0} \Delta_\mu \mathfrak{f}_t\big|_{x_0} = -\frac{2}{3}\operatorname{Ric}(\lambda).$$

The proof Theorem C is postponed to Section 8.

Observe that only the first order term in t of Eq. (37) depends on the choice of the volume. The explicit expression of ω is not relevant here, and requires the premature introduction of some technical tools which we deemed not necessary at this point. We only anticipate that ω, which indeed depends on γ, is related with a generalization of the parallel transport of the volume form along the geodesic. On a Riemannian manifold, ω does not depend on γ and, up to a sign, is equal to the Riemannian volume form. Therefore the first order term in Eq. (37) vanishes. This is not true, in general, for sub-Riemannian manifolds.

5.5 Equiregular distributions

In this section we focus on equiregular sub-Riemannian structures, endowed with a smooth, intrinsic volume form, called Popp's volume. Then we introduce a special class of equiregular distributions, that we call *slow growth*. In this case, we define a family of smooth operators in terms of which the asymptotic expansion of Theorem C (and in particular its linear term) can be expressed explicitly.

Recall that a bracket generating sub-Riemannian manifold M is *equiregular* if $\dim \mathscr{D}^i_x$ does not depend on $x \in M$, for every $i \geq 0$, where $\mathscr{D}^0_x \subset \mathscr{D}^1_x \subset \mathscr{D}^2_x \subset \ldots \subset T_xM$ is the flag of the distribution at a point $x \in M$ (see Section 5).

5.5.1 Popp's volume

In this section we provide the definition of Popp's volume for an equiregular sub-Riemannian structure. Our presentation follows closely the one of [12, 26]. The definition rests on the following lemmas, whose proof is not repeated here.

Lemma 5.28. *Let E be an inner product space, and let $\pi : E \to V$ be a surjective linear map. Then π induces an inner product on V such that the norm of $v \in V$ is*

$$\|v\|_V = \min\{\|e\|_E \ \ s.t.\ \ \pi(e) = v\}.$$

Lemma 5.29. *Let E be a vector space of dimension n with a flag of linear subspaces $\{0\} = F^0 \subset F^1 \subset F^2 \subset \ldots \subset F^m = E$. Let $\mathrm{gr}(F) \doteq F^1 \oplus F^2/F^1 \oplus \ldots \oplus F^m/F^{m-1}$ be the associated graded vector space. Then there is a canonical isomorphism $\theta : \wedge^n E \to \wedge^n \mathrm{gr}(F)$.*

The idea behind Popp's volume is to define an inner product on each $\mathscr{D}^i_x/\mathscr{D}^{i-1}_x$ which, in turn, induces an inner product on the orthogonal direct sum

$$\mathrm{gr}_x(\mathscr{D}) = \mathscr{D}_x \oplus \mathscr{D}^2_x/\mathscr{D}_x \oplus \ldots \oplus \mathscr{D}^{\mathfrak{m}}_x/\mathscr{D}^{\mathfrak{m}-1}_x.$$

The latter has a natural volume form, which is the canonical volume of an inner product space obtained by wedging the elements an orthonormal dual basis. Then, we employ Lemma 5.29 to define an element of $(\wedge^n T_x M)^* \simeq \wedge^n T^*_x M$, which is Popp's volume form computed at x.

Fix $x \in M$. Then, let $v, w \in \mathscr{D}_x$, and let V, W be any horizontal extensions of v, w. Namely, $V, W \in \overline{\mathscr{D}}$ and $V(x) = v$, $W(x) = w$. The linear map $\pi : \mathscr{D}_x \otimes \mathscr{D}_x \to \mathscr{D}^2_x/\mathscr{D}_x$

$$\pi(v \otimes w) := [V, W]_x \mod \mathscr{D}_x, \tag{38}$$

is well defined, and does not depend on the choice the horizontal extensions. Similarly, let $1 \leq i \leq \mathfrak{m}$. The linear maps $\pi_i : \otimes^i \mathscr{D}_x \to \mathscr{D}^i_x/\mathscr{D}^{i-1}_x$

$$\pi_i(v_1 \otimes \cdots \otimes v_i) = [V_1, [V_2, \ldots, [V_{i-1}, V_i]]]_x \mod \mathscr{D}^{i-1}_x, \tag{39}$$

are well defined and do not depend on the choice of the horizontal extensions $V_1, \ldots, V_i$ of $v_1, \ldots, v_i$.

By the bracket-generating condition, the maps π_i are surjective and, by Lemma 5.28, they induce an inner product space structure on $\mathscr{D}^i_x/\mathscr{D}^{i-1}_x$. Therefore, the nilpotentization of the distribution at x, namely $\mathrm{gr}_x(\mathscr{D})$, is an inner product space, as the orthogonal direct sum of a finite number of inner product spaces. As such, it is endowed with a canonical volume (defined up to a sign) $\eta_x \in \wedge^n \mathrm{gr}_x(\mathscr{D})^*$, which is the volume form obtained by wedging the elements of an orthonormal dual basis.

Finally, Popp's volume (computed at the point x) is obtained by transporting the volume of $\mathrm{gr}_x(\mathscr{D})$ to $T_x M$ through the map $\theta_x : \wedge^n T_x M \to \wedge^n \mathrm{gr}_x(\mathscr{D})$ defined in Lemma 5.29. Namely

$$\mu_x = \eta_x \circ \theta_x, \tag{40}$$

where we employ the canonical identification $(\wedge^n T_x M)^* \simeq \wedge^n T^*_x M$. Eq. (40) is defined only in the domain of the chosen local frame. If M is orientable, with a standard argument, these n-forms can be glued together to obtain Popp's volume $\mu \in \Omega^n(M)$. Notice that Popp's volume is smooth by construction.

Remark 5.30. From Eq. (38) and (39) it follows that, for any $i \geq 0$ and $V \in \mathscr{D}_x$ the linear maps $\mathrm{ad}^i_x V : \mathscr{D}_x \to \mathscr{D}^{i+1}_x/\mathscr{D}^i_x$ given by

$$\mathrm{ad}^i_x V(W) \doteq \underbrace{[V, [V, \ldots, [V}_{i \text{ times}}, W]]]_x \mod \mathscr{D}^i_x, \qquad W \in \mathscr{D}_x,$$

are well-defined.

5.5.2 Slow growth distributions

Now we are ready to introduce the following class of equiregular distributions.

Definition 5.31. An equiregular distribution is *slow growth* at $x \in M$ if there exists a vector $\mathsf{T} \in \mathscr{D}_x$ such that the linear map $\mathrm{ad}^i_x \mathsf{T}$ is surjective for all $i \geq 0$.

This condition is actually generic in T, as stated by the following proposition.

Proposition 5.32. *Let $\mathscr{D}$ be a slow growth distribution at x. Then, for T in a nonempty open Zariski subset of $\mathscr{D}_x$, all the linear maps $\mathrm{ad}_x^i\mathsf{T}$ are surjective.*

Proof. Let X_i be an orthonormal basis for $\mathscr{D}_x$ and write $\mathsf{T} = \sum_{j=1}^k \alpha_j X_j$, where $k = \dim \mathscr{D}_x$ and the α_j are constant. The definition of slow growth is a maximal rank condition on the operators $\mathrm{ad}_x^i\mathsf{T} = (\sum_{j=1}^k \alpha_j \mathrm{ad}_x X_j)^i$, which is satisfied by at least one element of $\mathscr{D}_x$. Then, the result follows from the fact that $\mathrm{ad}_x^i\mathsf{T}$ depends polynomially on the α_j. □

We say that a distribution $\mathscr{D}$ is *slow growth* if it is slow growth at every point $x \in M$. Familiar sub-Riemannian structures such as contact, quasi-contact, fat, Engel, Goursat-Darboux distributions (see [16]) are examples of slow growth distributions.

Now, for any fixed equiregular, ample (of step m) geodesic $\gamma : [0, T] \to M$, with flag $0 = \mathscr{F}^0_{\gamma(t)} \subset \mathscr{F}^1_{\gamma(t)} \subset \ldots \subset \mathscr{F}^m_{\gamma(t)} = T_{\gamma(t)}M$ recall the smooth families of operators

$$\mathcal{L}^i_{\mathsf{T}} : \mathscr{F}_{\gamma(t)} \to \mathscr{F}^{i+1}_{\gamma(t)}/\mathscr{F}^i_{\gamma(t)}, \qquad i = 0, \ldots, m-1,$$

defined for all $t \in [0, T]$ in terms of an admissible extension T of $\dot\gamma$ (see Remark 3.18). If the distribution is slow growth, we have the identities $\mathcal{L}^i_{\mathsf{T}} = \mathrm{ad}^i_{\gamma(t)}\mathsf{T}$ which, in particular, say that $\mathcal{L}^i_{\mathsf{T}}$ depend only on the value of T at $\gamma(t)$. Moreover, the following growth condition is satisfied

$$\dim \mathscr{F}^i_\gamma = \dim \mathscr{D}^i, \qquad \forall\, i \geq 0. \tag{41}$$

As a consequence of Proposition 5.32 it follows that, for a nonempty Zariski open set of initial covectors, the corresponding geodesic is ample (of step $m = \mathfrak{m}$, the step of the distribution), equiregular and satisfies the growth condition of Eq. (41).

Next, recall that given V, W inner product spaces, any surjective linear map $L : V \to W$ descends to an isomorphism $L : V/\ker L \to W$. Then, thanks to the inner product structure, we can consider the map $L^* \circ L : V/\ker L \to V/\ker L$ obtained by composing L with its adjoint L^*, which is a symmetric invertible operator. Applying this construction to our setting, we define the smooth families of symmetric operators

$$M_i(t) := (\mathcal{L}^{i-1}_{\mathsf{T}})^* \circ \mathcal{L}^{i-1}_{\mathsf{T}} : \mathscr{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}} \to \mathscr{D}_{\gamma(t)}/\ker \mathcal{L}^{i-1}_{\mathsf{T}}, \qquad i = 1, \ldots, m. \tag{42}$$

We are now ready to specify Theorem C for any ample, equiregular geodesic satisfying the growth condition of Eq. (41). First, let us discuss the zeroth order term of the expansion. Recall that the Hausdorff dimension of an equiregular sub-Riemannian manifold is computed by Mitchell's formula (see [15, 25]), namely

$$Q = \sum_{i=1}^m i(\dim \mathscr{D}^i - \dim \mathscr{D}^{i-1}).$$

Thus, for a slow growth distribution and a geodesic γ with initial covector $\lambda \in T^*_{x_0}M$ satisfying the growth condition of Eq. (41), we have the following identity (see also Remark 4.10)

$$\begin{aligned}\operatorname{tr} \mathcal{I}_\lambda &= \sum_{i=1}^m (2i-1)(\dim \mathscr{F}^i_\gamma - \dim \mathscr{F}^{i-1}_\gamma) = \\ &= \sum_{i=1}^m (2i-1)(\dim \mathscr{D}^i - \dim \mathscr{D}^{i-1}) = 2Q - n.\end{aligned}$$

This formula gives the zeroth order term of the following theorem.

Theorem 5.33. *Let M be a sub-Riemannian manifold with a slow growth distribution $\mathscr{D}$. Let γ be an ample, equiregular geodesic with initial covector $\lambda \in T^*_{x_0}M$ satisfying the growth condition* (41)*. Then*

$$\Delta_\mu \mathfrak{f}_t|_{x_0} = (2Q - n) - \frac{1}{2}\sum_{i=1}^{m} \operatorname{tr}\left(M_i(0)^{-1}\dot{M}_i(0)\right) t - \frac{1}{3}\operatorname{Ric}(\lambda)t^2 + O(t^3). \tag{43}$$

where the smooth families of operators $M_i(t)$ are defined by Eq. (42)*.*

Remark 5.34. Equivalently we can write Eq. (43) in the following form

$$\Delta_\mu \mathfrak{f}_t|_{x_0} = (2Q - n) - \frac{1}{2}\left(\frac{d}{ds}\bigg|_{s=0}\sum_{i=1}^{m} \log\det M_i(s)\right) t - \frac{1}{3}\operatorname{Ric}(\lambda)t^2 + O(t^3).$$

The proof of Theorem 5.33 is postponed to the end of Section 8. We end this section with an example.

Example 5.35 (Riemannian structures)**.** In a Riemannian structure (see Section 4.4.2), any nontrivial geodesic has the same flag $\mathscr{F}_{\gamma(t)} = \mathscr{D}_{\gamma(t)} = T_{\gamma(t)}M$. In particular, it is a trivial example of slow growth distribution. Notice that Popp's volume reduces to the usual Riemannian volume form. Since every geodesic is ample with step $m = 1$, there is only one family of operators associated with $\gamma(t)$, namely the constant operator $M_1(t) = \mathbb{I}|_{T_{\gamma(t)}M}$. Thus, in this case, the linear term of Theorem 5.33 vanishes, and we obtain

$$\Delta \mathfrak{f}_t|_{x_0} = n - \frac{1}{3}\operatorname{Ric}(\lambda)t^2 + O(t^3),$$

where $\operatorname{Ric}(\lambda)$ is the classical Ricci curvature in the direction of the geodesic.

In Section 5.7 we compute explicitly the asymptotic expansion of Theorem 5.33 in the case of the Heisenberg group, endowed with its canonical volume. A more general class of slow growth sub-Riemannian distributions, in which the operators $M_i(t)$ are not trivial and can be computed explicitly, namely contact structures, will appear in a forthcoming paper [4].

5.6 Geodesic dimension and sub-Riemannian homotheties

In this section, M is a complete, connected, orientable sub-Riemannian manifold, endowed with a smooth volume form μ. With a slight abuse of notation, we denote by the same symbol the induced measure on M. We are interested in sub-Riemannian homotheties, namely contractions along geodesics. To this end, let us fix $x_0 \in M$, which will be the center of the homothety. Recall that Σ_{x_0} is the set of points $x \in M$ such that there exists a unique non-conjugate and non-abnormal minimizer $\gamma : [0,1] \to M$ that joins x_0 with x. Recall also that, by Theorem 5.16, $\Sigma_{x_0} \subset M$ is the open and dense set where the function $\mathfrak{f} = \frac{1}{2}\mathsf{d}^2(x_0, \cdot)$ is smooth.

Definition 5.36. For any $x \in \Sigma_{x_0}$ and $t \in [0,1]$, the *sub-Riemannian geodesic homothety of center x_0 at time t* is the map $\phi_t : \Sigma_{x_0} \to M$ that associates x with the point at time t of the unique geodesic connecting x_0 with x.

As a consequence of Theorem 5.16 and the smooth dependence on initial data, it is easy to prove that $(t, x) \mapsto \phi_t(x)$ is smooth on $[0,1] \times \Sigma_{x_0}$, and is given by the explicit formula

$$\phi_t(x) = \pi \circ e^{(t-1)\vec{H}}(d_x\mathfrak{f}). \tag{44}$$

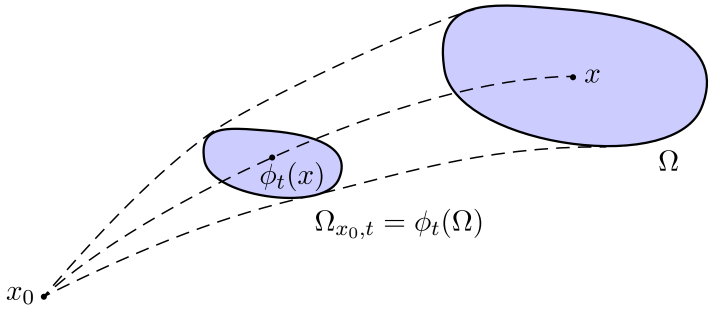


Figure 4: Sub-Riemannian homothety of the set Ω with center x_0.

Let now $\Omega \subset \Sigma_{x_0}$ be a bounded, measurable set, with $0 < \mu(\Omega) < +\infty$, and let $\Omega_{x_0,t} \doteq \phi_t(\Omega)$. The map $t \mapsto \mu(\Omega_{x_0,t})$ is smooth on $[0,1]$. As shown in Fig. 4, the homothety shrinks Ω to the center x_0. Indeed $\Omega_{x_0,0} = \{x_0\}$, and $\mu(\Omega_{x_0,t}) \to 0$ for $t \to 0$. For a Riemannian structure, a standard computation in terms of Jacobi fields shows that

$$\mu(\Omega_{x_0,t}) \sim t^{\dim M}, \qquad \text{for } t \to 0, \tag{45}$$

where we write $f(t) \sim g(t)$ if there exists $C \neq 0$ such that $f(t) = g(t)(C + o(1))$.

In the sub-Riemannian case, we have a similar power-law behaviour, but the exponent is a different dimensional invariant, which we call *geodesic dimension*. The main result of this section is a formula for the geodesic dimension, in terms of the growth vector of the geodesic.

Definition 5.37. Let $\lambda \in T^*_{x_0}M$. Assume that the corresponding geodesic $\gamma : [0,1] \to M$ is ample (at $t = 0$) of step m, with growth vector $\mathcal{G}_\gamma = \{k_1, k_2, \dots, k_m\}$ (at $t = 0$). Then we define

$$\mathcal{N}_\lambda \doteq \sum_{i=1}^{m} (2i-1)(k_i - k_{i-1}), \tag{46}$$

and $\mathcal{N}_\lambda \doteq +\infty$ if the geodesic is not ample.

Observe that Eq. (46) closely resembles the formula for Hausdorff dimension of an equiregular sub-Riemannian manifold. In the latter, each direction has a weight according to the flag of the distribution, while in Eq. (46), the weights depend on the flag of the geodesic.

Remark 5.38. Assume that λ is associated with an equiregular geodesic γ. Then, by Remark 4.10 and Eq. (46) it follows that

$$\mathcal{N}_\lambda = \operatorname{tr} \mathcal{I}_\lambda.$$

Moreover, as a consequence of Theorem C, under these assumption $\mathcal{N}_\lambda$ can be recovered from the sub-Laplacian of $\mathfrak{f}_t$ by the following formula

$$\mathcal{N}_\lambda = \lim_{t \to 0} \Delta_\mu \mathfrak{f}_t \big|_{x_0}.$$

Proposition 5.39. *The function $\lambda \mapsto \mathcal{N}_\lambda$ is constant a.e. on $T^*_{x_0}M$, assuming its minimum value. Therefore, we define the* geodesic dimension *at x_0 as*

$$\mathcal{N}_{x_0} \doteq \min\{\mathcal{N}_\lambda \,|\, \lambda \in T^*_{x_0}M\} < +\infty.$$

Remark 5.40. For every $x_0 \in M$ we have the inequality $\mathcal{N}_{x_0} \geq \dim M$ and the equality holds if and only if the structure is Riemannian at x_0. Notice that, if the distribution is equiregular at x_0, it follows from Lemma 5.15 and Mitchell's formula for Hausdorff dimension (see [25]) that $\mathcal{N}_{x_0} > \dim_{\mathcal{H}} M$. For genuine sub-Riemannian structures then, the geodesic dimension is a new invariant, related with the structure of the distribution along geodesics.

The geodesic dimension is the exponent of the sub-Riemannian analogue of Eq. (45).

Theorem D. *Let μ be a smooth volume. For any bounded, measurable set $\Omega \subset \Sigma_{x_0}$, with $0 < \mu(\Omega) < +\infty$ we have*

$$\mu(\Omega_{x_0,t}) \sim t^{\mathcal{N}_{x_0}}, \qquad \textit{for } t \to 0.$$

Observe also that homotheties with different center may have different asymptotic exponents. This can happen, for example, in non-equiregular sub-Riemannian structures.

The proof of Proposition 5.39 and Theorem D is postponed to the end of Section 6.

Example 5.41 (Geodesic dimension in contact structures)**.** Let $(M, \mathscr{D}, \langle\cdot|\cdot\rangle)$ be a contact sub-Riemannian structure. In this case, for any $x_0 \in M$, $\dim M = 2\ell + 1$ and $\dim \mathscr{D}_{x_0} = 2\ell$. Any non-trivial geodesic γ is ample with the same growth vector $\mathcal{G}_\gamma = \{2\ell, 2\ell + 1\}$. Therefore, by Eq. (46), $\mathcal{N}_{x_0} = 2\ell + 3$ (notice that it does not depend on x_0). Theorem D is an asymptotic generalization of the results obtained in [20], where the exponent $2\ell+3$ appears in the context of measure contraction property in the Heisenberg group. For a more recent overview on measure contraction property in Carnot groups, see [31].

5.7 Heisenberg group

Before entering into details of the proofs, we repeat the construction introduced in the previous sections for one of the simplest sub-Riemannian structures: the Heisenberg group. We provide an explicit expression for the geodesic cost function and, applying Definition 4.8, we obtain a formula for the operators $\mathcal{I}_\lambda$ and $\mathcal{R}_\lambda$. In particular, we recover by a direct computation the results of Theorems A, B and C.

The Heisenberg group $\mathbb{H}$ is the equiregular sub-Riemannian structure on $\mathbb{R}^3$ defined by the global (orthonormal) frame

$$X = \partial_x - \frac{y}{2}\partial_z, \qquad Y = \partial_y + \frac{x}{2}\partial_z. \tag{47}$$

Notice that the distribution is bracket-generating, for $Z := [X, Y] = \partial_z$. Let us introduce the linear on fibers functions $h_x, h_y, h_z : T^*\mathbb{R}^3 \to \mathbb{R}$

$$h_x := p_x - \frac{y}{2}p_z, \qquad h_y := p_y + \frac{x}{2}p_z, \qquad h_z := p_z,$$

where (x, y, z, p_x, p_y, p_z) are canonical coordinates on $T^*\mathbb{R}^3$ induced by coordinates (x, y, z) on $\mathbb{R}^3$. Notice that h_x, h_y, h_z are the linear on fibers functions associated with the fields X, Y, Z, respectively (i.e. $h_x(\lambda) = \langle\lambda, X\rangle$, and analogously for h_y, h_z).

The sub-Riemannian Hamiltonian is $H = \frac{1}{2}(h_x^2 + h_y^2)$ and the coordinates (x, y, z, h_x, h_y, h_z) define a global chart for T^*M. It is useful to introduce the identification $\mathbb{R}^3 = \mathbb{C}\times\mathbb{R}$, by defining the complex variable $w := x + iy$ and the complex "momentum" $h_w := h_x + ih_y$. Let $q = (w, z)$ and $q' = (w', z')$ be two points in $\mathbb{H}$. The Heisenberg group law, in complex coordinates, is given by

$$q \cdot q' = \left(w + w', z + z' - \frac{1}{2}\Im\left(w\overline{w'}\right)\right). \tag{48}$$

Observe that the frame (47) is left-invariant for the group action defined by Eq. (48). Notice also that h_z is constant along any geodesic due to the identity $[X, Z] = [Y, Z] = 0$.

The geodesic $\gamma(t) = (w(t), z(t))$ starting from $(w_0, z_0) \in \mathbb{H}$ and corresponding to the initial covector $(h_{w,0}, h_z)$, with $h_z \neq 0$ is given by

$$w(t) = w_0 + \frac{h_{w,0}}{ih_z}\left(e^{ih_z t} - 1\right),$$

$$z(t) = z_0 + \frac{1}{2}\int_0^t \Im(\overline{w}dw).$$

In the following, we assume that the geodesic is parametrized by arc length, i.e. $|h_{w,0}|^2 = 1$. We fix $h_{w,0} = ie^{i\phi}$, i.e. ϕ parametrizes the (unit) velocity of the geodesic $\dot{\gamma}(0) = -\sin\phi X + \cos\phi Y$. Finally, the geodesics corresponding to covectors with $h_z = 0$ are straight lines

$$w(t) = w_0 + h_{w,0}t,$$
$$z(t) = z_0 + \frac{1}{2}\Im(h_{w,0}\overline{w_0})t.$$

In the following, we employ both real (x, y, z, h_x, h_y, h_z) and complex (w, z, h_w, h_z) coordinates when convenient.

5.7.1 Distance in the Heisenberg group

Let $\mathsf{d}_0 = \mathsf{d}(0, \cdot) : \mathbb{H} \to \mathbb{R}$ be the sub-Riemannian distance from the origin and introduce cylindrical coordinates (r, φ, z) on $\mathbb{H}$ defined by $x = r\cos\varphi$, $y = r\sin\varphi$. In order to write an explicit formula for d recall that

(i) $\mathsf{d}_0^2(r, \varphi, z)$ does not depend on φ.

(ii) $\mathsf{d}_0^2(\alpha r, \varphi, \alpha^2 z) = \alpha^2 \mathsf{d}_0^2(r, \varphi, z)$, where $\alpha > 0$.

Then, for $r \neq 0$, one has

$$\mathsf{d}_0^2(r, \varphi, z) = r^2 \mathsf{d}_0^2\left(1, 0, \frac{z}{r^2}\right). \tag{49}$$

It is then sufficient to compute the squared distance of the point $q = (1, 0, \xi)$ from the origin.

Consider the minimizing geodesic joining the origin with the point $(1, 0, \xi)$. Its projection on the xy-plane is an arc of circle with radius ρ, connecting the origin with the point $(1, 0)$. In what follows we refer to notation of Fig. 5.

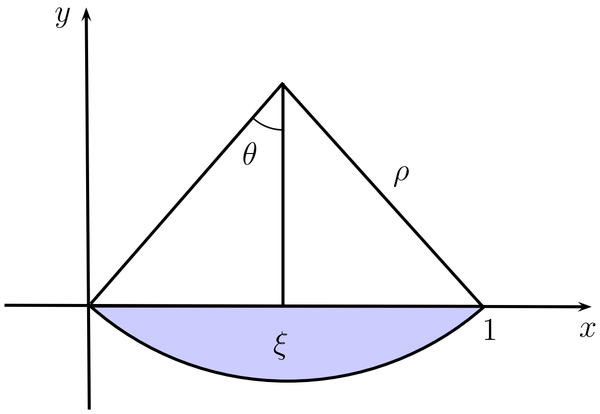


Figure 5: Projection of the geodesic joining the origin with $(1, 0, \xi)$ in $\mathbb{H}$.

The highlighted circle segment has area equal to ξ. Observe that $\theta \in (-\pi, \pi)$, with $\theta = 0$ corresponding to $\xi = 0$ and $\theta \to \pm\pi$ corresponding to $\xi \to \pm\infty$. Then

$$\xi = \theta\rho^2 - \frac{\rho\cos\theta}{2}.$$

Since $2\rho\sin\theta = 1$, we obtain the following equation

$$4\xi = \frac{\theta}{\sin^2\theta} - \cot\theta. \tag{50}$$

The right hand side of Eq. (50) is a smooth and strictly monotone function of θ, for $\theta \in (-\pi, \pi)$. Therefore the function $\theta : \xi \mapsto \theta(\xi)$ is well defined and smooth. Moreover θ is an odd function and, by Eq. (50), it satisfies the following differential equation

$$\frac{d}{d\xi}\left(\frac{\theta^2}{\sin^2\theta}\right) = 4\theta.$$

Finally, the squared distance from the origin of the point $(1, 0, \xi)$ is the Euclidean squared length of the arc, i.e.

$$\mathsf{d}_0^2(1, 0, \xi) = \frac{\theta^2(\xi)}{\sin^2\theta(\xi)}. \tag{51}$$

Plugging Eq. (51) in Eq. (49), we obtain the formula for the squared distance:

$$\mathsf{d}_0^2(r, \phi, z) = r^2 \frac{\theta^2(z/r^2)}{\sin^2\theta(z/r^2)}. \tag{52}$$

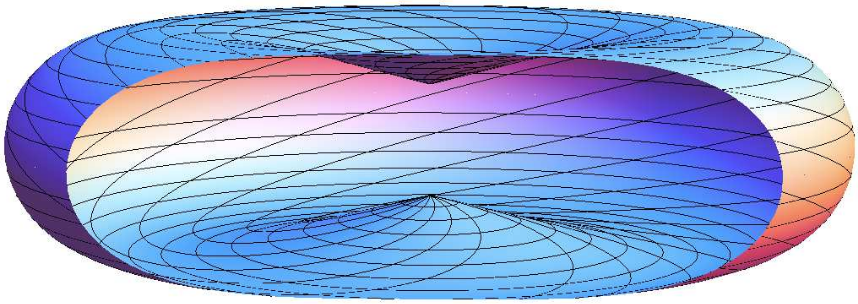


Figure 6: A picture of the sub-Riemannian sphere defined by $\mathsf{d}_0 = 1$.

5.7.2 Asymptotic expansion of the distance

Next we investigate, for two given geodesics γ_1, γ_2 in $\mathbb{H}$ starting from the origin and associated with covectors $\lambda_1, \lambda_2 \in T_0^*M$, the regularity of the function

$$C(t, s) := \frac{1}{2}\mathsf{d}^2(\gamma_1(t), \gamma_2(s))$$

in a neighbourhood of $(t, s) = (0, 0)$. By left-invariance, one has

$$C(t, s) = \frac{1}{2}\mathsf{d}_0^2(\gamma_1(t)^{-1} \cdot \gamma_2(s)).$$

Let $(W_{t,s}, Z_{t,s})$ be the complex coordinates for the point $\gamma_1(t)^{-1} \cdot \gamma_2(s) \in \mathbb{H}$. Moreover, let $R_{t,s}^2 := |W_{t,s}|^2$, and $\xi_{t,s} := Z_{t,s}/R_{t,s}^2$. Then, by Eq. (52),

$$C(t, s) = \frac{1}{2}R_{t,s}^2 \frac{\theta^2(\xi_{t,s})}{\sin^2\theta(\xi_{t,s})}.$$

A long computation, that is sketched in Appendix E, leads to the following result.

Proposition 5.42. *The function $C(t,s)$ is C^1 in a neighbourhood of the origin, but not C^2. In particular, the function $\partial_{ss}C(t,0)$ is not continuous at the origin. However, the singularity at $t=0$ is removable, and the following expansion holds, for $t>0$*

$$\frac{\partial^2 C}{\partial s^2}(t,0) = 1 + 3\sin^2(\phi_2-\phi_1) + \frac{1}{2}[2h_{z,2}\sin(\phi_2-\phi_1) - h_{z,1}\sin(2\phi_2-2\phi_1)]t - \\ - \frac{2}{15}h_{z,1}^2\sin^2(\phi_2-\phi_1)t^2 + O(t^3).$$

If the geodesic γ_2 is chosen to be a straight line (i.e. $h_{z,2}=0$), then

$$\frac{\partial^2 C}{\partial s^2}(t,0) = 1 + 3\sin^2(\phi_2-\phi_1) - \frac{h_{z,1}}{2}\sin(2\phi_2-2\phi_1)t - \frac{2}{15}h_{z,1}^2\sin^2(\phi_2-\phi_1)t^2 + O(t^3), \quad (53)$$

where $\lambda_j = (-\sin\phi_j, \cos\phi_j, h_{z,j}) \in T_0^*M$ is the initial covector of the geodesic γ_j.

We stress once again that, for a Riemannian structure, the function $C(t,s)$ (which can be defined in a completely analogous way as the squared distance between two Riemannian geodesics) is smooth at the origin.

5.7.3 Second differential of the geodesic cost

We are now ready to compute explicitly the asymptotic expansion of $\mathcal{Q}_\lambda$. Fix $w \in T_{x_0}M$ and let $\alpha(s)$ be any geodesic in $\mathbb{H}$ such that $\dot\alpha(0)=w$. Then we compute the quadratic form $d^2_{x_0}\dot c_t(w)$ for $t>0$

$$\begin{aligned}\langle \mathcal{Q}_\lambda(t)w|w\rangle = d^2_{x_0}\dot c_t(w) &= \left.\frac{\partial^2}{\partial s^2}\right|_{s=0}\frac{\partial}{\partial t}c_t(\alpha(s)) = \\ &= \left.\frac{\partial^2}{\partial s^2}\right|_{s=0}\frac{\partial}{\partial t}\left(-\frac{1}{2t}\mathsf{d}^2(\gamma(t),\alpha(s))\right) = \frac{\partial}{\partial t}\left(-\frac{1}{t}\frac{\partial^2 C}{\partial s^2}(t,0)\right) = \\ &= \frac{1}{t^2}\left(\lim_{t\to0^+}\frac{\partial^2 C}{\partial s^2}(t,0)\right) + \frac{1}{3}\left(-\frac{3}{2}\lim_{t\to0^+}\frac{\partial^4 C}{\partial t^2\partial s^2}(t,0)\right) + O(t),\end{aligned}$$

where, in the second line, we exchanged the order of derivations by smoothness of $C(t,s)$ for $t>0$. It is enough to compute the value of $\mathcal{Q}_\lambda(t)$ on an orthonormal basis $v := \dot\gamma(0)$ and $v^\perp := \dot\gamma(0)^\perp$. By using the results of Proposition 5.42, we obtain

$$\langle \mathcal{Q}_\lambda(t)v|v\rangle = \frac{1}{t^2} + O(t), \qquad \langle \mathcal{Q}_\lambda(t)v^\perp|v^\perp\rangle = \frac{4}{t^2} + \frac{2}{15}h_z^2 + O(t).$$

By polarization we obtain $\langle \mathcal{Q}_\lambda(t)v|v^\perp\rangle = O(t)$. Thus the matrices representing the symmetric operators $\mathcal{I}_\lambda$ and $\mathcal{R}_\lambda$ in the basis $\{v, v^\perp\}$ of $\mathscr{D}_{x_0}$ are

$$\mathcal{I}_\lambda = \begin{pmatrix}1 & 0\\ 0 & 4\end{pmatrix}, \qquad \mathcal{R}_\lambda = \frac{2}{5}h_z^2\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \quad (54)$$

where, we recall, λ has coordinates (h_x, h_y, h_z).

Another way to obtain Eq. (54) is to exploit the connection between the curvature operator and the invariants of the Jacobi curves obtained in the proof of Theorem B (see Eqs. (79)–(80)), in terms of a canonical frame. The latter is not easy to compute, even though, in principle, an algorithmic construction is possible. The interested reader can see [36] for the general case, [22] for contact structures with symmetries, and [8] for a more explicit expression for the 3D contact case and, in particular, the Heisenberg group.

Explicit computations for the curvature of a contact sub-Riemannian structure will appear in a forthcoming paper [4].

5.7.4 Sub-Laplacian of the geodesic cost

By using the results of Proposition 5.42, we explicitly compute the asymptotics of the sub-Laplacian Δ_μ of the function $\mathfrak{f}_t = \frac{1}{2}\mathsf{d}^2(\cdot,\gamma(t))$ at x_0, at the second order in t. In the Heisenberg group, we fix $\mu = dx \wedge dy \wedge dz$ (i.e. the Popp's volume of $\mathbb{H}$), and we suppress the explicit dependence of Δ_μ from the volume form.

Since the sub-Riemannian structure of the Heisenberg group is left-invariant, we can reduce the computation of the asymptotic of $\Delta\mathfrak{f}_t$ to the case of a geodesic γ starting from the origin. Indeed, let us denote by $L_g : \mathbb{H} \to \mathbb{H}$ the left multiplication by $g \in \mathbb{H}$. It is easy to show that if $\gamma(t) = \mathcal{E}_{x_0}(t,\lambda)$ is a geodesic, then $\widetilde{\gamma}(t) := L_g(\gamma(t))$ is a geodesic too. If $\mathfrak{f}_t$ and $\widetilde{\mathfrak{f}}_t$ denote the squared distance along the geodesics γ and $\widetilde{\gamma}$, respectively, we have

$$\widetilde{\mathfrak{f}}_t(L_g(x)) = \frac{1}{2}\mathsf{d}^2(L_g(x),\widetilde{\gamma}(t)) = \frac{1}{2}\mathsf{d}^2(L_g(x),L_g(\gamma(t))) = \frac{1}{2}\mathsf{d}^2(x,\gamma(t)) = \mathfrak{f}_t(x).$$

Moreover, by using Proposition 4.3, and recalling the relation $c_t = -t\mathfrak{f}_t$, it is easy to show that

$$\widetilde{\gamma}(t) = \mathcal{E}_{y_0}(t,\eta), \qquad \text{where} \qquad y_0 = L_g(x_0), \quad \eta = (L_g^*)^{-1}\lambda \in T^*_{y_0}M.$$

Moreover Δ is left-invariant hence $\Delta(f \circ L_g) = \Delta f \circ L_g$ for every $f \in C^\infty(M)$, and we have

$$\Delta\widetilde{\mathfrak{f}}_t|_{y_0} = \Delta\mathfrak{f}_t|_{x_0}.$$

In terms of an orthonormal frame, the sub-Laplacian is $\Delta = X^2 + Y^2$ hence

$$\Delta\mathfrak{f}_t|_{x_0} = \left.\frac{d^2}{ds^2}\right|_{s=0} \mathfrak{f}_t(e^{sX}(x_0)) + \left.\frac{d^2}{ds^2}\right|_{s=0} \mathfrak{f}_t(e^{sY}(x_0)), \tag{55}$$

where $e^{sX}(x_0)$ denote the integral curve of the vector field X starting from x_0 (and similarly for Y). Observe that the integral curves of the vector fields X and Y, starting from the origin, are two orthogonal straight lines contained in the xy-plane. Thus we can compute Eq. (55) (where $x_0 = 0$) by summing two copies of Eq. (53) for $\phi_2 = -\pi/2$ and $\phi_2 = 0$ respectively. By left-invariance we immediately find, for any $x_0 \in \mathbb{H}$

$$\Delta\mathfrak{f}_t|_{x_0} = 5 - \frac{2}{15}h_z^2 t^2 + O(t^3),$$

where, we recall, the initial covector associated with the geodesic γ is $\lambda = (h_x,h_y,h_z) \in T^*_{x_0}M$.

Part II

Technical tools and proofs

6 Jacobi curves

In this section we introduce the notion of Jacobi curve associated with a normal geodesic, that is a curve of Lagrangian subspaces in a symplectic vector space. This curve arises naturally from the geometric interpretation of the second derivative of the geodesic cost, and is closely related with the asymptotic expansion of Theorem A.

We start with a brief description of the properties of curves in the Lagrange Grassmannian. For more details, see [5, 10, 36].

6.1 Curves in the Lagrange Grassmannian

Let (Σ, σ) be a $2n$-dimensional symplectic vector space. A subspace $\Lambda \subset \Sigma$ is called *Lagrangian* if it has dimension n and $\sigma|_\Lambda \equiv 0$. The *Lagrange Grassmannian* $L(\Sigma)$ is the set of all n-dimensional Lagrangian subspaces of Σ.

Proposition 6.1. *$L(\Sigma)$ is a compact $n(n+1)/2$-dimensional submanifold of the Grassmannian of n-planes in Σ.*

Proof. Let $\Delta \in L(\Sigma)$, and consider the set $\Delta^{\pitchfork} \doteq \{\Lambda \in L(\Sigma) \,|\, \Lambda \cap \Delta = 0\}$ of all Lagrangian subspaces transversal to Δ. Clearly, the collection of these sets for all $\Delta \in L(\Sigma)$ is an open cover of $L(\Sigma)$. Then it is sufficient to find submanifold coordinates on each $\Delta^{\pitchfork}$.

Let us fix any Lagrangian complement Π of Δ (which always exists, tough it is not unique). Every n-dimensional subspace $\Lambda \subset \Sigma$ that is transversal to Δ is the graph of a linear map from Π to Δ. Choose an adapted Darboux basis on Σ, namely a basis $\{e_i, f_i\}_{i=1}^n$ such that

$$\Delta = \mathrm{span}\{f_1, \dots, f_n\}, \qquad \Pi = \mathrm{span}\{e_1, \dots, e_n\},$$
$$\sigma(e_i, f_j) - \delta_{ij} = \sigma(f_i, f_j) = \sigma(e_i, e_j) = 0, \qquad i, j = 1, \dots, n.$$

In these coordinates, the linear map is represented by a matrix S_Λ such that

$$\Lambda \cap \Delta = 0 \Leftrightarrow \Lambda = \{z = (p, S_\Lambda p),\, p \in \Pi \simeq \mathbb{R}^n\}.$$

Moreover it is easily seen that $\Lambda \in L(\Sigma)$ if and only if $S_\Lambda = S_\Lambda^*$. Hence, the open set $\Delta^{\pitchfork}$ of all Lagrangian subspaces transversal to Δ is parametrized by the set of symmetric matrices, and this gives smooth submanifold coordinates on $\Delta^{\pitchfork}$. This also proves that the dimension of $L(\Sigma)$ is $n(n+1)/2$. Finally, as a closed subset of a compact manifold, $L(\Sigma)$ is compact. □

Fix now $\Lambda \in L(\Sigma)$. The tangent space $T_\Lambda L(\Sigma)$ to the Lagrange Grassmannian at the point Λ can be canonically identified with the set of quadratic forms on the space Λ itself, namely

$$T_\Lambda L(\Sigma) \simeq Q(\Lambda).$$

Indeed, consider a smooth curve $\Lambda(\cdot)$ in $L(\Sigma)$ such that $\Lambda(0) = \Lambda$, and denote by $\dot\Lambda \in T_\Lambda L(\Sigma)$ its tangent vector. For any point $z \in \Lambda$ and any smooth extension $z(t) \in \Lambda(t)$, we define the quadratic form

$$\dot\Lambda \doteq z \mapsto \sigma(z, \dot z),$$

where $\dot z \doteq \dot z(0)$. A simple check shows that the definition does not depend on the extension $z(t)$. Finally, if in local coordinates $\Lambda(t) = \{(p, S(t)p),\, p \in \mathbb{R}^n\}$, the quadratic form $\dot\Lambda$ is represented by the matrix $\dot S(0)$. In other words, if $z \in \Lambda$ has coordinates $p \in \mathbb{R}^n$, then $\dot\Lambda : p \mapsto p^* \dot S(0) p$.

6.1.1 Ample, equiregular, monotone curves

Let $J(\cdot) \in L(\Sigma)$ be a smooth curve in the Lagrange Grassmannian. For $i \in \mathbb{N}$, consider

$$J^{(i)}(t) = \operatorname{span}\left\{ \frac{d^j}{dt^j}\ell(t) \middle| \; \ell(t) \in J(t),\, \ell(t) \text{ smooth},\, 0 \leq j \leq i \right\} \subset \Sigma, \qquad i \geq 0.$$

Definition 6.2. The subspace $J^{(i)}(t)$ is the *i-th extension* of the curve $J(\cdot)$ at t. The flag

$$J(t) = J^{(0)}(t) \subset J^{(1)}(t) \subset J^{(2)}(t) \subset \ldots \subset \Sigma,$$

is the *associated flag of the curve* at the point t. The curve $J(\cdot)$ is called:

(i) *equiregular* at t if $\dim J^{(i)}(\cdot)$ is locally constant at t, for all $i \in \mathbb{N}$,

(ii) *ample* at t if there exists $N \in \mathbb{N}$ such that $J^{(N)}(t) = \Sigma$,

(iii) *monotone increasing* (resp. *decreasing*) at t if $\dot{J}(t)$ is non-negative definite (resp. non-positive definite) as a quadratic form.

The *step* of the curve at t is the minimal $N \in \mathbb{N}$ such that $J^{(N)}(t) = \Sigma$.

In coordinates, $J(t) = \{(p, S(t)p) |\; p \in \mathbb{R}^n\}$ for some smooth family of symmetric matrices $S(t)$. The curve is ample at t if and only if there exists $N \in \mathbb{N}$ such that

$$\operatorname{rank}\{\dot{S}(t), \ddot{S}(t), \ldots, S^{(N)}(t)\} = n.$$

The *rank* of the curve at t is the rank of $\dot{J}(t)$ as a quadratic form (or, equivalently, the rank of $\dot{S}(t)$). We say that the curve is equiregular, ample or monotone (increasing or decreasing) if it is equiregular, ample or monotone for all t in the domain of the curve.

In the subsequent sections we show that with any ample (resp. equiregular) geodesic, we can associate in a natural way an ample (resp. equiregular) curve in an appropriate Lagrange Grassmannian. This justifies the terminology introduced in Definition 6.2.

An important property of ample, monotone curves is described in the following lemma.

Lemma 6.3. *Let $J(\cdot) \in L(\Sigma)$ be a monotone, ample curve at t_0. Then, there exists $\varepsilon > 0$ such that $J(t) \cap J(t_0) = \{0\}$ for $0 < |t - t_0| < \varepsilon$.*

Proof. Without loss of generality, assume $t_0 = 0$. Choose a Lagrangian splitting $\Sigma = \Lambda \oplus \Pi$, with $\Lambda = J(0)$. For $|t| < \varepsilon$, the curve is contained in the chart defined by such a splitting. In coordinates, $J(t) = \{(p, S(t)p) |\; p \in \mathbb{R}^n\}$, with $S(t)$ symmetric and $S(0) = 0$. The curve is monotone, then $\dot{S}(t)$ is a semidefinite symmetric matrix. It follows that $S(t)$ is semidefinite too.

Suppose that, for some τ, $J(\tau) \cap J(0) \neq \{0\}$ (w.l.o.g. assume $\tau > 0$). This means that $\exists p \in \mathbb{R}^n$ such that $S(\tau)p = 0$. Indeed also $p^*S(\tau)p = 0$. The function $t \mapsto p^*S(t)p = 0$ is monotone, vanishing at $t = 0$ and $t = \tau$. Therefore $p^*S(t)p = 0$ for all $0 \leq t \leq \tau$. Being a semidefinite, symmetric matrix, $p^*S(t)p = 0$ if and only if $S(t)p = 0$. Therefore, we conclude that $p \in \ker S(t)$ for $0 \leq t \leq \tau$. This implies that, for any $i \in \mathbb{N}$, $p \in \ker S^{(i)}(0)$, which is a contradiction, since the curve is ample at 0. ∎

Remark 6.4. Ample curves with $N = 1$ are also called *regular*. See in particular [5, 10], where the authors discuss geometric invariants of these curves. Notice that a curve $J(\cdot)$ is regular at t if and only if its tangent vector at t is a non degenerate quadratic form, i.e. the matrix $\dot{S}(t)$ is invertible.

6.1.2 The Young diagram of an equiregular Jacobi curve

Let $J(\cdot) \in L(\Sigma)$ be smooth, ample and equiregular. We can associate in a standard way a Young diagram with the curve $J(\cdot)$ as follows. Consider the restriction of the curve to a neighbourhood of t such that, for all $i \in N$, $\dim J^{(i)}(\cdot)$ is constant. Let $h_i \doteq \dim J^{(i)}(\cdot)$. By hypotesis, there exists a minimal $N \in \mathbb{N}$ such that $h_i = \dim \Sigma$ for all $i \geq N$. It follows from the definition of extension that, for $i \in \mathbb{N}$, we have the inequalities $h_{i+1} - h_i \leq h_i - h_{i-1}$. Then, we build a Young diagram with N columns, with $h_i - h_{i-1}$ boxes in the i-th column. This is the *Young diagram of the curve* $J(\cdot)$. In particular, notice that the number of boxes in the first column is equal to the rank of $J(\cdot)$.

6.2 The Jacobi curve and the second differential of the geodesic cost

Recall that T^*M has a natural structure of symplectic manifold, with the canonical symplectic form defined as the differential of the Liouville form, namely $\sigma = d\varsigma$. In particular, for any $\lambda \in T^*M$, $T_\lambda(T^*M)$ is a symplectic vector space with the canonical symplectic form σ. Therefore, we can specify the construction above to $\Sigma \doteq T_\lambda(T^*M)$. In this section we show that the second derivative of the geodesic cost (associated with an ample geodesic γ with initial covector $\lambda \in T^*M$) can be naturally interpreted as a curve in the Lagrange Grassmannian of $T_\lambda(T^*M)$, which is ample in the sense of Definition 6.2.

6.2.1 Second differential at a non critical point

Let $f \in C^\infty(M)$. As we explained in Section 4.3, the second differential of f, which is a symmetric bilinear form on the tangent space, is well defined only at critical points of f. If $x \in M$ is not a critical point, it is still possible to define the second differential of f, as the differential of df, thought as a section of T^*M.

Definition 6.5. Let $f \in C^\infty(M)$, and

$$df : M \to T^*M, \qquad df : x \mapsto d_x f.$$

Fix $x \in M$, and let $\lambda \doteq d_x f \in T^*M$. The *second differential* of f at $x \in M$ is the linear map

$$d_x^2 f \doteq d_x(df) : T_x M \to T_\lambda(T^*M), \qquad d_x^2 f : v \mapsto \left.\frac{d}{ds}\right|_{s=0} d_{\gamma(s)} f,$$

where $\gamma(\cdot)$ is a curve on M such that $\gamma(0) = x$ and $\dot\gamma(0) = v$.

Definition 6.5 generalizes the concept of "second derivatives" of f, as the linearisation of the differential.

Remark 6.6. The image of the differential $df : M \to T^*M$ is a Lagrangian submanifold of T^*M. Thus, by definition, the image of the second differential $d_x^2 f(T_x M)$ at a point x is the tangent space of $df(M)$ at $\lambda = d_x f$, which is an n-dimensional Lagrangian subspace of $T_\lambda(T^*M)$ transversal to the vertical subspace $T_\lambda(T_x^*M)$.

By a dimensional argument and the fact that $\pi \circ df = \mathbb{I}_M$ (hence $\pi_* \circ d_x^2 f = \mathbb{I}_{T_x M}$), we obtain the following formula for the image of a subspace through the second differential.

Lemma 6.7. *Let $f : M \to \mathbb{R}$ and $W \subset T_x M$. Then $d_x^2 f(W) = d_x^2 f(T_x M) \cap \pi_*^{-1}(W)$.*

The next lemma describes the affine structure on the space of second differentials.

Lemma 6.8. *Let $\lambda \in T_x^*M$. The set $\mathcal{L}_\lambda \doteq \{d_x^2 f |\, f \in C^\infty(M), d_x f = \lambda\}$ is an affine space over the vector space $Q(T_x M)$ of the quadratic forms over $T_x M$.*

Proof. Consider two functions f_1, f_2 such that $d_x f_1 = d_x f_2 = \lambda$. Then $f_1 - f_2$ has a critical point at x. We define the difference between $d_x^2 f_1$ and $d_x^2 f_2$ as the quadratic form $d_x^2(f_1 - f_2)$. □

Remark 6.9. When $\lambda = 0 \in T_x^*M$, $\mathcal{L}_\lambda$ is the space of the second derivatives of the functions with a critical point at x. In this case we can fix a canonical origin in $\mathcal{L}_\lambda$, namely the second differential of any constant function. This gives the identification of $\mathcal{L}_\lambda$ with the space of quadratic forms on $T_x M$, recovering the standard notion of Hessian discussed in Section 4.3.

6.2.2 Second differential of the geodesic cost

Let $\gamma : [0, T] \to M$ be a strongly normal geodesic. Let $x = \gamma(0)$. Without loss of generality, we can choose T sufficiently small so that the map $(t, x) \to c_t(x)$ is smooth in a neighbourhood of $(0, T) \times \{x\} \subset \mathbb{R} \times M$, and $d_x c_t = \lambda$ is the initial covector associated with γ (see Theorem 4.2 and Proposition 4.3).

The second differential of c_t defines a curve in the Lagrange Grassmannian $L(T_\lambda(T^*M))$. For any $\lambda \in T^*M$, $\pi(\lambda) = x$, we denote with the symbol $\mathcal{V}_\lambda = T_\lambda(T_x^*M) \subset T_\lambda(T^*M)$ the vertical subspace, namely the tangent space to the fiber T_x^*M. Observe that, if $\pi : T^*M \to M$ is the bundle projection, $\mathcal{V}_\lambda = \ker \pi_*$.

Definition 6.10. The *Jacobi curve* associated with γ is the smooth curve $J_\lambda : [0, T] \to L(T_\lambda(T^*M))$ defined by

$$J_\lambda(t) \doteq d_x^2 c_t(T_x M),$$

for $t \in (0, T]$, and $J_\lambda(0) \doteq \mathcal{V}_\lambda$.

The Jacobi curve is smooth as a consequence of the next proposition, which provides an equivalent characterization of the Jacobi curve in terms of the Hamiltonian flow on T^*M.

Proposition 6.11. *Let $\lambda : [0, T] \to T^*M$ be the unique lift of γ such that $\lambda(t) = e^{t\vec{H}}(\lambda)$. Then the associated Jacobi curve satisfies the following properties for all t, s such that both sides of the statements are defined:*

(i) $J_\lambda(t) = e_*^{-t\vec{H}} \mathcal{V}_{\lambda(t)}$,

(ii) $J_\lambda(t + s) = e_*^{-t\vec{H}} J_{\lambda(t)}(s)$,

(iii) $\dot{J}_\lambda(0) = -d_\lambda^2 H_x$ *as quadratic forms on* $\mathcal{V}_\lambda \simeq T_x^*M$.

Proof. In order to prove (i) it is sufficient to show that $\pi_* \circ e_*^{t\vec{H}} \circ d_x^2 c_t = 0$. Then, let $v \in T_x M$, and $\alpha(\cdot)$ a smooth arc such that $\alpha(0) = x$, $\dot{\alpha}(0) = v$. Recall that, for s sufficiently small, $d_{\alpha(s)} c_t$ is the initial covector of the unique normal geodesic which connects $\alpha(s)$ with $\gamma(t)$ in time t, i.e. $\pi \circ e^{t\vec{H}} \circ d_{\alpha(s)} c_t = \gamma(t)$. Then

$$\pi_* \circ e_*^{t\vec{H}} \circ d_x^2 c_t(v) = \left.\frac{d}{ds}\right|_{s=0} \pi \circ e^{t\vec{H}} \circ d_{\alpha(s)} c_t = 0.$$

Statement (ii) follows from (i) and the group property of the Hamiltonian flow. To prove (iii), introduce canonical coordinates (p, x) in the cotangent bundle. Let $\xi \in \mathcal{V}_\lambda$, such that $\xi = \sum_{i=1}^n \xi_i \partial_{p_i}|_\lambda$. By (i), the smooth family of vectors in $\mathcal{V}_\lambda$ defined by

$$\xi(t) \doteq e_*^{-t\vec{H}} \left(\sum_{i=1}^n \xi^i \partial_{p_i}|_{\lambda(t)} \right),$$

satisfies $\xi(0) = \xi$ and $\xi(t) \in J_\lambda(t)$. Therefore

$$\dot{J}_\lambda(0)\xi = \sigma(\xi, \dot{\xi}) = -\sum_{i,j=1}^{n} \frac{\partial^2 H}{\partial p_i \partial p_j} \xi^i \xi^j = -\langle \xi, (d^2_\lambda H_x)\xi \rangle,$$

where the last equality follows from the definition of $d^2_\lambda H_x$ after the identification $\mathcal{V}_\lambda \simeq T^*_x M$ (see Section 4.2). □

Remark 6.12. Point (i) of Proposition 6.11 can be used to associate a Jacobi curve with any integral curve of the Hamiltonian flow, without any further assumptions on the underlying trajectory on the manifold. In particular we associate with any initial covector $\lambda \in T_x M$ the Jacobi curve $J_\lambda(t) \doteq e^{-t\vec{H}} \mathcal{V}_{\lambda(t)}$. Observe that, in general, $\gamma(\cdot) \doteq \pi \circ \lambda(\cdot)$ may be also abnormal.

Proposition 6.11 and the fact that the quadratic form $d^2_\lambda H_x$ is non-negative definite imply the next corollary.

Corollary 6.13. *The Jacobi curve J_λ is monotone nonincreasing for every $\lambda \in T^*M$.*

The following proposition provides the connection between the flag of a normal geodesic and the flag of the associated Jacobi curve.

Proposition 6.14. *Let $\gamma(t) = \pi \circ e^{t\vec{H}}(\lambda)$ be a normal geodesic associated with the initial covector λ. The flag of the Jacobi curve J_λ projects to the flag of the geodesic γ at $t = 0$, namely*

$$\pi_* J^{(i)}_\lambda(0) = \mathscr{F}^i_\gamma(0), \qquad \forall\, i \in \mathbb{N}. \tag{56}$$

Moreover, $\dim J^{(i)}_\lambda(t) = n + \dim \mathscr{F}^i_\gamma(t)$. *Therefore γ is ample of step m (resp. equiregular) if and only if J_λ is ample of step m (resp. equiregular).*

Proof. The last statement follows directly from Eq. (56), Proposition 6.11 (point (ii)) and the definition of $\mathscr{F}_{\gamma(s)}(t) = (P_{s,s+t})^{-1}_* \mathscr{D}_{\gamma(s+t)}$. In order to prove Eq. (56), let $\bar{u} : T^*M \to L^\infty([0,T], \mathbb{R}^k)$ be the map that associates to any covector the corresponding normal control:

$$\bar{u}_i(\lambda)(\cdot) = \langle e^{\cdot \vec{H}}(\lambda), f_i \rangle, \qquad i = 1, \dots, k,$$

where we assume, without loss of generality, that the Hamiltonian field $\vec{H}$ is complete. For any control $v \in L^\infty([0,T], \mathbb{R}^k)$ and initial point $x \in M$, consider the non-autonomous flow $P^v_{0,t}(x)$. We have the following identity, for any $\lambda \in T^*M$ and $t \in [0,T]$

$$\pi \circ e^{t\vec{H}}(\lambda) = P^{\bar{u}(\lambda)}_{0,t}(\pi(\lambda)).$$

Remember that, as a function of the control, $P^v_{0,t}(x) = E_{x,t}(v)$ (i.e. the endpoint map with basepoint x and endtime t). Therefore, by taking the differential at λ (such that $\pi(\lambda) = x$), we obtain

$$\pi_* \circ e^{t\vec{H}}_*|_\lambda = \left(P^{\bar{u}(\lambda)}_{0,t}\right)_* \circ \pi_*|_\lambda + D_{\bar{u}(\lambda)} E_{x,t} \circ \bar{u}_*|_\lambda,$$

Then, by the explicit formula for the differential of the endpoint map, we obtain, for any vertical field $\xi(t) \in \mathcal{V}_{e^{t\vec{H}}(\lambda)}$

$$\pi_* \circ e^{-t\vec{H}}_* \xi(t) = -\int_0^t (P_{0,\tau})^{-1}_* \overline{f}(v(t,\tau), \gamma(t)) d\tau,$$

where $\gamma(t) = \pi \circ e^{t\vec{H}}(\lambda)$ is the normal geodesic with initial covector λ and, for any $t \in [0,T]$,

$$v_i(t,\cdot) \doteq \bar{u}_* \circ e^{-t\vec{H}}_* \xi(t) = \left(\bar{u} \circ e^{-t\vec{H}}\right)_* \xi(t), \qquad v(t,\cdot) \in L^\infty([0,T], \mathbb{R}^k).$$

More precisely, $v(t,\cdot)$ has components

$$v_i(t,\tau) = \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \langle e^{(\tau-t)\vec{H}}(\lambda(t)+\varepsilon\xi(t)), f_i\rangle, \qquad i=1,\dots,k,$$

where $\lambda(t) = e^{t\vec{H}}(\lambda)$, and we identified $\mathcal{V}_{e^{t\vec{H}}(\lambda)} \simeq T^*_{\gamma(t)}M$. Observe that, on the diagonal, $v_i(t,t) = \langle \xi(t), f_i\rangle = \xi_i(t)$. It is now easy to show that, for any positive $i \in \mathbb{N}$

$$\frac{d^i}{dt^i}\bigg|_{t=0} \pi_* \circ e_*^{-t\vec{H}}\xi(t) = -\frac{d^{i-1}}{dt^{i-1}}\bigg|_{t=0}\left[(P_{0,t})_*^{-1}\sum_{j=1}^k \xi_j(t)\overline{f}_j(\gamma(t))\right] \mod \mathscr{F}_\gamma^{i-1}(0). \tag{57}$$

By point (i) of Proposition 6.11, any smooth family $\ell(t) \in J_\lambda(t)$ is of the form $e_*^{-t\vec{H}}\xi(t)$ for some smooth $\xi(t) \in \mathcal{V}_{e^{t\vec{H}}(\lambda)}$. Therefore, Eq. (57) for $i=1$ implies that $J_\lambda^{(1)} = \mathscr{F}_\gamma^1(0)$. The same equation and an easy induction argument, together with the definitions of the flags show that $J_\lambda^{(i)}(0) = \mathscr{F}_\gamma^i(0)$ for any positive $i \in \mathbb{N}$. □

Remark 6.15. Observe that, if γ is equiregular, ample of step m with growth vector $\mathcal{G}_\lambda = (k_1, k_2, \dots, k_m)$, the Young diagram of J_λ has m columns, with $d_i \doteq k_i - k_{i-1}$ boxes in the i-th column (recall that $k_0 = \dim \mathscr{F}_\gamma^0(t) = 0$).

Remark 6.16. We already observed that the ampleness condition is a rank condition on the derivatives of the symmetric matrix $S_\lambda(t)$ that represents the Jacobi curve $J_\lambda(t)$. In particular the curve is ample at $t=0$ if and only if there exists $N \in \mathbb{N}$ such that

$$\operatorname{rank}\{\dot{S}_\lambda(0), \ddot{S}_\lambda(0), \dots, S_\lambda^{(N)}(0)\} = n.$$

By point (i) of Proposition 6.11 it follows that, for any fibre-wise polynomial Hamiltonian, $S_\lambda^{(i)}(0)$ is a polynomial function of the initial covector $\lambda \in T_x^*M$, for any $i \in \mathbb{N}$. Therefore, under this assumption (which is true, for example, in the sub-Riemannian case), $J_\lambda(\cdot)$ is ample on an open Zariski subset of the fibre T_x^*M.

6.3 The Jacobi curve and the Hamiltonian inner product

The following is an elementary, albeit very useful property of the symplectic form σ.

Lemma 6.17. *Let $\xi \in \mathcal{V}_\lambda$ a vertical vector. Then, for any $\eta \in T_\lambda(T^*M)$*

$$\sigma(\xi,\eta) = \langle \xi, \pi_*\eta\rangle,$$

where we employed the canonical identification $\mathcal{V}_\lambda = T_x^*M$.

Proof. In any Darboux basis induced by canonical local coordinates (p,x) on T^*M, we have $\sigma = \sum_{i=1}^n dp_i \wedge dx_i$ and $\xi = \sum_{i=1}^n \xi^i \partial_{p_i}$. The result follows immediately. □

In Section 4.2 we introduced the Hamiltonian inner product on $\mathscr{D}_x$, which, in general, depends on λ. Such an inner product is defined by the quadratic form $d_\lambda^2 H_x : T_x^*M \to T_xM$ on $\mathscr{D}_x = \operatorname{Im}(d_\lambda^2 H_x)$. The following lemma allows the practical computation of the Hamiltonian inner product through the Jacobi curve.

Lemma 6.18. *Let $\xi \in T_x^*M$. Then*

$$d_\lambda^2 H_x(\xi) = -\pi_*\dot{\xi},$$

where $\dot{\xi}$ is the derivative, at $t=0$, of any extension $\xi(t)$ of ξ such that $\xi(0)=\xi$ and $\xi(t) \in J_\lambda(t)$.

Proof. By point (iii) of Proposition 6.11, $d^2_\lambda H_x = -\dot{J}_\lambda(0)$. By definition of $\dot{J}_\lambda(0) : \mathcal{V}_\lambda \to \mathbb{R}$ as a quadratic form, $\dot{J}_\lambda(0)(\xi) = \sigma(\xi, \dot{\xi})$. Then, by Lemma 6.17, $\dot{J}_\lambda(0)(\xi) = \langle \xi, \pi_* \dot{\xi} \rangle$. This implies the statement after identifying again the quadratic form with the associated symmetric map. □

By Lemma 6.18, for any $v \in \mathscr{D}_x$ there exists a $\xi \in \mathcal{V}_\lambda$ such that, for any extension $\xi(t) \in J_\lambda(t)$, with $\xi(0) = \xi$, we have $v = \pi_* \dot{\xi}$. Indeed ξ may not be unique. Besides, if $v = \pi_* \dot{\xi}$ and $w = \pi_* \dot{\eta}$, the Hamiltonian inner product rewrites

$$\langle v | w \rangle_\lambda = \sigma(\xi, \dot{\eta}) = -\sigma(\eta, \dot{\xi}). \tag{58}$$

We now have all the tools required for the proof of Theorem A.

6.4 Proof of Theorem A

The statement of Theorem A is related with the analytic properties of the functions $t \mapsto \langle \mathcal{Q}_\lambda(t) v | v \rangle_\lambda$ for $v \in \mathscr{D}_x$. By definition, $\langle \mathcal{Q}_\lambda(t) v | v \rangle_\lambda = d^2_x \dot{c}_t(v)$.

As a first step, we compute a coordinate formula for such a function in terms of a splitting $\Sigma = \mathcal{V}_\lambda \oplus \mathcal{H}_\lambda$, where $\mathcal{V}_\lambda$ is the vertical space and $\mathcal{H}_\lambda$ is any Lagrangian complement. Observe that $\mathcal{V}_\lambda = J_\lambda(0) = \ker \pi_*$ and π_* induces an isomorphism between $\mathcal{H}_\lambda$ and $T_x M$. $J_\lambda(t)$ is the graph of a linear map $S(t) : \mathcal{V}_\lambda \to \mathcal{H}_\lambda$. Equivalently, by Lemma 6.3, for $0 < t < \varepsilon$, $J_\lambda(t)$ is the graph of $S(t)^{-1} : \mathcal{H}_\lambda \to \mathcal{V}_\lambda$. Once a Darboux basis (adapted to the splitting) is fixed, as usual one can identify these maps with the representative matrices.

Fix $v \in \mathscr{D}_x \subset T_x M$ and let $\widetilde{v} \in \mathcal{H}_\lambda$ be the unique horizontal lift such that $\pi_* \widetilde{v} = v$. Then, by definition of Jacobi curve, and the standard identification $\mathcal{V}_\lambda \simeq T^*_x M$

$$\langle \mathcal{Q}_\lambda(t) v | v \rangle_\lambda = \frac{d}{dt} \sigma(S(t)^{-1} \widetilde{v}, \widetilde{v}). \tag{59}$$

Since $J_\lambda(0) = \mathcal{V}_\lambda$, it follows that $S(t)^{-1}$ is singular at $t = 0$. In what follows we prove Theorem A, by computing the asymptotic expansion of the matrix $S(t)^{-1}$. More precisely, from (59) it is clear that we need only a "block" of $S(t)^{-1}$ since it acts only on vectors $\widetilde{v} \in \pi_*^{-1}(\mathscr{D}_x) \cap \mathcal{H}_\lambda$. In what follows we build natural coordinates on the space Σ in such a way that Eq. (59) is given by the derivative of the first $k \times k$ block of $S(t)^{-1}$ where, we recall, $k = \dim \mathscr{D}_x$. Notice that this restriction is crucial in the proof since only the aforementioned block has a simple pole. This is not true, in general, for the whole matrix $S(t)^{-1}$.

6.4.1 Coordinate presentation of the Jacobi curve

In order to obtain a convenient expression for the matrix $S(t)$ we introduce a set of coordinates (p, x) induced by a particular Darboux frame adapted to the splitting $\Sigma = \mathcal{V}_\lambda \oplus \mathcal{H}_\lambda$. Namely

$$\Sigma = \{(p, x) | \, p, x \in \mathbb{R}^n\}, \qquad \mathcal{V}_\lambda = \{(p, 0) | \, p \in \mathbb{R}^n\}, \qquad \mathcal{H}_\lambda = \{(0, x) | \, x \in \mathbb{R}^n\}.$$

Besides, if $\xi = (p, x)$, $\bar{\xi} = (\bar{p}, \bar{x}) \in \Sigma$ the symplectic product is $\sigma(\xi, \bar{\xi}) = p^* \bar{x} - \bar{p}^* x$. In these coordinates, $J_\lambda(t) = \{(p, S(t)p) | \, p \in \mathbb{R}^n\}$, and $S(0) = 0$. The symmetric matrix $S(t)$ represents a monotone Jacobi curve, hence $\dot{S}(t) \leq 0$. Moreover, since the curve is ample, by Lemma 6.3, $S(t) < 0$ for $0 < t < \varepsilon$. Moreover we introduce the coordinate splitting $\mathbb{R}^n = \mathbb{R}^k \oplus \mathbb{R}^{n-k}$ (accordingly we write $p = (p_1, p_2)$ and $x = (x_1, x_2)$), such that $\pi_*(\mathbb{R}^k) = \mathscr{D}_x$. In blocks notation

$$S(t) = \begin{pmatrix} S_{11}(t) & S_{12}(t) \\ S_{12}^*(t) & S_{22}(t) \end{pmatrix}, \qquad \text{with} \quad S_{11}(t), S_{22}(t) < 0 \quad \text{for} \quad 0 < t < \varepsilon.$$

By point (iii) of Proposition 6.11, in these coordinates we also have

$$\dot{S}(0) = \begin{pmatrix} \dot{S}_{11}(0) & 0 \\ 0 & 0 \end{pmatrix}, \qquad \text{with} \quad \operatorname{rank} \dot{S}_{11}(0) = \dim \mathscr{D}_x.$$

Therefore, we obtain the following coordinate formula for the Hamiltonian inner product. Let $v, w \in \mathscr{D}_x$, with coordinates $v = (v_1, 0)$, $w = (w_1, 0)$ then

$$\langle v|w\rangle_\lambda = -v_1^* \dot{S}_{11}(0)^{-1} w_1, \qquad v_1, w_1 \in \mathbb{R}^k,$$

Remark 6.19. In other words, the quadratic form associated with the operator $\mathbb{I} : \mathscr{D}_x \to \mathscr{D}_x$ via the Hamiltonian inner product is represented by the matrix $-\dot{S}_{11}(0)^{-1}$.

Moreover the horizontal lift of v is $\widetilde{v} = ((0,0),(v_1,0))$ and analogously for w. Thus, by (59)

$$\langle \mathcal{Q}_\lambda(t) v|w\rangle_\lambda = \frac{d}{dt} v_1^* [S(t)^{-1}]_{11} w_1, \qquad v_1, w_1 \in \mathbb{R}^k, \quad t > 0. \tag{60}$$

For convenience, for $t > 0$, we introduce the smooth family of $k \times k$ matrices $S^\flat(t)$ defined by

$$S^\flat(t)^{-1} \doteq [S(t)^{-1}]_{11}, \qquad t > 0.$$

Then, the quadratic form associated with the operator $\mathcal{Q}_\lambda(t) : \mathscr{D}_x \to \mathscr{D}_x$ via the Hamiltonian inner is represented by the matrix $\frac{d}{dt} S^\flat(t)^{-1}$.

The proof of Theorem A is based upon the following result.

Theorem 6.20. *The map $t \mapsto S^\flat(t)^{-1}$ has a simple pole at $t = 0$.*

Proof. The expression of $S^\flat(t)$ in terms of the blocks of $S(t)$ is given by the following lemma.

Lemma 6.21. *Let $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right)$ be a sign definite matrix, and denote by $[A^{-1}]_{11}$ the first block of the inverse of A. Then $[A^{-1}]_{11} = (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1}$.*

Then, by definition of $S^\flat$, we have the following formula (where we suppress t):

$$S^\flat = S_{11} - S_{12} S_{22}^{-1} S_{12}^*. \tag{61}$$

Lemma 6.22. *As quadratic forms on $\mathbb{R}^k$, $S_{11}(t) \leq S^\flat(t) < 0$ for $t > 0$.*

Proof of Lemma 6.22. Let $t > 0$. $S(t)$ is symmetric and negative, then also its inverse $S(t)^{-1}$ is symmetric and negative. This implies that $S^\flat(t)^{-1} = [S(t)^{-1}]_{11} < 0$ and so is $S^\flat(t)$. This proves the right inequality. By Eq. (61) and the fact that $S_{22}(t)$ is negative definite (and so is $S_{22}^{-1}(t)$) one also gets (we suppress $t > 0$)

$$p_1^*(S_{11} - S^\flat) p_1 = p_1^* S_{12} S_{22}^{-1} S_{12}^* p_1 = (S_{12}^* p_1)^* S_{22}^{-1} (S_{12}^* p_1) \leq 0, \qquad p_1 \in \mathbb{R}^k.$$

□

Lemma 6.23. *The map $t \mapsto S^\flat(t)$ can be extended by smoothness at $t = 0$.*

Proof. Indeed, by the coordinate expression of Eq. (61), it follows that the only term that can give rise to singularities is the inverse matrix $S_{22}^{-1}(t)$. Since, by assumption, the curve is ample, $t \mapsto \det S_{22}(t)$ has a finite order zero at $t = 0$, thus the singularity can be only a finite order pole. On the other hand $S(t) \to 0$ for $t \to 0$, thus $S_{11}(t) \to 0$ as well. Then, by Lemma 6.22, $S^\flat(t) \to 0$ for $t \to 0$, hence can be extended by smoothness at $t = 0$. □

We are now ready to prove that $t \mapsto S^\flat(t)^{-1}$ has a simple pole at $t=0$. As a byproduct, we obtain an explicit form for its residue. As usual, for $i>0$, we set $k_i \doteq \dim J_\lambda^{(i)}(0) - n$, and $d_i \doteq k_i - k_{i-1}$. In coordinates, this means that

$$\operatorname{rank}\{\dot{S}(0), \dots, S^{(i)}(0)\} = k_i, \qquad i = 1, \dots, m.$$

By hypotesis, the curve is ample at $t=0$, then there exists m such that $k_m = n$. Since we are only interested in Taylor expansions, we may assume $S(t)$ to be real-analytic in $[0,\varepsilon]$ by replacing, if necessary, $S(t)$ with its Taylor polynomial of sufficient high order. Then, let us consider the analytic family of symmetric matrices $\dot{S}(t)$. For $i = 1, \dots, n$, the family $w_i(t)$ of eigenvectors of $\dot{S}(t)$ (and the relative eigenvalues) are an analytic family (see [21, Theorem 6.1, Chapter II]). Therefore, $\dot{S}(t) = W(t)D(t)W(t)^*$, where $W(t)$ is the $n \times n$ matrix whose columns are the vectors $w_i(t)$, and $D(t)$ is a diagonal matrix. Recall that $\dot{S}(t)$ is non-positive definite. Then $\dot{S}(t) = -V(t)V(t)^*$, for some analytic family of $n \times n$ matrices $V(t)$. Let $v_i(t)$ denote the columns of $V(t)$.

Now, let us consider the flag $E_1 \subset E_2 \subset \ldots \subset E_m = \mathbb{R}^n$ defined as follows

$$E_i = \operatorname{span}\{v_j^{(\ell)}(0),\, 1 \leq j \leq n,\, 0 \leq \ell \leq i-1\}.$$

Let $\operatorname{span}\{A\}$ denote the column space of a matrix A. Indeed $\operatorname{span}\{\dot{S}(t)\} \subseteq \operatorname{span}\{V(t)\}$. Besides, $\operatorname{rank}\{\dot{S}(t)\} = \operatorname{rank}\{V(t)V(t)^*\} = \operatorname{rank}\{V(t)\} = \dim \operatorname{span}\{V(t)\}$. Therefore, $\operatorname{span}\{\dot{S}(t)\} = \operatorname{span}\{V(t)\}$, for all $|t| < \varepsilon$. Thus, for $i = 1, \dots, m$

$$E_i = \operatorname{span}\{V(0), V^{(1)}(0), \dots, V^{(i-1)}(0)\} = \operatorname{span}\{\dot{S}(0), \dots, S^{(i)}(0)\}.$$

Therefore $\dim E_i = k_i$. Choose coordinates in $\mathbb{R}^n$ adapted to this flag, i.e. $\operatorname{span}\{e_1, \dots, e_{k_i}\} = E_i$. In these coordinates, $V(t)$ has a peculiar structure, namely

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

where $\widehat{v}_i$ is a $d_i \times n$ matrix of maximal rank (notice that the $\widehat{v}_i$ are not directly related with the columns $v_i(t)$ of $V(t)$). Let $\widehat{V}(t)$ denote the "principal part" of $V(t)$. In other words, $\widehat{V}(t) = (\widehat{v}_1, t\widehat{v}_2, \dots, t^{m-1}\widehat{v}_m)^*$. Then, remember that $S(0) = 0$ and

$$S(t) = \int_0^t \dot{S}(\tau)d\tau = -\int_0^t V(\tau)V(\tau)^* d\tau = -\int_0^t \widehat{V}(\tau)\widehat{V}(\tau)^* d\tau + r(t),$$

where $r(t)$ is a remainder term. Observe that the matrix

$$\widehat{S}(t) = -\int_0^t \widehat{V}(\tau)\widehat{V}(\tau)^* d\tau$$

is negative definite for $t > 0$. In fact, a non trivial kernel for some $t > 0$ would contradict the hypotesis $\operatorname{span}\{V(0), V^{(1)}(0), \dots, V^{(m-1)}(0)\} = \mathbb{R}^n$. In components, we write $S(t)$ as a $m \times m$ block matrix, $S_{ij}(t)$ being a $d_i \times d_j$ block, as follows:

$$S_{ij}(t) = \int_0^t \dot{S}_{ij}(\tau)d\tau = -\left(\frac{\widehat{v}_i\widehat{v}_j^*}{i+j-1}\right) t^{i+j-1} + O(t^{i+j}) = \chi_{ij} t^{i+j-1} + O(t^{i+j}),$$

where we introduced the negative definite constant matrix $\chi \doteq \widehat{S}(1) < 0$. By computing the determinant of $\widehat{S}(t)$, we obtain

$$\det \widehat{S}(t) = \det \begin{pmatrix} t\chi_{11} & t^2\chi_{12} & \cdots & t^m\chi_{1m} \\ t^2\chi_{21} & t^3\chi_{22} & \cdots & t^{m+1}\chi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ t^m\chi_{m1} & t^{m+1}\chi_{m2} & \cdots & t^{2m-1}\chi_{mm} \end{pmatrix} = t^{d_1+3d_2+\ldots+(2m-1)d_m} \det \chi. \tag{62}$$

We now compute the inverse of $S(t)$. First, the inverse of the principal part $\widehat{S}(t)$ is

$$\widehat{S}(t)^{-1}_{ij} = \frac{(\chi^{-1})_{ij}}{t^{i+j-1}},$$

as we readily check:

$$\sum_{\ell=1}^{m} \widehat{S}(t)^{-1}_{i\ell}\widehat{S}(t)_{\ell j} = \sum_{\ell=1}^{m}(\chi^{-1})_{i\ell}\chi_{\ell j}\frac{t^{\ell+j-1}}{t^{i+\ell-1}} = \sum_{\ell=1}^{m}(\chi^{-1})_{i\ell}\chi_{\ell j}t^{j-i} = \delta_{ij}.$$

The (block-wise) principal part of the inverse $S(t)^{-1}$ is equal to the inverse of the (block-wise) principal part of $S(t)$. Then we obtain, in blocks notation, for $i = 1, \ldots, m$

$$[S(t)^{-1}]_{ij} = \frac{(\chi^{-1})_{ij}}{t^{i+j-1}} + O\left(\frac{1}{t^{i+j-2}}\right).$$

Finally, by definition, $(S^\flat)^{-1} = [S^{-1}]_{11}$. Thus

$$S^\flat(t)^{-1} = \frac{(\chi^{-1})_{11}}{t} + O(1).$$

Therefore the reduced curve has a simple pole at $t = 0$, with a negative definite residue, and the proof is complete. □

Remark 6.24. Notice that, as a consequence of Eq. (62), the order of $\det S(t)$ at $t = 0$ is equal to the order of its principal part $\widehat{S}(t)$. Namely

$$\det S(t) \sim \det \widehat{S}(t) \sim t^{\mathcal{N}}, \qquad \mathcal{N} = \sum_{i=1}^{m}(2i-1)d_i. \tag{63}$$

Proof of the Theorem A. It is now clear that, in coordinates

$$\mathcal{Q}_\lambda(t) = \frac{d}{dt}S^\flat(t)^{-1},$$

as quadratic forms on $(\mathscr{D}_x, \langle\cdot,\cdot\rangle_\lambda)$ (see Eq. (60)). By Theorem 6.20, the map $t \mapsto S^\flat(t)^{-1}$ has a simple pole at $t = 0$, and its residue is a negative definite matrix. Then, $\mathcal{Q}_\lambda(t)$ has a second order pole at $t = 0$, and $t^2\mathcal{Q}_\lambda(t)$ can be extended smoothly also at $t = 0$. In particular, $\mathcal{I}_\lambda \doteq \lim_{t\to 0^+} t^2\mathcal{Q}_\lambda(t) > 0$.

Besides, by Lemma 6.22, $S_{11}(t) \leq S^\flat(t) < 0$, which implies $S^\flat(t)^{-1} \leq S_{11}(t)^{-1} < 0$. Then,

$$\mathcal{I}_\lambda = \lim_{t\to 0^+} t^2\frac{d}{dt}S^\flat(t)^{-1} = -\lim_{t\to 0^+} tS^\flat(t)^{-1} \geq -\lim_{t\to 0^+} tS_{11}(t)^{-1} = -\dot{S}_{11}(0)^{-1} > 0,$$

which, according to Remark 6.19, implies $\mathcal{I}_\lambda \geq \mathbb{I} > 0$ as operators on $\mathscr{D}_x$.

Finally, $\mathcal{Q}_\lambda(t)$ cannot have a term of order -1 in the Laurent expansion, which is tantamount to $\frac{d}{dt}\Big|_{t=0} t^2\mathcal{Q}_\lambda(t) = 0$. □

6.5 Proof of Theorem D

The purpose of this section is the proof of the main result of Section 5.6, namely a formula for the exponent of the asymptotic volume growth of geodesic homotheties.

Fix $x_0 \in M$ and let $\gamma : [0,1] \to M$ be the geodesic associated with the covector $\lambda \in T^*_{x_0}M$. Moreover, let J_λ be the associated Jacobi curve. As usual, we fix a Lagrangian splitting $T_\lambda(T^*M) = \mathcal{V}_\lambda \oplus \mathcal{H}_\lambda$, in terms of which $J_\lambda(t)$ is the graph of the map $S(t) : \mathcal{V}_\lambda \to \mathcal{H}_\lambda$. The reader can easily check that the statements that follow do not depend on the choice of the Lagrangian subspaces $\mathcal{H}_\lambda$. The following lemma relates $\mathcal{N}_\lambda$ with the Jacobi curve.

Lemma 6.25. *Assume that γ is ample, of step m, with growth vector $\mathcal{G}_\lambda = \{k_1, \dots, k_m\}$ (at $t = 0$). Then the order of $\det S(t)$ at $t = 0$ is*

$$\det S(t) \sim t^{\mathcal{N}_\lambda}, \qquad \mathcal{N}_\lambda = \sum_{i=1}^{m} (2i-1)(k_i - k_{i-1}).$$

If γ is not ample, the order of $\det S(t)$ at $t = 0$ is $+\infty$.

Proof. Indeed the order of $\det S(t)$ does not depend on the choice of the horizontal complement $\mathcal{H}_\lambda$ and Darboux coordinates. Then, for an ample curve, the statement is precisely Eq. (63), obtained in the proof of Theorem 6.20. Finally, if γ is not ample, the Taylor polynomial of arbitrary order of $S(t)$ is singular, thus the order of $\det S(t)$ at $t = 0$ is $+\infty$. □

Proof of Proposition 5.39. By Lemma 5.17, on a sub-Riemannian manifold there always exists at least an ample geodesic γ, with covector λ. Then it is well defined

$$\mathcal{N}_{x_0} = \min\{\mathcal{N}_\lambda | \, \lambda \in T^*_{x_0}M\} < +\infty.$$

Let $\lambda \in T^*_{x_0}M$ be a covector at which the minimum is attained. Then, by definition of ample Jacobi curve

$$\operatorname{rank}\{\dot{S}_\lambda(0), \ddot{S}_\lambda(0), \dots, S_\lambda^{(\mathcal{N}_{x_0})}(0)\} = n,$$

where $S_\lambda(t)$ is the matrix associated with the Jacobi curve $J_\lambda(t)$. We already observed (see Remark 6.16) that, for any Hamiltonian that is fibre-wise polynomial, $S_\lambda^{(i)}(0)$ is a polynomial function of the initial covector $\lambda \in T^*_{x_0}M$. Then $\operatorname{rank}\{\dot{S}_\lambda(0), \ddot{S}_\lambda(0), \dots, S_\lambda^{(\mathcal{N}_{x_0})}(0)\} < n$ on a closed Zariski subset $\mathcal{Z} \subset T^*_{x_0}M$, which has indeed zero measure. □

We are now ready to prove the main result of Section 5.6.

Proof of Theorem D. Without loss of generality, we can assume that Ω is contained in a single coordinate patch $\{x_i\}_{i=1}^n$. In terms of such coordinates, $\mu = e^a dx^1 \wedge \ldots \wedge dx^n$ and

$$\mu(\Omega_{x_0,t}) = \int_\Omega |\det(d_x\phi_t)| e^{a \circ \phi_t(x)} dx. \tag{65}$$

By smoothness, it is clear that the order of $\mu(\Omega_{x_0,t})$ at $t = 0$ is equal to the order of the map $t \mapsto \det(d_x\phi_t)$. In the following, $\mathcal{E}_{x_0} : T^*_{x_0}M \to M$ denotes the sub-Riemannian exponential map at time 1. Let us define $\Sigma^*_{x_0} \doteq \mathcal{E}_{x_0}^{-1}(\Sigma_{x_0}) \subset T^*_{x_0}M$. Indeed, if $\lambda \in \Sigma^*_{x_0}$, the associated geodesic $\gamma(t) = \mathcal{E}_{x_0}(t\lambda)$ is the unique one connecting x_0 with $x = \mathcal{E}_{x_0}(\lambda)$. We now compute the order of the map $t \mapsto \det(d_x\phi_t)$.

Lemma 6.26. *For every $x \in \Sigma_{x_0}$ the order of $t \mapsto \det(d_x\phi_t)$ is equal to $\mathcal{N}_\lambda$, where $\lambda = \mathcal{E}_{x_0}^{-1}(x)$.*

Proof. Recall that the order of a family of linear maps does not depend on the choice of the representative matrices. By Eq. (44),

$$d_x\phi_t = \pi_* \circ e_*^{(t-1)\vec{H}} \circ d_x^2\mathfrak{f}.$$

Let us focus on the linear map $e_*^{(t-1)\vec{H}} \circ d_x^2\mathfrak{f} : T_xM \to T_{\lambda(t)}(T^*M)$, where $\lambda(t) = e^{t\vec{H}}(\lambda)$ is the normal lift of γ. Let us choose a smooth family of Darboux bases $\{E_i|_{\lambda(t)}, F_i|_{\lambda(t)}\}_{i=1}^n$ of $T_{\lambda(t)}(T^*M)$, such that $\mathcal{V}_{\lambda(t)} = \operatorname{span}\{E_i|_{\lambda(t)}\}_{i=1}^n$ and $\mathcal{H}_{\lambda(t)} = \operatorname{span}\{F_i|_{\lambda(t)}\}_{i=1}^n$. Let us define the column vectors $E|_{\lambda(t)} \doteq (E_1|_{\lambda(t)}, \dots, E_n|_{\lambda(t)})^*$ and $F|_{\lambda(t)} \doteq (F_1|_{\lambda(t)}, \dots, F_n|_{\lambda(t)})^*$. Observe that the elements of $\pi_*F|_{\lambda(t)}$ are a smooth family of bases for $T_{\gamma(t)}M$. Then

$$e_*^{(t-1)\vec{H}} \circ d_x^2\mathfrak{f}(\pi_*F|_{\lambda(1)}) = A(t)E|_{\lambda(t)} + B(t)F|_{\lambda(t)}, \tag{66}$$

for some smooth families of $n \times n$ matrices $A(t)$ and $B(t)$. Then, by definition, the order of the map $t \mapsto \det(d_x\phi_t)$ is the order of $\det B(t)$ at $t = 0$. By acting with $e_*^{-t\vec{H}}$ in Eq. (66), we obtain

$$A(t)e_*^{-t\vec{H}}E|_{\lambda(t)} = e_*^{-\vec{H}} \circ d_x^2\mathfrak{f}(\pi_*F|_{\lambda(1)}) - B(t)e_*^{-t\vec{H}}F|_{\lambda(t)}. \tag{67}$$

Notice that $A(0)$ is nonsingular. Then, for t sufficiently close to 0, the l.h.s. of Eq. (67) is a smooth basis for the Jacobi curve J_λ. We rewrite the r.h.s. of Eq. (67) in terms of the fixed basis $\{E|_{\lambda(0)}, F|_{\lambda(0)}\}$. To this end, observe that

$$e_*^{-t\vec{H}}F|_{\lambda(t)} = C(t)E|_{\lambda(0)} + D(t)F|_{\lambda(0)},$$
$$e_*^{-\vec{H}} \circ d_x^2\mathfrak{f}(\pi_*F|_{\lambda(1)}) = GE|_{\lambda(0)}.$$

For some $n \times n$ smooth matrices $C(t), D(t), G$. Observe that $C(0) = 0$ and $D(t)$ is nonsingular for t sufficiently close to 0. Moreover, since $x \in \Sigma_{x_0}$ is a regular value for the sub-Riemannian exponential map $\mathcal{E}_{x_0} = \pi \circ e^{\vec{H}}$, G is nonsingular. Then

$$A(t)e_*^{-t\vec{H}}E|_{\lambda(t)} = [G - B(t)C(t)]E|_{\lambda(0)} - B(t)D(t)F|_{\lambda(0)}.$$

Therefore, the representative matrix of $J_\lambda(t)$ in terms of the basis $\{E|_{\lambda(0)}, F|_{\lambda(0)}\}$ is

$$S(t) = -[G - B(t)C(t)]^{-1}B(t)D(t), \qquad |t| < \varepsilon.$$

By the properties of the matrices G, $C(t)$ and $D(t)$ for sufficiently small t, $\det S(t) \sim \det B(t)$, and the two determinants have the same order. Then the statement follows from Lemma 6.25. ☐

By Proposition 5.39, $\mathcal{N}_\lambda = \mathcal{N}_{x_0}$ a.e. on $T_{x_0}^*M$. Then the order of $t \mapsto \det(d_x\phi_t)$ is equal to $\mathcal{N}_{x_0}$ up to a zero measure set on Σ_{x_0} and the statement of Theorem D follows from (65), since $\mu(\Omega) > 0$. ☐

7 Asymptotics of the Jacobi curve: equiregular case

In this section, we introduce a key technical tool, the so-called *canonical frame*, associated with a monotone, ample, equiregular curve in the Lagrange Grassmannian $L(\Sigma)$. This is a special moving frame in the symplectic space Σ which satisfies a set of differential equations encoding the dynamics of the underlying curve, which has been introduced for the first time in [36].

The main result of this section is an asymptotic formula for the curve, written in coordinates induced by the canonical frame. Finally, we exploit this result to prove Theorem B.

7.1 The canonical frame

Let $J(\cdot) \subset L(\Sigma)$ be an ample, monotone nonincreasing, equiregular curve of rank k. Suppose that its Young diagram D has k rows, of length n_a, for $a = 1, \dots, k$. Let us fix some terminology about the frames, indexed by the boxes of the Young diagram D. Each box of the diagram is labelled "ai", where $a = 1, \dots, k$ is the row index, and $i = 1, \dots, n_a$ is the progressive box number, starting from the left, in the specified row. Indeed n_a is the length of the a-th row, and $n_1 + \dots + n_k = n = \dim \Sigma$. Briefly, the notation $ai \in D$ denotes a generic box of the diagram.

From now on, we employ letters from the beginning of the alphabet $a, b, c, d, \dots$ for rows, and letters from the middle of the alphabet $i, j, h, k, \dots$ for the position of the box in the row. According to this notation, a frame $\{E_{ai}, F_{ai}\}_{ai \in D}$ for Σ is Darboux if, for any $ai, bj \in D$,

$$\sigma(E_{ai}, E_{bj}) = \sigma(F_{ai}, F_{bj}) = \sigma(E_{ai}, F_{bj}) - \delta_{ab}\delta_{ij} = 0,$$

where $\delta_{ab}\delta_{ij}$ is the Kronecker delta defined on $D \times D$.

7.1.1 A remark on the notation

Any Darboux frame indexed by the boxes of the Young diagram defines a Lagrangian splitting $\Sigma = \mathcal{V} \oplus \mathcal{H}$, where

$$\mathcal{V} = \operatorname{span}\{E_{ai}\}_{ai \in D}, \qquad \mathcal{H} = \operatorname{span}\{F_{ai}\}_{ai \in D}.$$

In the following, we deal with linear maps $S : \mathcal{V} \to \mathcal{H}$ (and their inverses), written in coordinates induced by the frame. The corresponding matrices have a peculiar block structure, associated with the Young diagram. The F_{bj} component of $S(E_{ai})$ is denoted by $S_{ab,ij}$. As a matrix, S can be naturally thought as a $k \times k$ block matrix. The block ab is a $n_a \times n_b$ matrix. This structure is the key of the calculations that follow, and we provide an example. Consider the Young diagram D, together with the "reflected" diagram $\overline{D}$ in Fig. 7. We labelled the boxes of

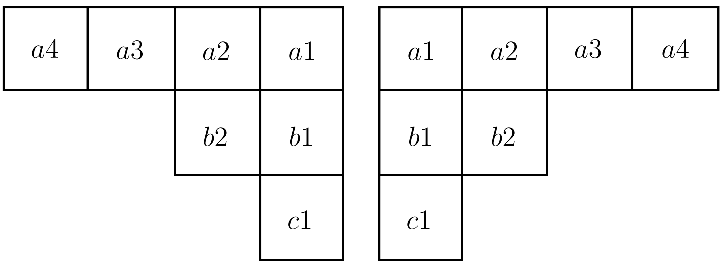


Figure 7: The Young diagrams $\overline{D}$ (left) and D (right).

the diagrams according to the convention introduced above. It is useful to think at each box of the diagram D as a one dimensional subspace of $\mathcal{V}$, and at each box of the diagram $\overline{D}$ as a one dimensional subspace of $\mathcal{H}$. Namely, the box $ai \in D$ corresponds to the subspace $\mathbb{R}E_{ai}$ (respectively, the box $bj \in \overline{D}$ corresponds to the subspace $\mathbb{R}F_{bj}$). Then the matrix S has the following block structure.

$$S = \begin{pmatrix} S_{aa} & S_{ab} & S_{ac} \\ S_{ba} & S_{bb} & S_{bc} \\ S_{ca} & S_{cb} & S_{cc} \end{pmatrix},$$

where each block is a matrix of the appropriate dimension, e.g. S_{ab} is a 4×2 matrix as explained pictorially in Fig. 8.

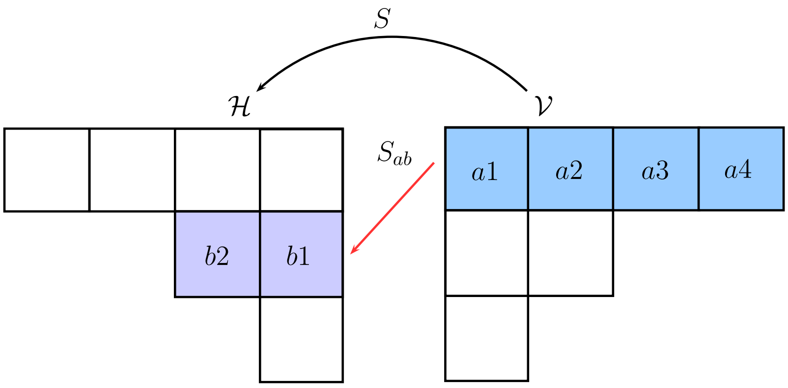


Figure 8: The 4×2 block S_{ab} of the map S.

Definition 7.1. A moving Darboux frame $\{E_{ai}(t), F_{ai}(t)\}_{ai\in D}$ is called a *canonical frame* of a monotonically nonincreasing curve $J(\cdot)$ with Young diagram D if $J(t) = \mathrm{span}\{E_{ai}(t)\}_{ai\in D}$ for any t, and there exists a one-parametric family of $n \times n$ symmetric matrices $R(t)$ such that the moving frame satisfies the *structural equations*

$$\begin{aligned}
&\dot{E}_{ai}(t) = E_{a(i-1)}(t), && a = 1,\dots,k, i = 2,\dots,n_a,\\
&\dot{E}_{a1}(t) = -F_{a1}(t), && a = 1,\dots,k,\\
&\dot{F}_{ai}(t) = \sum_{b=1}^{k}\sum_{j=1}^{n_b} R_{ab,ij}(t)E_{bj}(t) - F_{a(i+1)}(t), && a = 1,\dots,k, i = 1,\dots,n_a - 1,\\
&\dot{F}_{an_a}(t) = \sum_{b=1}^{k}\sum_{j=1}^{n_b} R_{ab,n_aj}(t)E_{bj}(t), && a = 1,\dots,k.
\end{aligned}$$

Notice that the matrix $R(t)$ is labelled according to the convention introduced above. The canonical frame for curves in a Lagrange Grassmannian has been introduced for the first time in [36]. In the aforementioned reference, the authors prove that such a frame always exists. Moreover, by requiring some algebraic condition on the family $R(t)$, the authors also proved that the canonical frame is unique up to orthogonal transformations which, in a sense, preserve the structure of the Young diagram. In this case, the family $R(t)$ (which is said to be *normal*) can be associated with a well defined operator which, together with the Young diagram D, completely classify the curve up to symplectic transformations (see also Section 7.2.2). At the end of this section, we also find a formula which connects the curvature operator $\mathcal{R}_\lambda$ of Definition 4.8 with some of the symplectic invariants $R(t)$ of the Jacobi curve (see Eq. (80)).

7.2 Main result

Fix a canonical frame, associated with $J(\cdot)$. Let $\mathcal{V} = \mathrm{span}\{E_{ai}(0)\}_{ai\in D}$ be the *vertical* subspace, and $\mathcal{H} = \mathrm{span}\{F_{bj}(0)\}_{bj\in D}$ be the *horizontal* subspace of Σ. Observe that $\mathcal{V} = J(0)$. The splitting $\Sigma = \mathcal{V} \oplus \mathcal{H}$ induces a coordinate chart in $L(\Sigma)$, such that $J(t) = \{(p, S(t)p) | \, p \in \mathbb{R}^n\}$. Recall that $S(0) = 0$ and, being the curve ample, $S(t)$ is invertible for $|t| < \varepsilon$ (see Lemma 6.3).

We introduce the constant $n \times n$ symmetric matrices, $\widehat{S}$, its inverse $\widehat{S}^{-1}$ and C, defined by

$$\widehat{S}_{ab,ij} = \frac{\delta_{ab}(-1)^{i+j-1}}{(i-1)!(j-1)!(i+j-1)},$$
$$\widehat{S}^{-1}_{ab,ij} = \frac{-\delta_{ab}}{i+j-1}\binom{n_a+i-1}{i-1}\binom{n_b+j-1}{j-1}\frac{(n_a)!(n_b)!}{(n_a-i)!(n_b-j)!},$$
$$C_{ab,ij} = \frac{(-1)^{i+j}(i+j+2)}{(i-1)!(j-1)!(i+j+1)(i+1)(j+1)}.$$

where, as usual, $a, b = 1, \dots, k$, $i = 1, \dots, n_a$, $j = 1, \dots, n_b$.

Theorem 7.2. *Let $J(\cdot)$ be a monotone, ample, equiregular curve of rank k, with a given Young diagram D with k rows, of length n_a, for $a = 1, \dots, k$. Then, for $|t| < \varepsilon$*

$$S_{ab,ij}(t) = \widehat{S}_{ab,ij}t^{i+j-1} - R_{ab,11}(0)C_{ab,ij}t^{i+j+1} + O(t^{i+j+2}). \tag{68}$$

Moreover, for $0 < |t| < \varepsilon$, the following asymptotic expansion holds for the inverse matrix:

$$S^{-1}_{ab,ij}(t) = \frac{\widehat{S}^{-1}_{ab,ij}}{t^{i+j-1}} + R_{ab,11}(0)\frac{(\widehat{S}^{-1}C\widehat{S}^{-1})_{ab,ij}}{t^{i+j-3}} + O\left(\frac{1}{t^{i+j-4}}\right). \tag{69}$$

Eqs. (68) and (69) highlight the block structure of the S matrix and its inverse at the leading orders. In particular, they give the leading order of the principal part of S^{-1} on the diagonal blocks (i.e. when $a = b$). The leading order terms of the diagonal blocks of S (and its inverse S^{-1}) only depend on the structure of the given Young diagram. Indeed the dependence on $R(t)$ appears in the higher order terms of Eqs. (68) and (69).

7.2.1 Restriction

At the end of this section, we apply Theorem 7.2 to compute the expansion of the family of operators $\mathcal{Q}_\lambda(t)$. According to the discussion that follows Eq. (59), we only need a block of the matrix $S(t)^{-1}$, namely $S^\flat(t)^{-1}$. As we explain below, it turns out that this corresponds to consider only the restriction of S^{-1} to the first columns of the Young diagram D and $\overline{D}$ (see Fig. 9). In terms of the frame $\{F_{a1}(0), E_{a1}(0)\}_{a=1}^k$, the map $S^\flat(t)^{-1}$ is a $k \times k$ matrix, with entries $S^\flat(t)^{-1}_{ab} = (S^{-1})_{ab,11}$. The following corollary is a consequence of Theorem 7.2, and gives the principal part of the aforementioned block.

Corollary 7.3. *Let $J(\cdot)$ be a monotone, ample, equiregular curve of rank k, with a given Young diagram D with k rows, of length n_a, for $a = 1, \dots, k$. Then, for $0 < |t| < \varepsilon$*

$$S^\flat(t)^{-1}_{ab} = -\delta_{ab}\frac{n_a^2}{t} + R_{ab,11}(0)\Omega(n_a, n_b)t + O(t^2), \tag{70}$$

where

$$\Omega(n_a, n_b) = \begin{cases} 0 & |n_a - n_b| \geq 2, \\ \frac{1}{4(n_a+n_b)} & |n_a - n_b| = 1, \\ \frac{n_a}{4n_a^2-1} & n_a = n_b. \end{cases} \tag{71}$$

Remark 7.4. If the Young diagram consists in a single column, with n boxes, $n_a = 1$ for all $a = 1, \dots, n$ and

$$S^\flat(t)^{-1}_{ab} = -\frac{\delta_{ab}}{t} + \frac{1}{3}R_{ab}(0)t + O(t^2).$$

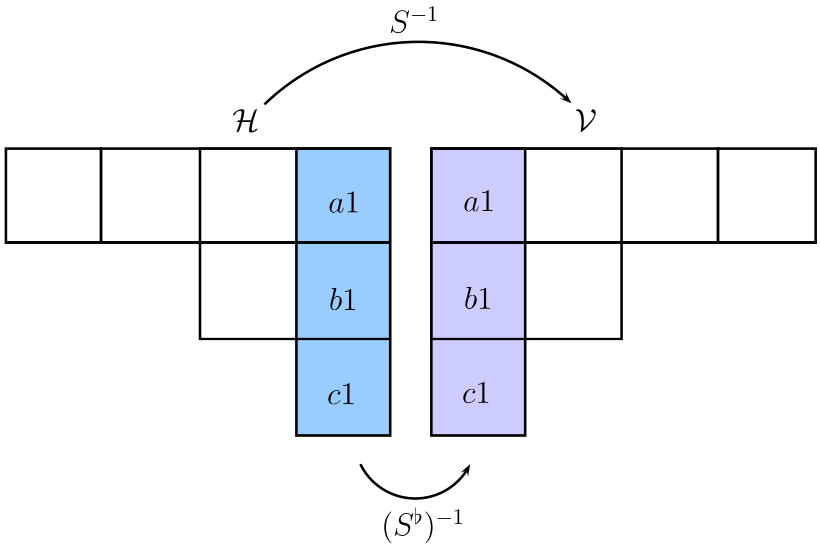


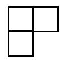
Figure 9: The block $S^\flat(t)^{-1}$ of the map $S(t)^{-1}$. Namely $(S^\flat)^{-1}_{ab} = S^{-1}_{ab,11}$.

7.2.2 A remark on the coefficients

Let us discuss the consequences of the peculiar form of the coefficients of Eq. (71). If $|n_a - n_b| \geq 2$, $\Omega(n_a, n_b) = 0$ and the corresponding $R_{ab,11}$ does not appear in the first order asymptotic. Nevertheless, if we assume that $R(t)$ is a *normal family* in the sense of [36], the "missing" entries are precisely the ones that vanish due to the assumptions on $R(t)$. It is natural to expect that some of the $R_{ab,ij}$ do not appear also in the higher orders of the asymptotic expansion. This may suggest the algebraic conditions to enforce on a generic family $R_{ab,ij}$ in order to obtain a truly canonical moving frame for the Jacobi curve.

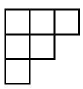
7.2.3 Examples

In this section we provide two practical examples of the asymptotic form of $S^\flat(t)^{-1}$. We suppress the subscript "$_{11}$" and the evaluation at $t = 0$ from each entry $R_{ab,11}(0)$.

A) Consider the 3-dimensional Jacobi curve with Young diagram: 

$$S^\flat(t)^{-1} = -\frac{1}{t}\begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} \frac{2}{5}R_{11} & \frac{1}{4}R_{12} \\ \frac{1}{4}R_{21} & R_{22} \end{pmatrix} t + O(t^2).$$

This corresponds to the case of the Jacobi curve associated with the geodesics of a 3D contact sub-Riemannian structure.

B) Consider the diagram: 

$$S^\flat(t)^{-1} = -\frac{1}{t}\begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} \frac{9}{35}R_{11} & \frac{3}{20}R_{12} & 0 \\ \frac{3}{20}R_{21} & \frac{2}{5}R_{22} & \frac{1}{4}R_{23} \\ 0 & \frac{1}{4}R_{23} & R_{33} \end{pmatrix} t + O(t^2).$$

This corresponds to the case of the Jacobi curve associated with a generic ample geodesics of a $(3, 6)$ Carnot group. In this example we can appreciate that some of the $R_{ab,11}$ do not appear in the linear term of the reduced matrix.

7.3 Proof of Theorem 7.2

The proof boils down to a careful manipulation of the structural equations, and matrices inversions. We prove Theorem 7.2 in three steps.

1. First, we consider the case of a rank 1 curve, and we assume $R(t) = 0$. In this case, the Young diagram is a single row and the structural equations are very simple. The canonical frame at time t is a polynomial in terms of the canonical frame at $t = 0$, and we compute explicitly the matrix $S(t)$ and its inverse.

2. Then, we consider a general rank 1 curve. The canonical frame at time t is no longer a polynomial in terms of the canonical frame at $t = 0$, but we can control the higher order terms. The non-vanishing $R(t)$ gives a contribution of higher order in t in each entry of the matrix $S(t)$ and its inverse.

3. Finally, we consider a general rank k curve. We show that, at the leading orders, we can "split" the curve in k rank 1 curves, and employ the results of the previous steps.

7.3.1 Rank 1 curve with vanishing $R(t)$

With these assumptions, the canonical frame is $\{E_i(t), F_i(t)\}_{i=1}^n$ (we suppress the row index, as D has a single row). The structural equations are

$$
\begin{aligned}
\dot{E}_1(t) &= -F_1(t), & \dot{F}_1(t) &= -F_2(t),\\
\dot{E}_2(t) &= E_1(t), & \dot{F}_2(t) &= -F_3(t),\\
&\vdots & &\vdots\\
\dot{E}_n(t) &= E_{n-1}(t), & \dot{F}_n(t) &= 0.
\end{aligned}
$$

Pictorially, in the double Young diagram the derivative shifts each element of the frame to the left by one box (see Fig. 10).

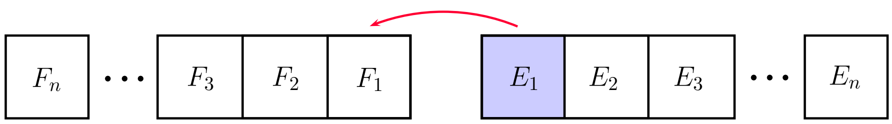


Figure 10: The action of the derivative on E_1.

Let $E(t) = (E_1, \dots, E_n)^*$ and $F(t) = (F_1, \dots, F_n)^*$, where each element is computed at t. Then there exist one parameter families of $n \times n$ matrices $A(t), B(t)$ such that

$$E(t) = A(t)E(0) + B(t)F(0).$$

$A(t)$ and $B(t)$ have monomial entries w.r.t. t. For $i, j = 1, \dots, n$

$$A_{ij}(t) = \frac{t^{i-j}}{(i-j)!} = \widehat{A}_{ij} t^{i-j}, \qquad (i \geq j), \tag{72}$$

$$B_{ij}(t) = \frac{(-1)^j t^{i+j-1}}{(i+j-1)!} = \widehat{B}_{ij} t^{i+j-1}. \tag{73}$$

Observe that A is a lower triangular matrix. A straightforward computation shows that

$$A^{-1}_{ij}(t) = \frac{(-1)^{i-j} t^{i-j}}{(i-j)!} = \widehat{A}^{-1}_{ij} t^{i-j}, \qquad (i \geq j). \tag{74}$$

Eqs. (72), (73) and (74) implicitly define the constant matrices $\widehat{A}$, $\widehat{B}$ and $\widehat{A}^{-1}$. The matrix $S(t)$ can be computed directly in terms of $A(t)$ and $B(t)$. Indeed $S(t) = A(t)^{-1}B(t)$.

Proposition 7.5 (Special case of Theorem 7.2)**.** *Let $J(\cdot)$ a curve of rank 1, with vanishing $R(t)$. The matrix $S(t)$, in terms of a canonical frame, is*

$$S(t)_{ij} = \frac{(-1)^{i+j-1}}{(i-1)!(j-1)!}\frac{t^{i+j-1}}{(i+j-1)} = \widehat{S}_{ij}t^{i+j-1}. \tag{75}$$

Its inverse is

$$S^{-1}(t)_{ij} = \frac{-1}{i+j-1}\binom{n+i-1}{i-1}\binom{n+j-1}{j-1}\frac{(n!)^2}{(n-i)!(n-j)!} = \frac{\widehat{S}^{-1}_{ij}}{t^{i+j-1}}. \tag{76}$$

As expected, $S(t)$ is symmetric, since the canonical frame is Darboux. The proof of Proposition 7.5 is a straightforward but long computation, which can be found in Appendix A. Eqs. (75) (76) implicitly define the constant matrix $\widehat{S}$ and its inverse $\widehat{S}^{-1}$. Observe that the entries of the latter depend explicitly on the dimension n.

7.3.2 General rank 1 curve

Now consider a general rank 1 curve. Its Young diagram is still a single row but, in general, $R(t) \neq 0$. As a consequence, the elements of the moving frame are no longer polynomial in t. However, we can still expand each $E_i(t)$ and obtain a Taylor approximation of its components w.r.t. the frame at $t = 0$. Each derivative at $t = 0$, up to order $i - 1$, is still a vertical vector

$$\frac{d^k E_i}{dt^k}(0) = E_{i-k}(0), \qquad k = 0, \dots, i-1.$$

The i-th derivative at $t = 0$ gives the lowest order horizontal term, i.e.

$$\frac{d^i E_i}{dt^i}(0) = -F_1(0).$$

Henceforth, each additional derivative, computed at $t = 0$, gives higher order horizontal terms, but also new vertical terms, depending on $R(t)$. Let us see a particular example, for $E_1(t)$. $\dot{E}_1(0) = -F_1(0)$, and $\ddot{E}_1(0) = F_2(0) - \sum_{j=1}^n R_{1j}(0)E_j(0)$ (see Fig. 11).

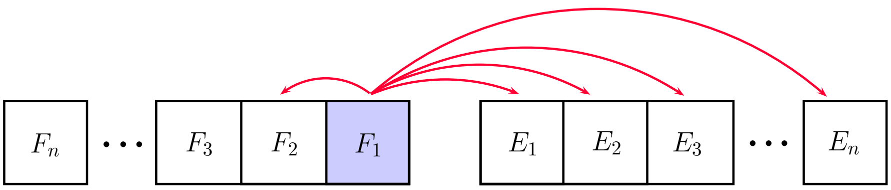


Figure 11: The action of the derivative of an horizontal element of the frame when $R \neq 0$.

Indeed $E_1(t)$ has a zeroth order term (w.r.t. the variable t) in the direction $E_1(0)$. The next term in the direction $E_1(0)$ is of order 2 or more. Besides, $E_1(t)$ has vanishing zeroth order term in each other vertical direction (i.e. $E_j(0)$, $j \neq 1$), but non vanishing components in each other vertical direction can appear, at orders greater or equal than 2. Let us turn to the horizontal components. $E_1(t)$ has a first order term in the direction $F_1(0)$. The next term in the same direction can appear only after two additional derivatives, or more. Therefore, the next term in the direction $F_1(0)$ is of order 3 or more in t. The "gaps" in the orders appearing in a given directions are precisely the key to the proof.

Let $E(t) = (E_1, \dots, E_n)^*$ and $F(t) = (F_1, \dots, F_n)^*$, where each element is computed at t. Then, as in the previous step, there exist one parameter families of $n \times n$ matrices $A(t)$, $B(t)$ such that

$$E(t) = A(t)E(0) + B(t)F(0).$$

The discussion above, and a careful application of the structural equations give us asymptotic formulae for the matrices $A(t)$ and $B(t)$. Let $\widehat{A}$ and $\widehat{B}$ defined as in Eqs. (72)-(73), corresponding to the case of a rank 1 curve with vanishing $R(t)$. Then, for $i, j = 1, \dots, n$

$$A(t)_{ij} = \widehat{A}_{ij} t^{i-j} - R_{1j}(0)\frac{t^{i+1}}{(i+1)!} + O(t^{i+2}),$$

$$B(t)_{ij} = \widehat{B}_{ij} t^{i+j-1} + R_{11}(0)\frac{(-1)^{j+1}t^{i+j+1}}{(i+j+1)!} + O(t^{i+j+2}).$$

The matrix A is no longer triangular, due to the presence of higher order terms in each entry. Besides, the order of the remainder grows only with the row index for $A(t)$ and it grows with both the column and row indices for $B(t)$. This reflects the different role played by the horizontal and vertical terms in the structural equations. We are now ready to consider the general case.

7.3.3 General rank k curve

The last step, which concludes the proof of the theorem, is built upon the previous cases. It is convenient to split a frame in subframes, relative to the rows of the Young diagram. For $a = 1, \dots, k$, the symbol E_a denotes the n_a-dimensional column vector

$$E_a = (E_{a1}, E_{a2}, \dots, E_{an_a})^* \in \Sigma^{n_a},$$

and analogously for F_a. Similarly, the symbol E denotes the n-dimensional column vector

$$E = (E_1, \dots, E_k)^* \in \Sigma^n,$$

and similarly for F. Once again, we express the elements of the Jacobi curves $E(t)$ in terms of the canonical frame at $t = 0$. With the notation introduced above

$$E(t) = A(t)E(0) + B(t)F(0).$$

This time, $A(t)$ and $B(t)$ are $k \times k$ block matrices, the ab block being a $n_a \times n_b$ matrix. For $a, b = 1, \dots, k$, $i = 1, \dots, n_a$, $j = 1, \dots, n_b$

$$A(t)_{ab,ij} = \delta_{ab}\widehat{A}_{ij} t^{i-j} - R_{ab,1j}(0)\frac{t^{i+1}}{(i+1)!} + O(t^{i+2}), \tag{77}$$

$$B(t)_{ab,ij} = \delta_{ab}\widehat{B}_{ij} t^{i+j-1} + R_{ab,11}(0)\frac{(-1)^{j+1}t^{i+j+1}}{(i+j+1)!} + O(t^{i+j+2}),$$

where, once again, the constant matrices $\widehat{A}$, $\widehat{B}$ correspond to the matrices defined for the rank 1 and $R(t) = 0$ case, of the appropriate dimension. Notice that we do not need explicitly the leading terms on the off-diagonal blocks. The knowledge of the leading terms on the diagonal blocks is sufficient for our purposes.

Remember that $S(t) = A(t)^{-1}B(t)$. In order to compute the inverse of $A(t)$ at the relevant order, we rewrite the matrix $A(t)$ as

$$A(t) = \widehat{A}(t) - M(t),$$

where $\widehat{A}(t)$ is the matrix corresponding to a rank k curve with vanishing $R(t)$, namely

$$\widehat{A}(t)_{ab,ij} = \delta_{ab}\widehat{A}_{ij}t^{i-j}, \qquad i=1,\ldots,n_a, \quad j=1,\ldots,n_b,$$

and, from Eq. (77), we get

$$M(t)_{ab,ij} = R_{ab,1j}(0)\frac{t^{i+1}}{(i+1)!} + O(t^{i+2}).$$

A standard inversion of the Neumann series leads to

$$A(t)^{-1} = \widehat{A}(t)^{-1} + \widehat{A}(t)^{-1}M(t)\widehat{A}(t)^{-1} + \sum_{n=2}^{\infty}\left(\widehat{A}(t)^{-1}M(t)\right)^n \widehat{A}(t)^{-1},$$

where the reminder term in the r.h.s. converges uniformly in the operator norm small t. Then, a long computation gives

$$A(t)^{-1}_{ab,ij} = \delta_{ab}\widehat{A}^{-1}_{ij}t^{i-j} - R_{ab,11}(0)\frac{(-1)^i t^{i+1}}{(i+1)(i-1)!} + O(t^{i+2}).$$

The matrix $S(t)$ can be computed explicitly, at the leading order, by the usual formula $S(t) = A(t)^{-1}B(t)$, and we obtain, for $a,b=1,\ldots,k$, $i=1,\ldots,n_a$, $j=1,\ldots,n_b$,

$$S(t)_{ab,ij} = \widehat{S}_{ab,ij}t^{i+j-1} - R_{ab,11}(0)C_{ab,ij}t^{i+j+1} + O(t^{i+j+2}),$$

where $\widehat{S}_{ab,ij} = \delta_{ab}\widehat{S}_{ij}$ of the appropriate dimension, and

$$C_{ab,ij} = \frac{(-1)^{i+j}(i+j+2)}{(i-1)!(j-1)!(i+j+1)(i+1)(j+1)}, \qquad i=1,\ldots,n_a, \quad j=1,\ldots,n_b.$$

The computation of $S(t)^{-1}$ follows from another inversion of the Neumann series, and a careful estimate of the remainder. We obtain

$$S^{-1}_{ab,ij}(t) = \frac{\widehat{S}^{-1}_{ab,ij}}{t^{i+j-1}} + R_{ab,11}(0)\frac{(\widehat{S}^{-1}C\widehat{S}^{-1})_{ab,ij}}{t^{i+j-3}} + O\left(\frac{1}{t^{i+j-4}}\right),$$

where

$$\widehat{S}^{-1}_{ab,ij} = \frac{-\delta_{ab}}{i+j-1}\binom{n_a+i-1}{i-1}\binom{n_b+j-1}{j-1}\frac{n_a!n_b!}{(n_a-i)!(n_b-j)!}.$$

This concludes the proof of Theorem 7.2. □

7.3.4 Proof of Corollary 7.3

Corollary 7.3 follows easily from Theorem 7.2. The only non trivial part is the explicit form of the coefficient $\Omega(n_a,n_b)$ in Eq. (70). By the results of Theorem 7.2,

$$\Omega(n_a,n_b) = (\widehat{S}^{-1}C\widehat{S}^{-1})_{ab,11}.$$

By replacing the explicit expression of $\widehat{S}^{-1}$ and C, the proof of Corollary 7.3 is reduced to the following lemma, which we prove in Appendix B.

Lemma 7.6. *Let $\Omega(n,m)$ be defined by the formula*

$$\Omega(n,m) = \frac{nm}{(n+1)(m+1)} \sum_{j=1}^{n}\sum_{i=1}^{m}(-1)^{i+j}\binom{n+i-1}{i-1}\binom{n+1}{i+1}\binom{m+j-1}{j-1}\binom{m+1}{j+1}\frac{i+j+2}{i+j+1}.$$

Then

$$\Omega(n,m) = \begin{cases} 0 & |n-m| \geq 2, \\ \frac{1}{4(n+m)} & |n-m| = 1, \\ \frac{n}{4n^2-1} & n = m. \end{cases}$$

The proof of Corollary 7.3 is now complete. □

7.4 Proof of Theorem B

In this section $J_\lambda : [0,T] \to L(T_\lambda(T^*M))$ is the Jacobi curve associated with an ample, equiregular geodesic γ, with initial covector $\lambda \in T_x^*M$. The next lemma shows that the projection of the horizontal part of the canonical frame corresponding to the first column of the Young diagram is an orthonormal basis for the Hamiltonian product on the distribution.

Lemma 7.7. *Let $X_a \doteq \pi_* F_{a1}(0) \in T_xM$. Then, the set $\{X_a\}_{a=1}^k$ is an orthonormal basis for $(\mathscr{D}_x, \langle\cdot|\cdot\rangle_\lambda)$.*

Proof. First, recall that $F_{a1}(0) = -\dot{E}_{a1}(0)$. Therefore $X_a = -\pi_*\dot{E}_{a1}(0)$. Then, by Eq. (58)

$$\langle X_a|X_b\rangle_\lambda = -\sigma(E_{a1}(0), \dot{E}_{b1}(0)) = \sigma(E_{a1}(0), F_{b1}(0)) = \delta_{ab}.$$

where we used the structural equations and the fact that the canonical frame is Darboux. □

We are now ready to prove one of the main results of Section 4.3, namely the one concerning the spectrum of the operator $\mathcal{I}_\lambda : \mathscr{D}_x \to \mathscr{D}_x$.

Proof of Theorem B. Actually, we prove something more: we use the basis $\{X_a\}_{a=1}^k$ obtained above to compute an asymptotic formula for the family $\mathcal{Q}_\lambda(t)$ introduced in Section 4.3.

Let $\Sigma = \mathcal{V}_\lambda \oplus \mathcal{H}_\lambda$ be the splitting induced by the canonical frame in $\Sigma = T_\lambda(T^*M)$. Let $S(t) : \mathcal{V}_\lambda \to \mathcal{H}_\lambda$ be the map which represents the Jacobi curve in terms of the canonical splitting. Then, by definition of Jacobi curve, it follows that, for any $v \in T_xM$ (see also Eq. (59)),

$$\langle \mathcal{Q}_\lambda(t)v|v\rangle_\lambda = \frac{d}{dt}\sigma(S(t)^{-1}\tilde{v}, \tilde{v}).$$

where $\tilde{v} \in \mathcal{H}_\lambda$ is the unique horizontal lift such that $\pi_*\tilde{v} = v$. In particular, if $v = \sum_{a=1}^k v_a X_a \in \mathscr{D}_x$, we have $\tilde{v} = \sum_{a=1}^k v_a F_{a1}(0)$. Thus,

$$\langle \mathcal{Q}_\lambda(t)v|v\rangle_\lambda = \frac{d}{dt}\sum_{a,b=1}^k S(t)^{-1}_{ab,11} v_a v_b = \frac{d}{dt}\sum_{a,b=1}^k S^\flat(t)^{-1}_{ab} v_a v_b.$$

By Corollary 7.3, we obtain finally the following asymptotic formula for $\mathcal{Q}_\lambda(t)$.

$$\langle \mathcal{Q}_\lambda(t)v|v\rangle_\lambda = \sum_{a,b=1}^k \left(\delta_{ab}\frac{n_a^2}{t^2} + R_{ab,11}(0)\Omega(n_a,n_b)\right) v_a v_b + O(t). \tag{78}$$

Equation (78), together with Lemma 7.7 imply that, for $a,b = 1,\ldots,k$,

$$\mathcal{I}_\lambda X_a = n_a^2 X_a, \tag{79}$$

$$\mathcal{R}_\lambda X_a = 3R_{ab,11}(0)\Omega(n_a,n_b)X_b. \tag{80}$$

Equation (79) completely characterizes the spectrum and the eigenvectors of $\mathcal{I}_\lambda$. □

Equation (80) is the anticipated formula which connects the curvature operator of Definition 4.8 with some of the symplectic invariants of the Jacobi curve, namely the elements of the matrix $R_{ab,ij}$ corresponding to the first column of the Young diagram.

8 Sub-Laplacian and Jacobi curves

Throughout this section, we assume M to be an equiregular sub-Riemannian manifold (that is, the rank of the distribution $\mathscr{D}$ is constant, equal to k). Nevertheless, most of the statements of this section hold true in the general case, by replacing the sub-Riemannian inner product on $\mathscr{D}$ with the Hamiltonian inner product. The final goal of this section is the proof of Theorem C, that is an asymptotic formula for the sub-Laplacian of the cost function. We start with a general discussion about the computation of the sub-Laplacian at a fixed point.

Let $f \in C^\infty(M)$, $x \in M$ and $\lambda = d_x f \in T_x^*M$. Moreover, let $X_1, \ldots, X_k$ be a local orthonormal frame for the sub-Riemannian structure. All our considerations are local, then we assume without loss of generality that the frame $X_1, \ldots, X_n$ is globally defined. Then, by Eq. (35), the sub-Laplacian associated with the volume form μ writes

$$\Delta_\mu f = \sum_{i=1}^k X_i^2(f) + \mathrm{div}_\mu(X_i)X_i(f).$$

As one can see, the sub-Laplacian is the sum of two terms. The first term, $\sum_{i=1}^k X_i^2(f)$, is a "sum of squares" which does not depend on the choice of the volume form. On the other hand, the second term, namely $\sum_{i=1}^k \mathrm{div}_\mu(X_i)X_i(f)$ depends on μ through the divergence operator. When x is a critical point for f, the second term vanishes, and the sub-Laplacian can be computed by taking the trace of the ordinary second differential of f (see Lemma 5.26). On the other hand, if x is non-critical, we need to compute both terms explicitly.

We start with the second term. Let $\theta_1, \ldots, \theta_n$ be the coframe dual to $X_1, \ldots, X_n$. Namely $\theta_i(X_j) = \delta_{ij}$. Then, there exists a smooth function $g \in C^\infty(M)$ such that $\mu = e^g \theta_1 \wedge \ldots \wedge \theta_n$. Finally, let $c_{ij}^k \in C^\infty(M)$ be the *structure functions* defined by $[X_i, X_j] = \sum_{k=1}^n c_{ij}^k X_k$. A standard computation using the definition of divergence gives

$$\mathrm{div}_\mu(X_i) = X_i(g) - \sum_{j=1}^n c_{ij}^j.$$

Thus, the second term of the sub-Laplacian is

$$\sum_{i=1}^k \mathrm{div}_\mu(X_i)X_i(f) = \langle \nabla f | \nabla g \rangle - \sum_{i=1}^k \sum_{j=1}^n c_{ij}^j X_i(f). \tag{81}$$

The first term of the sub-Laplacian can be computed through the generalized second differential introduced with Definition 6.5. Recall that the second differential at a non critical point x is a linear map $d_x^2 f : T_x M \to T_\lambda(T^*M)$.

8.1 Coordinate lift of a local frame

We introduce a special basis of $T_\lambda(T^*M)$, associated with a choice of the local frame $X_1, \ldots, X_n$, which is a powerful tool for explicit calculations. We define an associated frame on T^*M as follows. For $i = 1, \ldots, n$ let $h_i : T^*M \to \mathbb{R}$ be the linear-on-fibres function defined by $\lambda \mapsto h_i(\lambda) \doteq \langle \lambda, X_i \rangle$. The action of the derivations on T^*M is completely determined by the action on affine functions, namely functions $a \in C^\infty(T^*M)$ such that $a(\lambda) = \langle \lambda, Y \rangle + \pi^* g$ for some $Y \in \text{Vec}(M)$, $g \in C^\infty(M)$. Then, we define the *coordinate lift of a field* $X \in \text{Vec}(M)$ as the field $\widetilde{X} \in \text{Vec}(T^*M)$ such that $\widetilde{X}(h_i) = 0$ for $i = 1, \dots, n$ and $\widetilde{X}(\pi^* g) = X(g)$. This, together with Leibniz rule, characterize the action of $\widetilde{X}$ on affine functions, and then completely define $\widetilde{X}$. Indeed, by definition, $\pi_* \widetilde{X} = X$. On the other hand, we define the (vertical) fields ∂_{h_i} such that $\partial_{h_i}(\pi^* g) = 0$, and $\partial_{h_i}(h_j) = \delta_{ij}$. It is easy to check that $\{\partial_{h_i}, \widetilde{X}_i\}_{i=0}^n$ is a frame on T^*M. We call such a frame the *coordinate lifted frame*, and we employ the shorthand $\partial_i \doteq \partial_{h_i}$. Observe that, by the same procedure, we can define the coordinate lift of a vector $X \in T_xM$ (i.e. not necessarily a field) at any point $\lambda \in T_x^*M$.

Remark 8.1. Remember that we require $X_1, \dots, X_n$ to be *fields* (and not simple vectors in T_xM) in order to define the coordinate lift. In particular, the lift $\widetilde{X}|_\lambda \in T_\lambda(T^*M)$ depends on the germ at x of the chosen frame $X_1, \dots, X_n$. On the other hand, $\partial_i|_\lambda$ depends only on the value of $X_1, \dots, X_n$ at x.

Lemma 8.2. *Let $X \in T_xM$. In terms of a coordinate lifted frame,*

$$d_x^2 f(X) = \widetilde{X} + \sum_{i=1}^n X(X_i(f))\partial_i,$$

where $X(X_i(f))$ is understood to be computed at x and $\widetilde{X}, \partial_i \in T_\lambda(T^*M)$.

Proof. We explicitly compute the action of the vector $d_x^2 f(X) \in T_\lambda(T^*M)$ on affine functions. First, for any $g \in C^\infty(M)$, $d_x^2 f(X)(\pi^* g) = \pi_* \circ d_x^2 f(X)(g) = X(g)$. Moreover, $d_x^2 f(X)(h_i) = X(h_i \circ df) = X(\langle df, X_i \rangle) = X(X_i(f))$. □

Lemma 8.2, when applied to the vectors $X_1, \dots, X_k$, completely characterize the second order component of the sub-Laplacian, in terms of the second differential $d_x^2 f$.

8.2 Sub-Laplacian of the geodesic cost

Assume $f = c_t$, that is the geodesic cost associated with an ample, equiregular geodesic $\gamma : [0, T] \to M$. As usual, let $x = \gamma(0)$ be the initial point, $\lambda = d_x c_t$ the initial covector, and $J_\lambda(\cdot)$ the associated Jacobi curve, with Young diagram D. As discussed in Section 7, there is a class of preferred frames in $T_\lambda(T^*M)$, namely the canonical moving frame $\{E_{ai}(t), F_{ai}(t)\}_{ai \in D}$. In order to employ the results of Theorem 7.2 for the computation of Δc_t, we first relate the canonical frame with a coordinate lifted frame. As a first step, we need the following lemma, which is an extension of Lemma 7.7 along the geodesic.

Lemma 8.3. *Let $\{E_{ai}(t), F_{ai}(t)\}_{ai \in D}$ be a canonical moving frame for $J_\lambda(\cdot)$ and consider the following vector fields along γ:*

$$X_{ai}(t) \doteq \pi_* \circ e_*^{t\vec{H}} F_{ai}(t) \in T_{\gamma(t)}M, \qquad ai \in D.$$

The set $\{X_{ai}(t)\}_{ai \in D}$ is a basis for $T_{\gamma(t)}M$. Moreover $\{X_{a1}(t)\}_{a=1}^k$ is an orthonormal basis for $\mathscr{D}_{\gamma(t)}$ along the geodesic. Finally, consider any smooth extension of $\{X_{ai}(t)\}_{ai \in D}$ in a neighbourhood of γ, and the associated coordinate lifted frame. Then

$$E_{ai}(t) = e_*^{-t\vec{H}} \partial_{ai}|_{\lambda(t)},$$

Lemma 8.3 states that the projection of the horizontal elements of the canonical frame (the "F"s) corresponding to the first column of the Young diagram are an orthonormal frame for the sub-Riemannian distribution along the geodesic. Moreover, if we complete the frame with the projections of the other horizontal elements, and we introduce the associated coordinate lifted frame along the extremal $e^{t\vec{H}}(\lambda)$, the vertical elements of the canonical frame (the "E"s) have a simple expression. Observe that, according to Remark 8.1, the last statement of the lemma does not depend on the choice of the extension of the vectors $X_{ai}(t)$ in a neighbourhood of γ.

Proof. Assume first that the statement is true at $t = 0$. Then, let $0 < t < T$. Point (ii) of Proposition 6.11 gives the relation between the Jacobi curves "attached" at different points $\lambda(t) = e^{t\vec{H}}(\lambda)$ along the lift of γ. Namely

$$J_{\lambda(t)}(\cdot) = e^{t\vec{H}}_* J_\lambda(t + \cdot).$$

As a consequence of this, and the definition of canonical frame, if $\{E_{ai}(\cdot), F_{ai}(\cdot)\}_{ai\in D}$ is a canonical frame for the Jacobi curve $J_\lambda(\cdot)$, it follows that, for any fixed t,

$$\widetilde{E}_{ai}(\cdot) \doteq e^{t\vec{H}}_* E_{ai}(t + \cdot),$$
$$\widetilde{F}_{ai}(\cdot) \doteq e^{t\vec{H}}_* F_{ai}(t + \cdot),$$

is a canonical frame for the Jacobi curve $J_{\lambda(t)}(\cdot)$. In particular, $X_{ai}(t) = \pi_* \widetilde{F}_{ai}(0)$, and the statements now follow from the assumption that the lemma is true at the initial time of the Jacobi curve $J_{\lambda(t)}(\cdot)$.

Then, we only need to prove the statement at $t = 0$. For clarity, we suppress the explicit evaluation at $t = 0$. As usual, let $\mathcal{H}_\lambda = \operatorname{span}\{F_{ai}\}_{ai\in D}$ be the horizontal subspace and $\mathcal{V}_\lambda = \operatorname{span}\{E_{ai}\}_{ai\in D}$ be the vertical subspace. By definition of canonical frame, $T_\lambda(T^*M) = \mathcal{H}_\lambda \oplus \mathcal{V}_\lambda$. Since $\mathcal{V}_\lambda = \ker \pi_*$, and π_* is a submersion, $\pi_* \mathcal{H}_\lambda = T_x M$. Thus $\{X_{ai}\}_{ai\in D}$ is a basis for $T_x M$. By Lemma 7.7, the set $\{X_{a1}\}_{a=1}^k$ is an orthonormal frame for the Hamiltonian inner product $\langle \cdot | \cdot \rangle_\lambda$ which, in the sub-Riemannian case, does not depend on λ and coincides with the sub-Riemannian inner product (see Remark 4.7). Now, we show that $E_{ai} = \partial_{ai}|_\lambda$. Since the canonical frame is Darboux, this is equivalent to $\sigma(\partial_{ai}, F_{bj}) = \delta_{ab}\delta_{ij}$. Indeed, in terms of the coframe $\{\theta_{ai}\}_{ai\in D}$, dual to $\{X_{ai}\}_{ai\in D}$

$$\sigma = \sum_{ai\in D} dh_{ai} \wedge \pi^*\theta_{ai} + h_{ai}\pi^* d\theta_{ai}.$$

Therefore

$$\sigma(\partial_{ai}, F_{bj}) = \theta_{ai}(\pi_* F_{bj}) = \theta_{ai}(X_{bj}) = \delta_{ab}\delta_{ij}.$$

□

We now have all the tools we need in order to prove Theorem C, concerning the asymptotic behaviour of Δc_t.

Proof of Theorem C. The idea is the compute the "hard" term of Δc_t, namely the sum of squares term, through the coordinate representation of the Jacobi curve. By Lemma 8.2, written in terms of the frame $X_{ai} \doteq X_{ai}(0) = \pi_* F_{ai}(0)$ of $T_x M$, and its coordinate lift, we have

$$d^2_x c_t(X_\rho) = \widetilde{X}_\rho + \sum_{\nu\in D} X_\rho(X_\nu(c_t))\partial_\nu, \tag{82}$$

where we used greek letters as a shorthand for boxes of the Young diagram D. When ρ belongs to the first column of the Young diagram D, namely $\rho = a1$ (in this case, we simply write a), we have, as a consequence of Lemma 8.3 and the structural equations

$$F_a(0) = -\dot{E}_a(0) = -[\vec{H}, \partial_a] = \widetilde{X}_a + \sum_{\nu\in D} \left(\sum_{\kappa\in D} c^\kappa_{a\nu} h_\kappa + \sum_{b=1}^k h_b c^a_{b\nu} \right) \partial_\nu,$$

where everything is evaluated at λ. Therefore, from Eq. (82), we obtain

$$d_x^2 c_t(X_a) = F_a(0) + \sum_{\nu \in D} \left(X_a(X_\nu(c_t)) - \sum_{\kappa \in D} c_{a\nu}^{\kappa} h_\kappa - \sum_{b=1}^{k} h_b c_{b\nu}^{a} \right) E_\nu(0).$$

Recall that $S(t)^{-1} : \mathcal{H}_\lambda \to \mathcal{V}_\lambda$ is the matrix that represents the Jacobi curve in the coordinates induced by the canonical frame (at $t = 0$). More explicitly

$$d_x^2 c_t(X_\rho) = F_\rho(0) + \sum_{\nu \in D} S(t)^{-1}_{\rho\nu} E_\nu(0).$$

Moreover, since we restricted $d_x^2 c_t$ to elements of $\mathscr{D}_x$, we obtain

$$\sum_{a=1}^{k} X_a^2(c_t) = \sum_{a=1}^{k} S^\flat(t)^{-1}_{aa} + \sum_{a=1}^{k} \sum_{b=1}^{k} h_a c_{ab}^{b}. \tag{83}$$

Now observe that, if ρ *does not* belong to the first column of the Young diagram, we have

$$\dot{E}_\rho(0) = [\vec{H}, \partial_\rho] = \sum_{a=1}^{k} \sum_{\nu \in D} h_a c_{a\nu}^{\rho} E_\nu(0).$$

On the other hand, by the structural equations, $\dot{E}_\rho(0)$ is a vertical vector that does not have $E_\rho(0)$ components. Then, when ρ is not in the first column of D, $\sum_{a=1}^{k} h_a c_{a\rho}^{\rho} = 0$. Thus we rewrite Eq. (83) as

$$\sum_{a=1}^{k} X_a^2(c_t) = \sum_{a=1}^{k} S^\flat(t)^{-1}_{aa} + \sum_{a=1}^{k} \sum_{\rho \in D} h_a c_{a\rho}^{\rho}. \tag{84}$$

By taking the sum of Eq. (81) and Eq. (84), we obtain

$$\Delta_\mu c_t|_x = \sum_{a=1}^{k} S^\flat(t)^{-1}_{aa} + \langle \nabla_x c_t | \nabla_x g \rangle,$$

where we recall that the function g is implicitly defined (in a neighbourhood of γ) by $\mu = e^g \theta_1 \wedge \ldots \wedge \theta_n$. Remember that, at $x = \gamma(0)$, $\nabla_x c_t = \dot{\gamma}(0)$. Then

$$\Delta_\mu c_t|_x = \sum_{a=1}^{k} S^\flat(t)^{-1}_{aa} + \left.\frac{d}{dt}\right|_{t=0} g(\gamma(t)).$$

Remark 8.4. Observe that if $P_t \doteq X_1(t) \wedge \ldots \wedge X_n(t) \in \bigwedge^n T_{\gamma(t)} M$ is the parallelotope whose edges are the elements of the frame $\{X_i(t)\}_{i=1}^n$, then $g(\gamma(t)) = \log |\mu(P_t)|$, that is the logarithm of the volume of the parallelotope P_t.

Thus, by replacing the results of Corollary 7.3 about the asymptotics of the reduced Jacobi curve, we obtain

$$\Delta_\mu c_t|_x = -\frac{\operatorname{tr} \mathcal{I}_\lambda}{t} + \dot{g}(0) + \frac{1}{3}\operatorname{Ric}(\lambda) t + O(t^2),$$

where $\dot{g}(0) \doteq \left.\frac{d}{dt}\right|_{t=0} g(\gamma(t))$. Since $\mathfrak{f}_t = -t c_t$, we obtain

$$\Delta_\mu \mathfrak{f}_t|_x = \operatorname{tr} \mathcal{I}_\lambda - \dot{g}(0) t - \frac{1}{3}\operatorname{Ric}(\lambda) t^2 + O(t^3),$$

which is the sought expansion, valid for small t. □

8.2.1 Computation of the linear term

Recall that, for any equiregular smooth admissible curve $\gamma : [0,T] \to M$, the Lie derivative in the direction of the curve defines surjective linear maps

$$\mathcal{L}_{\mathsf{T}} : \mathscr{F}^i_{\gamma(t)}/\mathscr{F}^{i-1}_{\gamma(t)} \to \mathscr{F}^{i+1}_{\gamma(t)}/\mathscr{F}^i_{\gamma(t)}, \qquad i \geq 1,$$

as defined in Section 5.5. In particular, notice that $\mathcal{L}^i_{\mathsf{T}} : \mathscr{D}_{\gamma(t)} \to \mathscr{F}^{i+1}_{\gamma(t)}/\mathscr{F}^i_{\gamma(t)}$, for $i \geq 1$ is a well defined, surjective linear map from the distribution (see also point (iv) of Remark 3.4).

Lemma 8.5. *For $t \in [0,T]$, we recover the projections $X_{ai}(t) = e^{t\vec{H}}_* F_{ai}(t) \in T_{\gamma(t)}M$ as*

$$X_{ai}(t) = (-1)^{i-1}\mathcal{L}^{i-1}_{\mathsf{T}}(X_{a1}(t)) \bmod \mathscr{F}^{i-1}_{\gamma(t)}, \qquad a = 1,\dots,k, \quad i = 1,\dots,n_a.$$

Proof. Fix $a = 1,\dots,k$. For $i = 1$ the statement is trivial. Assume the statement to be true for $j \leq i$. Recall that we can see $F_{ai}|_{\lambda(t)} = e^{t\vec{H}}_* F_{ai}(t)$ as a field along the extremal $\lambda(t)$. Then, by the structural equations for the canonical frame, $X_{a(i+1)} = -\pi_*[\vec{H}, F_{ai}]$. A quick computation in terms of a coordinate lifted frame proves that

$$X_{a(i+1)}(t) = -[\mathsf{T}, X_{ai}]|_{\gamma(t)} \bmod \mathscr{F}^i_{\gamma(t)},$$

for an admissible extension T of $\dot{\gamma}$. Thus, by induction, we obtain the statement. □

Proof of Theorem 5.33. We consider equiregular distributions and ample geodesics γ that obey the growth condition

$$\dim \mathscr{F}^i_{\gamma(t)} = \dim \mathscr{D}^i, \qquad \forall\, i \geq 0. \tag{85}$$

We only need to compute explicitly the term $\dot{g}(0)$ of the asymptotic expansion in Theorem C. Recall that, according to the proof of Theorem C, the coefficient of the linear term is given by the following formula (see Remark 8.4)

$$\dot{g}(0) = \left.\frac{d}{dt}\right|_{t=0} \log|\mu(P_t)|,$$

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

$$P_t = \bigwedge_{ai\in D} X_{ai}(t). \tag{86}$$

By definition of canonical frame, Proposition 6.14, and the growth condition (85) we have that the elements $\{X_{ai}(t)\}_{ai\in D}$ are a frame along the curve $\gamma(t)$ adapted to the flag of the distribution. More precisely

$$\mathscr{D}^i_{\gamma(t)} = \operatorname{span}\{X_{aj}(t) \,|\, aj \in D,\, 1 \leq j \leq i\}.$$

By Lemma 8.5 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

$$P_t = \bigwedge_{i=1}^{m} \bigwedge_{a_i=1}^{d_i} X_{a_i i}(t) = \bigwedge_{i=1}^{m} \bigwedge_{a_i=1}^{d_i} \mathcal{L}^{i-1}_{\mathsf{T}}(X_{a_i 1}(t)).$$

Then, a standard linear algebra argument and the very definition of Popp's volume leads to

$$|\mu(P_t)| = \sqrt{\prod_{i=1}^{m} \det M_i(t)},$$

where the smooth families of operators $M_i(t)$, for $i = 1, \ldots, m$ are the one defined in Eq. (42). This, together with Eq. (86) completes the computation of the linear term of Theorem C for any ample geodesic satisfying the growth condition (85). □

Part III
Appendix

A Proof of Proposition 7.5

Proposition (Special case of Theorem 7.2)**.** *Let* $\Lambda(\cdot)$ *a Jacobi curve of rank* 1*, with vanishing* $R(t)$*. The matrix* S*, in terms of the canonical frame, is*

$$S_{ij}(t) = \frac{(-1)^{i+j-1}}{(i-1)!(j-1)!}\frac{t^{i+j-1}}{(i+j-1)} = \widehat{S}_{ij}t^{i+j-1}, \qquad i,j=1,\dots,n.$$

Its inverse is

$$S^{-1}(t)_{ij} = \frac{-1}{i+j-1}\binom{n+i-1}{i-1}\binom{n+j-1}{j-1}\frac{(n!)^2}{(n-i)!(n-j)!} = \frac{\widehat{S}^{-1}_{ij}}{t^{i+j-1}}, \qquad i,j=1,\dots,n.$$

Proof. From Eqs. (73) and (74), we obtain

$$\begin{aligned} S_{ij}(t) &= \sum_{k=1}^{n} A^{-1}_{ik}B_{kj} = \sum_{k=1}^{i}\frac{(-1)^{i-k}t^{i-k}}{(i-k)!}\frac{(-1)^j t^{k+j-1}}{(k+j-1)!} = (-1)^j t^{i+j-1}\sum_{k=1}^{i}\frac{(-1)^{i-k}}{(k+j-1)!(i-k)!} = \\ &= (-1)^j t^{i+j-1}\sum_{\ell=0}^{i-1}\frac{(-1)^\ell}{(i+j-1-\ell)!\ell!} = \frac{(-1)^j t^{i+j-1}}{(i+j-1)!}\sum_{\ell=0}^{i-1}\binom{i+j-1}{\ell}(-1)^\ell = \\ &= \frac{(-1)^{i+j-1}t^{i+j-1}}{(i+j-1)!}\binom{i+j-2}{j-1} = \frac{(-1)^{i+j-1}}{(i-1)!(j-1)!}\frac{t^{i+j-1}}{(i+j-1)}. \end{aligned}$$

By Cramer's rule, the inverse of $S(t)$ is

$$S^{-1}_{ij}(t) = \frac{(-1)^{i+j}\det\left[\frac{(-1)^{\ell+k-1}}{(\ell-1)!(k-1)!}\frac{t^{\ell+k-1}}{(\ell+k-1)}\right]_{\substack{\ell\neq j\\ k\neq i}}}{\det\left[\frac{(-1)^{\ell+k-1}}{(\ell-1)!(k-1)!}\frac{t^{\ell+k-1}}{(\ell+k-1)}\right]} = \frac{-(i-1)!(j-1)!\det\left[\frac{1}{\ell+k-1}\right]_{\substack{\ell\neq j\\ k\neq i}}}{t^{i+j-1}\det\left[\frac{1}{\ell+k-1}\right]}, \tag{87}$$

Now we compute the ratio of determinants in the last factor of Eq. (87). Consider a generic matrix of the form $H_{\ell k} = \frac{1}{x_\ell + x_k}$, for $\ell,k=1,\dots,n$. For fixed $i,j\in\{1,\dots,n\}$, we can express the determinant of H in terms of the the i,j-th minor, by rows and columns operations as follows. First, subtract the i-th column from each other column. We obtain a new matrix, H', whose i-th column is the same of H, while, for $k\neq i$

$$H'_{\ell k} = \frac{1}{x_\ell+y_k} - \frac{1}{x_\ell+y_i} = \frac{y_i-y_k}{(x_\ell+y_i)(x_\ell+y_k)}, \qquad \ell,k=1,\dots,n.$$

Indeed $\det H' = \det H$. Then, we collect the factor $\frac{1}{x_\ell+y_i}$ from each row, and the factor (y_i-y_k) from each column but the i-th. We obtain

$$\det\left[\frac{1}{x_\ell+x_k}\right] = \prod_{\ell=1}^{n}\frac{1}{x_\ell+y_i}\prod_{\substack{k=1\\k\neq i}}^{n}(y_i-y_k)\det\begin{bmatrix} \frac{1}{x_1+y_1} & \frac{1}{x_1+y_2} & \dots & 1 & \dots & \frac{1}{x_1+y_n} \\ \frac{1}{x_2+y_1} & \frac{1}{x_2+y_2} & \dots & 1 & \dots & \frac{1}{x_2+y_n} \\ \vdots & \vdots & & \vdots & & \vdots \\ \frac{1}{x_n+y_1} & \frac{1}{x_n+y_2} & \dots & 1 & \dots & \frac{1}{x_n+y_n} \end{bmatrix},$$

where the entries of the i-th column are equal to 1. Now, subtract the j-th row from each other row, but the j-th itself. Collect again the common factors. We obtain

$$\det\left[\frac{1}{x_\ell + x_k}\right] = (-1)^{i+j}\prod_{\ell=1}^{n}\frac{1}{x_\ell + y_i}\prod_{\substack{k=1\\k\neq i}}^{n}(y_i - y_k)\prod_{\substack{k=1\\k\neq i}}^{n}\frac{1}{x_j + y_k}\prod_{\substack{\ell=1\\\ell\neq j}}^{n}(x_j - x_\ell)\det\left[\frac{1}{x_\ell + x_k}\right]_{\substack{\ell\neq j\\k\neq i}}. \quad (88)$$

Now we apply the result of Eq. (88) to our case, i.e. $x_\ell = y_\ell = \ell - \frac{1}{2}$. Therefore we obtain

$$\frac{\det\left[\frac{1}{\ell+k-1}\right]_{\substack{\ell\neq j\\k\neq i}}}{\det\left[\frac{1}{\ell+k-1}\right]} = (-1)^{i+j}\prod_{\ell=1}^{n}(\ell+i-1)\prod_{\substack{k=1\\k\neq i}}^{n}\frac{1}{i-k}\prod_{\substack{k=1\\k\neq i}}^{n}(j+k-1)\prod_{\substack{\ell=1\\\ell\neq j}}^{n}\frac{1}{j-\ell} =$$

$$= \frac{1}{i+j-1}\frac{(n!)^2}{(i-1)!(j-1)!}\binom{i+n-1}{i-1}\binom{j+n-1}{j-1}. \quad (89)$$

Eq. (87) and Eq. (89), together, give the desired formula. □

B Proof of Lemma 7.6

Lemma. *Let*

$$\Omega(n,m) = \frac{nm}{(n+1)(m+1)}\sum_{j=1}^{n}\sum_{i=1}^{m}(-1)^{i+j}\binom{n+i-1}{i-1}\binom{n+1}{i+1}\binom{m+j-1}{j-1}\binom{m+1}{j+1}\frac{i+j+2}{i+j+1}.$$

Then

$$\Omega(n,m) = \begin{cases} 0 & |n-m| \geq 2, \\ \frac{1}{4(n+m)} & |n-m| = 1, \\ \frac{n}{4n^2-1} & n = m. \end{cases}$$

Proof. It is clear that $\Omega(n,m) = \Omega(m,n)$, then we can assume without loss of generality that $n \leq m$. The case $m = n = 1$ can be easily proved by a direct computation. Then, we also assume $m \geq 2$. Let us write $\Omega(n,m)$ in a more compact form. In order to do that, let $M(n,m)$ be the $n \times m$ matrix of components

$$M(n,m)_{ij} \doteq (-1)^{i+j}\frac{i+j+2}{i+j+1}, \qquad i = 1,\ldots,n, \quad j = 1,\ldots,m.$$

and let $v(m)$ be the m-dimensional column vector of components

$$v(m)_j = \frac{m}{m+1}\binom{m+1}{j+1}\binom{m+j-1}{j-1}, \qquad j = 1,\ldots,m.$$

Then

$$\Omega(n,m) = v(n)^* M(n,m) v(m).$$

Consider first the $i-th$ component of the n-dimensional vector $w(n,m) \doteq M(n,m)v(m)$, namely

$$w(n,m)_i = \sum_{j=1}^{m}(-1)^{i+j}\frac{i+j+2}{i+j+1}\frac{m}{m+1}\binom{m+1}{j+1}\binom{m+j-1}{j-1} = \frac{(-1)^i}{(m-1)!}\sum_{j=0}^{m}(-1)^j\binom{m}{j}Q_i(j),$$

where, for each $i = 1, \ldots, n$, $Q_i(j)$ is a rational function (in the variable j) defined by

$$Q_i(j) = \frac{(m+j-1)!}{(j-1)!(j+1)} \frac{i+j+2}{i+j+1} = j(j+2)(j+3)\ldots(j+m-1)\frac{i+j+2}{i+j+1}.$$

Notice that the factor $(j+1)$ does not appear (remember also that $m \geq 2$). The idea is to exploit the following beautiful identity.

Lemma B.1. *Let $m \geq 2$. Let $P(x)$ be any polynomial of degree smaller than m, then*

$$\sum_{j=0}^{m} (-1)^j \binom{m}{j} P(j) = 0.$$

Proof. It is sufficient to prove the statement for $P(x) = x^i$, with $0 \leq i < m$, since any polynomial of degree smaller than m is a linear combination of such monomials. By Newton's binomial formula, we have

$$(x-1)^m = (-1)^m \sum_{j=0}^{m} (-1)^j \binom{m}{j} x^j.$$

The result easily follows observing that any derivative of order strictly smaller than m, evaluated at $x = 1$ vanishes. □

We will see that, for many values of i, the denominator of $Q_i(j)$ factors the numerator, and then $Q_i(j)$ is actually a polynomial of degree $m-1$ in the variable j. Then we apply Lemma B.1 to show that $w(n,m)_i \neq 0$ only if $i = m-1, m$. In particular, since $w(n,m)$ is a n-dimensional vector, if $n \leq m-2$ then $w(n,m) = 0$ and $\Omega(n,m)$ vanishes too. Then we will explicitly compute the coefficient for $n = m-1$ and $n = m$.

Observe that, for each $i = 1, \ldots, n$, the numerator of $Q_i(j)$ is a polynomial of degree m in the variable j. Therefore there exists a polynomial $P_i(j)$ (of degree strictly smaller than m) and a number R_i such that

$$Q_i(j) = P_i(j) + \frac{R_i}{i+j+1}.$$

It is easy to compute the remainder. Observe that

$$R_i = -(i+j+1)P_i(j) + Q_i(j)(i+j+1).$$

Then, evaluating at $j = -i-1$, we obtain

$$R_i = \begin{cases} 0 & i = 1, 2, \ldots, m-2, \\ (-1)^{m-1} \dfrac{m!}{m-1} & i = m-1, \\ (-1)^{m-1} \dfrac{(m+1)!}{m} & i = m. \end{cases} \tag{90}$$

By Lemma B.1 we have

$$w(n,m)_i = \frac{(-1)^i}{(m-1)!} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \frac{R_i}{i+j+1},$$

which, by Eq. (90), is indeed zero if $i = 1, 2, \dots, m-2$. Then, since $\Omega(n,m) = v(n)^* w(n,m)$, we obtain after some straightforward computations the following formula:

$$\Omega(n,m) = \begin{cases} 0 & m-n > 2, \\ \binom{2m-3}{m-2} \sum_{j=0}^{m} \binom{m}{j} \frac{(-1)^j}{j+m} & n = m-1, \\ \binom{2m-2}{m-2} (m+1) \sum_{j=0}^{m} \binom{m}{j} \frac{(-1)^j}{j+m} - \binom{2m-1}{m-1} m \sum_{j=0}^{m} \binom{m}{j} \frac{(-1)^j}{j+m+1} & n = m. \end{cases} \tag{91}$$

In order to obtain the result, it only remains to compute the sums appearing in Eq. (91). Indeed these are of the form

$$S_k \doteq \sum_{j=0}^{m} (-1)^j \binom{m}{j} \frac{1}{j+k},$$

where k is a positive integer. We have the following, remarkable identity.

$$\sum_{j=0}^{m} (-1)^j \binom{m}{j} \frac{1}{j+k} = \frac{m!(k-1)!}{(m+k)!}. \tag{92}$$

By plugging Eq. (92) in Eq. (91) we obtain the result. Then we only need to prove Eq. (92). Indeed, for k a positive integer, let us define the following function

$$f_k(x) \doteq \sum_{j=0}^{m} (-1)^j \binom{m}{j} \frac{(-x)^{j+k}}{j+k}.$$

Indeed $S_k = f_k(-1)$. Let us compute the derivative of f.

$$f'(x) = -\sum_{j=0}^{m} (-1)^j \binom{m}{j} (-x)^{j+k-1} = (-1)^k x^{k-1} (1+x)^m.$$

where we used Newton's binomial formula. Then

$$S_k = f(-1) = (-1)^k \int_0^{-1} x^{k-1} (1+x)^m.$$

By integrating by parts $k-1$ times, we obtain the result

$$S_k = f(-1) = \frac{m!(k-1)!}{(m+k)!}.$$

□

C Proof of Theorem 2.19

The goal of this section is to prove Theorem 2.19 on the smoothness of the value function. All the relevant definitions can be found in Section 2. As a first step, we generalize the classical definition of conjugate points to our setting.

Definition C.1. Let $\gamma : [0, T] \to M$ be a strictly normal trajectory, such that $x_0 = \gamma(0)$ and $\gamma(t) = \mathcal{E}_{x_0}(t, \lambda_0)$. We say that $\gamma(t)$ is *conjugate with* x_0 *along* γ if λ_0 is a critical point for $\mathcal{E}_{x_0,t}$.

Observe that the relation "being conjugate with" is not reflexive in general. Indeed, even if $\gamma(t)$ is conjugate with x_0, there might not even exist an admissible curve starting from $\gamma(t)$ and ending at x_0.

We stress that, if γ is also abnormal, any $\gamma(t)$ is a critical value of the sub-Riemannian exponential map. Indeed, this is a consequence of the inclusion $\operatorname{Im} D_{\lambda_0}\mathcal{E}_{x_0,t} \subset \operatorname{Im} D_u E_{x_0,t} \neq T_{x_0}M$ for abnormal trajectories; being strongly normal is a necessary condition for the absence of critical values along a normal trajectory. Actually, a converse of this statement is true.

Proposition C.2. *Let $\gamma : [0,T] \to M$ be a strongly normal trajectory. Then, there exists an $\varepsilon > 0$ such that $\gamma(t)$ is not conjugate with $\gamma(0)$ along γ for all $t \in (0,\varepsilon)$.*

The proof of Proposition C.2 in the sub-Riemannian setting can be found in [3] and can be adapted to a general affine optimal control system. See also [9] for a more general approach.

We are now ready to prove Theorem 2.19 about smoothness of the value function which, for the reader's convenience, we restate here. Recall that $M' \subset M$ is the relatively compact subset chosen for the definition of the value function.

Theorem. *Let $\gamma : [0,T] \to M'$ be a strongly normal trajectory. Then there exists an $\varepsilon > 0$ and an open neighbourhood $U \subset (0,\varepsilon) \times M' \times M'$ such that:*

(i) $(t, \gamma(0), \gamma(t)) \in U$ *for all* $t \in (0,\varepsilon)$*,*

(ii) For any $(t,x,y) \in U$ there exists a unique (normal) minimizer of the cost functional J_t, among all the admissible curves that connect x with y in time t, contained in M',

(iii) The value function $(t,x,y) \mapsto S_t(x,y)$ is smooth on U.

Proof. We first prove the theorem in the case $M' = M$ compact. We need the following sufficient condition for optimality of normal trajectory. Let $a \in C^\infty(M)$. The graph of its differential is a smooth submanifold $\mathcal{L}_0 \doteq \{d_x a | \, x \in M\} \subset T^*M$, $\dim \mathcal{L}_0 = \dim M$. Translations of $\mathcal{L}_0$ by the flow of the Hamiltonian field $\mathcal{L}_\tau = e^{\tau \vec{H}}(\mathcal{L}_0)$ are also smooth submanifolds of the same dimension.

Lemma C.3 (see [9, Theorem 17.1]). *Assume that the restriction $\pi : \mathcal{L}_\tau \to M$ is a diffeomorphism for any $\tau \in [0,\varepsilon]$. Then, for any $\lambda_0 \in \mathcal{L}_0$, the normal trajectory*

$$\gamma(\tau) = \pi \circ e^{\tau \vec{H}}(\lambda_0), \qquad \tau \in [0,\varepsilon],$$

is a strict minimum of the cost functional J_ε among all admissible trajectories connecting $\gamma(0)$ with $\gamma(\varepsilon)$ in time ε.

Lemma C.3 is a sufficient condition for the optimality of a single normal trajectory. By building a suitable family of smooth functions $a \in C^\infty(M)$, one can prove that, for any sufficiently small compact set $K \subset T^*M$, we can find a $\varepsilon = \varepsilon(K) > 0$ sufficiently small such that, for any $\lambda_0 \in K$, and for any $t \leq \varepsilon$, the normal trajectory

$$\gamma(\tau) = \pi \circ e^{\tau \vec{H}}(\lambda_0), \qquad \tau \in [0,t], \qquad t \leq \varepsilon$$

is a strict minimum of the cost functional J_t among all admissible curves connecting $\gamma(0)$ with $\gamma(t)$ in time t.

We sketch the explicit construction of such a family. Let $K \subset T^*M$ sufficiently small such that it is contained in a trivial neighbourhood $\mathbb{R}^n \times U \subset T^*M$. Let (p,x) be coordinates on K induced by a choice of coordinates x on $O \subset M$. Then, consider the function $a : K \times O \to$

$\mathbb{R}$, defined in coordinates by $a(p_0, x_0; y) = p_0^* y$. Extend such a function to $a : K \times M \to \mathbb{R}$. For any $\lambda_0 \in K$, denote by $a^{(\lambda_0)} = a(\lambda_0; \cdot) \in C^\infty(M)$. Indeed, for $x_0 = \pi(\lambda_0)$, we have $\lambda_0 = d_{x_0} a^{(\lambda_0)}$. In other words we can recover any initial covector in K by taking the differential at x_0 of an appropriate element of the family. Therefore, let $\mathcal{L}_0^{(\lambda_0)} \doteq \{d_x a^{(\lambda_0)} | \, x \in M\}$, and $\mathcal{L}_\tau^{(\lambda_0)} \doteq e^{\tau \vec{H}}(\mathcal{L}_0^{(\lambda_0)})$. M is compact, then there exists $\varepsilon(K) = \sup\{\tau \geq 0 | \, \pi : \mathcal{L}_s^{(\lambda_0)} \to M$ is a diffeomorphism for all $s \in [0, \tau]$, $\lambda_0 \in K\} > 0$.

Let us go back to the proof. Set $x_0 = \gamma(0)$, and let $\gamma(t) = \mathcal{E}_{x_0}(t, \lambda_0)$. By Proposition C.2, we can assume that $\gamma(t)$ is not conjugate with $\gamma(0)$ along γ for all $t \in (0, \varepsilon)$. In particular, $D_{\lambda_0}\mathcal{E}_{x_0,t}$ has maximal rank for all $t \in (0, \varepsilon)$. Without loss of generality, assume that $\vec{H}$ is complete. Then, consider the map $\phi : \mathbb{R}^+ \times T^*M \to \mathbb{R}^+ \times M \times M$, defined by

$$\phi(t, \lambda) = (t, \pi(\lambda), \mathcal{E}_{\pi(\lambda)}(t, \lambda)).$$

The differential of ϕ, computed at (t, λ_0), is

$$D_{(t,\lambda_0)}\phi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \mathbb{I} & 0 \\ * & * & D_{\lambda_0}\mathcal{E}_{x_0,t} \end{pmatrix}, \qquad \forall t \in (0, \varepsilon),$$

which has maximal rank. Therefore, by the inverse function theorem, for each $t \in (0, \varepsilon)$, there exist an interval I_t and open sets W_t, U_t, V_t such that

$$t \in I_t \subset (0, \varepsilon), \qquad \lambda_0 \in W_t \subset T^*M, \qquad \gamma(0) \in U_t \subset M, \qquad \gamma(t) \in V_t \subset M,$$

and such that the restriction

$$\phi : I_t \times W_t \to I_t \times U_t \times V_t$$

is a smooth diffeomorphism. In particular, for any $(\tau, x, y) \in I_t \times U_t \times V_t$ there exists an unique initial covector $\lambda_0(\tau, x, y) \doteq \phi^{-1}(\tau, x, y)$ such that the corresponding normal trajectory starts from x and arrives at y in time τ, i.e. $\mathcal{E}_x(\tau, \lambda_0(\tau, x, y)) = y$. Moreover, we can choose $W_t \subset K$. Then such a normal trajectory is also a strict minimizer of J_τ among all the admissible curves connecting x with y in time τ. In particular, it is unique.

As a consequence of the smoothness of the local inverse, the value function $(t, x, y) \mapsto S_t(x, y)$ is smooth on each open set $I_t \times U_t \times V_t$. Indeed, for any $(\tau, x, y) \in I_t \times U_t \times V_t$, $S_t(x, y)$ is equal to the cost J_τ of the unique (normal) minimizer connecting x with y in time τ, namely

$$S_\tau(x, y) = \int_0^\tau L(\mathcal{E}_{x_0}(s, \lambda_0(\tau, x, y)), \bar{u}(e^{s\vec{H}}(\lambda_0(\tau, x, y))))ds, \qquad (\tau, x, y) \in I_t \times U_t \times V_t,$$

where $\bar{u} : T^*M \to \mathbb{R}^k$ is the smooth map which recovers the control associated with the lift on T^*M of the trajectory (see Theorem 2.17). Therefore the value function is smooth on $I_t \times U_t \times V_t$, as a composition of smooth functions. We conclude the proof by defining the open set

$$U \doteq \bigcup_{t \in (0,\varepsilon)} I_t \times U_t \times V_t \subset (0, \varepsilon) \times M \times M,$$

which is indeed open and contains $(t, \gamma(0), \gamma(t))$ for all $t \in (0, \varepsilon)$.

In the general case the proof follows the same lines, although the optimality of small segments of geodesics is only among all the trajectories not leaving M'. If we choose a different relatively compact $M'' \subset M$, we find a common ε such that the restriction to the interval $[0, \varepsilon]$ of all the normal geodesics with initial covector in K is a strict minimum of the cost function among all the admissible trajectories not leaving $M'' \cup M'$. Therefore, the value functions associated with the two different choices of the relatively compact subset agree on the intersection of the associated domains U.

□

D Proof of Proposition 5.14

The goal of this section is the proof of Proposition 5.14. Actually, we discuss a more general statement for the associated Hamiltonian system. All the relevant definitions can be found in Section 5.1.2.

Let $\lambda = (p, x) \in T^*\mathbb{R}^n = \mathbb{R}^{2n}$ any initial datum. Let ϕ^ε and $\widehat{\phi}$, respectively, the Hamiltonian flow of the ε-approximated system and of the nilpotent system, respectively. A priori, these local flows are defined in a neighbourhood of the initial condition and for small time which, in general, depend on ε. Notice that, by abuse of notation $\phi^0 = \widehat{\phi}$.

Lemma. *For $\varepsilon \geq 0$ sufficiently small, there exist common neighbourhood $I_0 \subset \mathbb{R}$ of 0 and $O_{\lambda_0} \subset \mathbb{R}^{2n}$ of λ_0, such that $\phi^\varepsilon : I_0 \times O_{\lambda_0} \to \mathbb{R}^{2n}$ is well defined. Moreover, $\phi^\varepsilon \to \widehat{\phi}$ in the C^∞ topology of uniform convergence of all derivatives on $I_0 \times O_{\lambda_0}$.*

Proof. Indeed, for any $\varepsilon \geq 0$, the Hamiltonian flow ϕ^ε is associated with the Cauchy problem

$$\dot{\lambda}(t) = H^\varepsilon(\lambda(t)), \qquad \lambda(0) = \lambda_0.$$

Moreover, ϕ^ε is well defined and smooth in a neighbourhood $I_0^\varepsilon \times O_{\lambda_0}^\varepsilon \subset \mathbb{R} \times \mathbb{R}^{2n}$ (that depends on ε). To find a common domain of definition, consider the associated Cauchy problem in $\mathbb{R}^{2n+1}$.

$$\begin{pmatrix} \dot{\lambda}(t) \\ \dot{\varepsilon}(t) \end{pmatrix} = \begin{pmatrix} H(\varepsilon(t), \lambda(t)) \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} \lambda(0) \\ \varepsilon(0) \end{pmatrix} = \begin{pmatrix} \lambda_0 \\ \varepsilon_0 \end{pmatrix}, \tag{93}$$

where $H(\varepsilon, \lambda) \doteq H^\varepsilon(\lambda)$ is smooth in both variables by construction. We denote by $\Phi(t; \lambda_0, \varepsilon_0)$ the flow associated with the Cauchy problem (93). By classical ODE theory, there exists a neighbourhood $I_0 \subset \mathbb{R}$ of 0 and $U_{\lambda_0,\varepsilon_0} \subset \mathbb{R}^{2n+1}$ of $(\lambda_0, \varepsilon_0)$ such that $\Phi : I_0 \times U_{\lambda_0,\varepsilon_0} \to \mathbb{R}^{2n+1}$ is well defined and smooth. Indeed $\Phi(t; \lambda_0, \varepsilon) = \phi^\varepsilon(t; \lambda_0)$ and $\Phi(t; \lambda_0, 0) = \widehat{\phi}(t; \lambda_0)$. Then, we can find an open neighbourhood $O_{\lambda_0} \subset \mathbb{R}^{2n}$ of λ_0 such that $O_{\lambda_0} \times [0, \delta] \subset U_{\lambda_0,0}$. Thus, the sought common domain of definition for all the ϕ^ε, with $0 \leq \varepsilon \leq \delta$, is $I_0 \times O_{\lambda_0}$.

Finally, Φ is smooth on $I_0 \times U_{\lambda_0,0}$. Then ϕ^ε (and all its derivatives) converge to $\widehat{\phi}$ (and all the corresponding derivatives) on $I_0 \times O_{\lambda_0}$. Up to restricting the domain of definition of Φ, we can always assume I_0 and O_λ to be compact, hence the convergence is also uniform. □

Without loss of generality, by homogeneity, we can always reduce to $I_0 = [0, T]$. Now Proposition 5.14 easily follows, since the exponential map is the projection of the Hamiltonian flow, restricted to the fiber $T_0^*\mathbb{R}^n$.

E Proof of Proposition 5.42

For the reader's convenience, we briefly recall the statement of Proposition 5.42. We refer to Section 5.7 for all the relevant definitions.

Proposition. *The function $C(t, s)$ is C^1 in a neighbourhood of the origin, but not C^2. In particular, the function $\partial_{ss}C(t, 0)$ is not continuous at the origin. However, the singularity at $t = 0$ is removable, and the following expansion holds, for $t > 0$*

$$\frac{\partial^2 C}{\partial s^2}(t, 0) = 1 + 3\sin^2(\phi_2 - \phi_1) + \frac{1}{2}[2h_{z,2}\sin(\phi_2 - \phi_1) - h_{z,1}\sin(2\phi_2 - 2\phi_1)]t - \\ - \frac{2}{15}h_{z,1}^2\sin^2(\phi_2 - \phi_1)t^2 + O(t^3).$$

If the geodesic γ_2 is chosen to be a straight line (i.e. $h_{z,2}=0$), then

$$\frac{\partial^2 C}{\partial s^2}(t,0) = 1 + 3\sin^2(\phi_2-\phi_1) - \frac{h_{z,1}}{2}\sin(2\phi_2-2\phi_1)t - \frac{2}{15}h_{z,1}^2\sin^2(\phi_2-\phi_1)t^2 + O(t^3). \quad (94)$$

where $\lambda_j = (ie^{i\phi_j}, h_{z,j}) = (-\sin\phi_j, \cos\phi_j, h_{z,h}) \in T_0^*M$ is the initial covector of the geodesic γ_j.

Proof. The proof is essentially a brute force computation. In the following, we show the relevant calculation to obtain the zeroth order term in Eq. (94), which is sufficient to prove the non-continuity of the function $t \mapsto \partial_{ss}C(t,0)$ at $t=0$. Indeed, since $C(0,s) = s^2/2$, we obtain $\partial_{ss}C(0,0)=1$, while from Eq. (94), $\lim_{t\to 0^+}\partial_{ss}C(0,s) = 1+3\sin^2(\phi_2-\phi_1)$.

For $i=1,2$, let $\gamma_i(\tau) = (w_i(\tau), z_i(\tau))$. Then

$$w_i(\tau) = \frac{e^{i\phi_i}}{a_i}\left(e^{ia_i\tau}-1\right) = ie^{i\phi_i}\tau - \frac{1}{2}a_i e^{i\phi_i}\tau^2 + O(\tau^3),$$
$$z_i(\tau) = \frac{a_i\tau - \sin(a_i\tau)}{2a_i^2} = O(\tau^3).$$

For $(t,s)\neq(0,0)$, dropping the subscripts from $R_{t,s}$ and $\xi_{t,s}$, we have

$$\partial_{tt}C(t,s) = \frac{1}{2}\partial_{tt}R^2\frac{\theta^2(\xi)}{\sin^2\theta(\xi)} + 4\partial_t R^2\theta(\xi)\partial_t\xi + 2R^2\dot\theta(\xi)(\partial_t\xi)^2 + 2R^2\theta(\xi)\partial_{tt}\xi =$$
$$= A_1(t,s)+A_2(t,s)+A_3(t,s)+A_4(t,s), \quad (95)$$

where A_i are the four addends of the upper line of Eq. (95). In order to compute Eq. (95), we employ the following calculations

$$R_{t,s}^2 = |w_2(s)-w_1(t)|^2,$$
$$\partial_t R_{t,s}^2 = \dot w_1(t)[\overline{w}_1(t)-\overline{w}_2(s)] + [w_1(t)-w_2(s)]\dot{\overline{w}}_1(t),$$
$$\partial_{tt}R_{t,s}^2 = \ddot w_1(t)[\overline{w}_1(t)-\overline{w}_2(s)] + 2|\dot w_1(t)|^2 + \ddot{\overline{w}}_1(t)[w_1(t)-w_2(s)],$$
$$Z_{t,s} = -z_1(t)+z_2(s)+\frac{1}{2}\Im(w_1(t)\overline{w}_2(s)),$$
$$\partial_t Z_{t,s} = -\dot z_1(t) + \frac{1}{2}\Im(\dot w_1(t)\overline{w}_2(s)),$$
$$\partial_{tt}Z_{t,s} = -\ddot z_1(t) + \frac{1}{2}\Im(\ddot w_1(t)\overline{w}_2(s)),$$
$$\xi_{t,s} = Z_{t,s}/R_{t,s}^2,$$
$$\partial_t\xi_{t,s} = \frac{\partial_t Z}{R^2} - \frac{Z}{R^4}\partial_t R^2,$$
$$\partial_{tt}\xi_{t,s} = \frac{\partial_{tt}Z}{R^2} - 2\frac{\partial_t Z}{R^4}\partial_t R^2 - \frac{Z}{R^4}\partial_{tt}R^2 + 4\frac{Z}{R^6}(\partial_t R^2)^2,$$

where $\Im$ is the imaginary part, the overline is the complex conjugate, and the dot is the derivative w.r.t. the argument. Moreover, the Taylor series for θ is

$$\theta(x) = 6x + O(x^3).$$

By computing everything at $t=0$, and then taking the limit $s\to 0$, we obtain

$$\lim_{s\to 0} A_1(0,s) = 1,$$
$$\lim_{s\to 0} A_2(0,s) = 0,$$
$$\lim_{s\to 0} A_3(0,s) = 3\sin^2(\phi_1-\phi_2),$$
$$\lim_{s\to 0} A_4(0,s) = 0,$$

therefore $\lim_{s\to 0}\partial_{tt}C(0,s) = 1 + 3\sin^2(\phi_1 - \phi_2)$, which is the zeroth order term of Eq. (94). The term arising from the addend $A_3(0,s)$ is responsible for the discontinuity of $\partial_{tt}C(0,s)$ at $s = 0$. The remaining terms can be obtained by taking expansions up to the fourth order of R^2, Z, θ, and replacing them in Eq. (95). □

Acknowledgements. We are grateful to Igor Zelenko and Paul Lee for very interesting stimulating discussions.

The first author has been supported by the grant of the Russian Federation for the state support of research, Agreement No 14 B25 31 0029.

The second author has been supported by the European Research Council, ERC StG 2009 "GeCoMethods", contract number 239748, by the ANR Project GCM, program "Blanche", project number NT09-504490.

The third author has been supported by INdAM (GDRE CONEDP) and Institut Henri Poincaré, Paris, where part of this research has been carried out.

References

[1] A. A. Agrachev. Geometry of optimal control problems and Hamiltonian systems. In *Nonlinear and optimal control theory*, volume 1932 of *Lecture Notes in Math.*, pages 1–59. Springer, Berlin, 2008.

[2] A. A. Agrachev. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.

[3] A. A. Agrachev, D. Barilari, and U. Boscain. Introduction to Riemannian and sub-Riemannian geometry (Lecture Notes), http://people.sissa.it/agrachev/agrachev_files/notes.html. 2012.

[4] A. A. Agrachev, D. Barilari, P. W. Y. Lee, and L. Rizzi. Curvature for contact sub-Riemannian manifolds. *In preparation.*

[5] A. A. Agrachev and R. V. Gamkrelidze. Feedback-invariant optimal control theory and differential geometry. I. Regular extremals. *J. Dynam. Control Systems*, 3(3):343–389, 1997.

[6] A. A. Agrachev, R. V. Gamkrelidze, and A. V. Sarychev. Local invariants of smooth control systems. *Acta Appl. Math.*, 14(3):191–237, 1989.

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

[8] A. A. Agrachev and P. Lee. Generalized Ricci Curvature Bounds for Three Dimensional Contact Subriemannian manifolds. *ArXiv preprint, arXiv:0903.2550*, Mar. 2009.

[9] 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.

[10] A. A. Agrachev and I. Zelenko. Geometry of Jacobi curves. I. *J. Dynam. Control Systems*, 8(1):93–140, 2002.

[11] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *ArXiv preprint, arXiv:1109.0222*, Sept. 2011.

[12] D. Barilari and L. Rizzi. A formula for Popp's volume in sub-Riemannian geometry. *Analysis and Geometry in Metric Spaces*, 2013.

[13] F. Baudoin and N. Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *ArXiv preprint, arXiv:1101.3590*, Jan. 2011.

[14] F. Baudoin and J. Wang. Curvature-dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds. *To appear on Potential Analysis*, 2013.

[15] A. Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian geometry*, volume 144 of *Progr. Math.*, pages 1–78. Birkhäuser, Basel, 1996.

[16] R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, and P. A. Griffiths. *Exterior differential systems*, volume 18 of *Mathematical Sciences Research Institute Publications*. Springer-Verlag, New York, 1991.

[17] J.-M. Coron. *Control and nonlinearity*, volume 136 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007.

[18] T. Gallouët. *Transport optimal : régularité et applications*. These, Ecole normale supérieure de lyon - ENS LYON, Dec. 2012.

[19] F. Jean. Control of Nonholonomic Systems and Sub-Riemannian Geometry. *ArXiv preprint, arXiv:1209.4387*, Sept. 2012.

[20] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *Int. Math. Res. Not. IMRN*, (13):2347–2373, 2009.

[21] T. Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[22] C. Li and I. Zelenko. Jacobi equations and comparison theorems for corank 1 sub-Riemannian structures with symmetries. *J. Geom. Phys.*, 61(4):781–807, 2011.

[23] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.

[24] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

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

[26] R. Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[27] S.-i. Ohta. On the curvature and heat flow on Hamiltonian systems. *ArXiv e-prints*, Aug. 2013.

[28] L. Pontryagin, V. Boltyanskij, R. Gamkrelidze, and E. Mishchenko. *Mathematische Theorie der optimalen Prozesse*. Oldenbourg, 1969.

[29] B. Riemann. Über die hypothesen, welche der geometrie zu grunde liegen. Göttengen, 1854.

[30] L. Rifford. *Sub-Riemannian Geometry and Optimal Transport*. Lecture Notes, 146 pp., 2012.

[31] L. Rifford. Ricci curvatures in Carnot groups. 2013. Math. Control Relat. Fields, to appear.

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

[33] K.-T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[34] K.-T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.

[35] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*.

[36] I. Zelenko and C. Li. Differential geometry of curves in Lagrange Grassmannians with given Young diagram. *Differential Geom. Appl.*, 27(6):723–742, 2009.

Andrei Agrachev

SISSA, Trieste, Italy and Steklov Mathematical Institute, Moscow, Russia.

E-mail: agrachev@sissa.it

Davide Barilari

CNRS, CMAP École Polytechnique and Équipe INRIA GECO Saclay Île-de-France, Paris, France.

E-mail: barilari@cmap.polytechnique.fr

Luca Rizzi
SISSA, Trieste, Italy.
E-mail: luca.rizzi@sissa.it