HAL
open science

Sub-Riemannian curvature in contact geometry

Andrei Agrachev, Davide Barilari, Luca Rizzi

To cite this version:

Andrei Agrachev, Davide Barilari, Luca Rizzi. Sub-Riemannian curvature in contact geometry. Journal of Geometric Analysis, 2016, 10.1007/s12220-016-9684-0 . hal-01160901v3

HAL Id: hal-01160901
https://hal.science/hal-01160901v3

Submitted on 21 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SUB-RIEMANNIAN CURVATURE IN CONTACT GEOMETRY

ANDREI AGRACHEV[1], DAVIDE BARILARI[2], AND LUCA RIZZI[3]

Abstract. We compare different notions of curvature on contact sub-Riemannian manifolds. In particular we introduce canonical curvatures as the coefficients of the sub-Riemannian Jacobi equation. The main result is that all these coefficients are encoded in the asymptotic expansion of the horizontal derivatives of the sub-Riemannian distance. We explicitly compute their expressions in terms of the standard tensors of contact geometry. As an application of these results, we obtain a sub-Riemannian version of the Bonnet-Myers theorem that applies to any contact manifold.

Contents

1. Introduction 1
2. Preliminaries 6
3. Jacobi fields revisited 9
4. Geodesic cost and its asymptotic 12
5. Bonnet-Myers type theorems 18
6. Computation of the canonical curvature 20
Appendix A. On the Yang-Mills condition 30
References 30

1. Introduction

The definition of general curvature-like invariants in sub-Riemannian geometry is a challenging and interesting topic, with many applications to the analysis, topology and geometry of these structures. In the general setting, there is no canonical connection à la Levi-Civita and thus the classical construction of the Riemann curvature tensor is not available. Nevertheless, in both the Riemannian and sub-Riemannian setting, the geodesic flow is a well defined Hamiltonian flow on the cotangent bundle: one can then generalize the classical construction of Jacobi fields and define the curvature as the invariants appearing in the Jacobi equation (i.e. invariants of the linearization of the geodesic flow). This approach has been extensively developed in [5,8,20,29]. This method, which leads to direct applications, has still some shortcomings since, even if these invariants could be a priori computed via an algorithm, it is extremely difficult to implement.

Another natural approach is to extract geometric invariants from the horizontal derivatives of the sub-Riemannian (squared) distance. This approach was developed in [4], where the authors introduce a family of symmetric operators canonically associated with a minimizing trajectory. Under some assumptions, this family admits an asymptotic expansion and each term of this expansion defines a metric invariant.

The goal of this paper is to revisit both constructions for contact manifolds, and to establish a bridge between the two approaches to curvature. The main result is how all invariants of

[1]SISSA, Italy, MI RAS and IM SB RAS, Russia
[2]Institut de Mathématiques de Jussieu-Paris Rive Gauche, UMR CNRS 7586 - Université Paris Diderot - Paris 7, Paris, France
[3]CMAP École Polytechnique and Équipe INRIA GECO Saclay Île-de-France, Paris, France
E-mail addresses: agrachev@sissa.it, davide.barilari@imj-prg.fr, luca.rizzi@cmap.polytechnique.fr.
Date: February 13, 2016.
2010 *Mathematics Subject Classification.* 53C17, 53C21, 53C22, 49N10.

the linearization of the geodesic flow are encoded in the asymptotic expansion of the squared distance. We explicitly compute these invariants in terms of the standard tensors of contact geometry (Tanno curvature, torsion and Tanno's tensor). Combining these formulas with the results of [9], we obtain a general Bonnet-Myers type result valid for every contact manifold.

1.1. **The contact setting.** Contact manifolds are an important sub-class of corank 1 sub-Riemannian structures that includes Yang-Mills type structures, Sasakian manifolds, (strongly pseudo-convex) CR structures, and the Heisenberg group.

More precisely, let M be a smooth manifold, with $\dim M = 2d+1$ and $\omega \in \Lambda^1 M$ be a one-form such that $\omega \wedge (d\omega)^d \neq 0$. The *contact distribution* is $\mathscr{D} := \ker \omega$. A sub-Riemannian structure on M is given by a smooth scalar product g on $\mathscr{D}$. In this case we say that (M, ω, g) is a *contact sub-Riemannian manifold*. This scalar product can be extended to the whole tangent bundle by requiring that the *Reeb vector field* X_0 is orthogonal to $\mathscr{D}$ and of norm one. The *contact endomorphism* $J : TM \to TM$ is defined by:

$$g(X, JY) = d\omega(X, Y), \qquad \forall\, X, Y \in \Gamma(TM).$$

By a classical result (see [14, Thm. 4.4]) there always exists a choice of the metric g on $\mathscr{D}$ such that $J^2|_{\mathscr{D}} = -\mathbb{I}$. The *Tanno's tensor* is the $(2,1)$ tensor field

$$Q(X, Y) := (\nabla_Y J)X, \qquad \forall\, X, Y \in \Gamma(TM),$$

where ∇ is the Tanno connection of the contact manifold (see Sec. 2.2 for precise definitions).

Horizontal curves (also called *Legendrian* in this setting), are curves γ such that $\dot\gamma(t) \in \mathscr{D}_{\gamma(t)}$. The length $\ell(\gamma)$ of an horizontal curve is well defined, and the *sub-Riemannian distance* is

$$\mathsf{d}(x, y) = \inf\{\ell(\gamma) \mid \gamma \text{ horizontal curve that joins } x \text{ with } y\}.$$

This turns a contact sub-Riemannian manifold in a metric space. Geodesics are horizontal curves such that sufficiently short segments realize the distance between their endpoints. They are projections of the integral lines of the geodesic flow on T^*M defined by the Hamiltonian function associated with the (sub-)Riemannian structure. In particular any geodesic is uniquely specified by its *initial covector* $\lambda \in T^*M$. In the Riemannian setting, this is nothing else than the geodesic flow seen on the cotangent space through the canonical isomorphism and the initial covector corresponds to the initial tangent vector of a geodesic. In the sub-Riemannian setting only the "dual viewpoint" survives.

1.2. **A family of operators.** For a fixed geodesic $\gamma(t)$, the *geodesic cost function* is

$$c_t(x) := -\frac{1}{2t}\mathsf{d}^2(x, \gamma(t)), \qquad x \in M,\, t > 0.$$

This function is smooth (as a function of t and x) for small $t > 0$ and x sufficiently close to $x_0 = \gamma(0)$. Moreover, its differential recovers the initial covector of the geodesic $\lambda = d_{x_0} c_t$, for all $t > 0$. As a consequence, the family of functions $\dot c_t := \frac{d}{dt} c_t$ has a critical point at x_0 and its second differential is a well defined quadratic form

$$d^2_{x_0} \dot c_t : T_{x_0} M \to \mathbb{R}.$$

Remark 1.1. The definition of $\dot c_t$ makes sense both in the Riemannian and sub-Riemannian setting. In both cases, for any unit-speed geodesic

$$\dot c_t(x) = \frac{1}{2}\|\dot\gamma(t) - W^t_{x,\gamma(t)}\|^2 - \frac{1}{2},$$

where $W^t_{x,\gamma(t)}$ is final tangent vector of the geodesic joining x with $\gamma(t)$ in time t (this is uniquely defined precisely where c_t is smooth). See Sec. 4.1.

Thanks to the sub-Riemannian metric, the restrictions $d^2_{x_0}\dot c_t|_{\mathscr{D}_{x_0}}$ define a family of symmetric operators $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$, by the formula

$$d^2_{x_0}\dot c_t(w) = g(\mathcal{Q}_\lambda(t)w, w), \qquad \forall\, t > 0, \quad \forall\, w \in \mathscr{D}_{x_0}.$$

The family $\mathcal{Q}_\lambda(t)$ is singular for $t \to 0$, but in this setting we have the following (Theorem 4.8).

Theorem 1.2. *The family $t \mapsto t^2\mathcal{Q}_\lambda(t)$ can be extended to a smooth family of operators on $\mathscr{D}_{x_0}$ for small $t \geq 0$, symmetric with respect to g. Moreover $\mathcal{I}_\lambda := \lim_{t\to 0^+} t^2\mathcal{Q}_\lambda(t) \geq \mathbb{I} > 0$.*

In particular, we have the Laurent expansion at $t = 0$

$$\mathcal{Q}_\lambda(t) = \frac{1}{t^2}\mathcal{I}_\lambda + \sum_{i=0}^{m} \mathcal{Q}_\lambda^{(i)} t^i + O(t^{m+1}). \tag{1}$$

Every operator $\mathcal{Q}_\lambda^{(i)} : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$, for $i \in \mathbb{N}$, is an invariant of the metric structure and, in this sense, $\mathcal{Q}_\lambda(t)$ can be thought of as a *generating function* for metric invariants. These operators, together with $\mathcal{I}_\lambda$, contain all the information on the germ of the structure along the fixed geodesic (clearly any modification of the structure that coincides with the given one on a neighborhood of the geodesic gives the same family of operators).

Remark 1.3. Applying this construction in the case of a Riemannian manifold, where $\mathscr{D}_{x_0} = T_{x_0}M$, one finds that $\mathcal{I}_\lambda = \mathbb{I}$ for any geodesic, and the operator $\mathcal{Q}_\lambda^{(0)}$ is the "directional" sectional curvature in the direction of the geodesic:

$$\mathcal{Q}_\lambda^{(0)} = \frac{1}{3}\overline{\mathrm{R}}(\cdot, \dot\gamma)\dot\gamma,$$

where $\dot\gamma$ is the initial vector of the geodesic and $\overline{\mathrm{R}}$ is the curvature of the Levi-Civita connection.

A similar construction has been carried out, in full generality, for the geometric structures arising from affine optimal control problems with Tonelli-type Lagrangian (see [4]). In that setting the geodesic is replaced by a minimizer of the optimization problem. Under generic assumptions on the extremal, the singularity of the family $\mathcal{Q}_\lambda(t)$ is controlled.

In the contact setting, due to the absence of abnormal minimizers, these assumptions are always satisfied for any non-trivial geodesic. One of the purposes of this paper is to provide a simpler and direct proof of the existence of the asymptotic.

1.3. **The singular term.** A first surprise is that, already in the contact case, $\mathcal{I}_\lambda$ is a non-trivial invariant. We obtain the complete characterization of the singular term of (1) (Theorem 4.10).

Theorem 1.4. *The symmetric operator $\mathcal{I}_\lambda : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ satisfies*

(i) $\operatorname{spec}\mathcal{I}_\lambda = \{1, 4\}$,
(ii) $\operatorname{tr}\mathcal{I}_\lambda = 2d + 3$.

More precisely, let $K_{\gamma(0)} \subset \mathscr{D}_{x_0}$ be the hyperplane $d\omega$-orthogonal to $\dot\gamma(0)$, that is

$$K_{\gamma(0)} := \{v \in \mathscr{D}_{\gamma(0)} \mid d\omega(v, \dot\gamma(0)) = 0\}.$$

Then $K_{\gamma(0)}$ is the eigenspace corresponding to eigenvalue 1 (and geometric multiplicity $2d-1$) and $K_{\gamma(0)}^\perp \cap \mathscr{D}_{x_0}$ is the eigenspace corresponding to eigenvalue 4 (and geometric multiplicity 1).

$\mathcal{I}_\lambda$ is a structural invariant that does not depend on the metric, but only on the fact that the distance function comes from a contact sub-Riemannian structure (operators $\mathcal{I}_\lambda$ coming from different metrics g on the same contact distribution have the same spectral invariants).

Remark 1.5. The trace $\mathcal{N} := \operatorname{tr}\mathcal{I}_\lambda = 2d + 3$, coincides with the *geodesic dimension* (of the contact sub-Riemannian structure) defined in [4, Sec. 5.6] (see also [25] for general metric measure spaces). If Ω_t is the geodesic homothety with center x_0 and ratio $t \in [0, 1]$ of a measurable, bounded set $\Omega \subset M$ with positive measure, we have the following *asymptotic measure contraction property*:

$$\lim_{t\to 0} \frac{\log \mu(\Omega_t)}{\log t} = \mathcal{N},$$

for any smooth measure μ. i.e. $\mathcal{N}$ is the order of vanishing of $\mu(\Omega_t)$ for $t \to 0$.

The regular terms of the asymptotic 1 are *curvature-like* invariants. The next main result is the explicit relation of the operators $\mathcal{Q}_\lambda^{(i)}$ with the symplectic invariants of the linearization of the geodesic flow, that we introduce now.

1.4. **Linearized Hamiltonian flow.** In the (sub-)Riemannian setting, the geodesic flow $\phi_t : T^*M \to T^*M$ is generated by the Hamiltonian function $H \in C^\infty(T^*M)$ (the co-metric of the sub-Riemannian structure). More precisely, if σ is the canonical symplectic structure on T^*M, then the Hamiltonian vector field $\vec{H}$ is defined by $\sigma(\cdot, \vec{H}) = dH$, and $\phi_t = e^{t\vec{H}}$. Integral lines $\lambda(t) = e^{t\vec{H}}(\lambda)$ of the geodesic flow are usually called *extremals*. Geodesics are then projections of non-trivial extremals $\gamma(t) = e^{t\vec{H}}(\lambda)$ (non-trivial is equivalent to $H(\lambda) \neq 0$). For any fixed extremal, and initial datum $\xi \in T_\lambda(T^*M)$, we define the vector field along the extremal

$$X_\xi(t) := e_*^{t\vec{H}}\xi \in T_{\lambda(t)}(T^*M).$$

The set of these vector fields is a $2n$-dimensional vector space that coincides with the space of solutions of the *(sub-)Riemannian Jacobi equation*

$$\dot{X} = 0,$$

where $\dot{X} := \mathcal{L}_{\vec{H}}X$ denotes the Lie derivative in the direction of $\vec{H}$.

Pick a Darboux frame $\{E_i(t), F_i(t)\}_{i=1}^n$ along $\lambda(t)$ (to fix ideas, one can think at the canonical basis $\{\partial_{p_i}|_{\lambda(t)}, \partial_{q_i}|_{\lambda(t)}\}$ induced by a choice of coordinates $(q_1, \dots, q_n)$ on M). In terms of this frame, the Jacobi field $X(t)$ has components $(p(t), q(t)) \in \mathbb{R}^{2n}$:

$$X(t) = \sum_{i=1}^n p_i(t)E_i(t) + q_i(t)F_i(t).$$

In the Riemannian case, one can choose a *canonical Darboux frame* (satisfying special equations, related with parallel transport) such that the components $(p(t), q(t)) \in \mathbb{R}^{2n}$ satisfy

$$\dot{p} = -R(t)q, \qquad \dot{q} = p, \tag{2}$$

for some smooth family of symmetric matrices $R(t)$. It turns out this class of frames is defined up to a constant orthogonal transformation and $R(t)$ is the matrix representing the *curvature operator* $\mathfrak{R}_{\gamma(t)} : T_{\gamma(t)}M \to T_{\gamma(t)}M$ in the direction of the geodesic, in terms of a parallel transported frame. In particular

$$\mathfrak{R}_{\gamma(t)}v = \overline{\mathrm{R}}(v, \dot{\gamma}(t))\dot{\gamma}(t), \qquad \forall\, v \in T_{\gamma(t)}M.$$

where $\overline{\mathrm{R}}$ is the Riemannian curvature tensor. From Eq. (2), it follows the classical classical Jacobi equation written in terms of a parallel transported frame: $\ddot{q} + R(t)q = 0$. In this language, the Riemann curvature arises as a set of invariants of the linearization of the geodesic flow.

1.4.1. *Canonical curvatures.* Analogously, in the sub-Riemannian setting, we might look for a *canonical Darboux frame* such that the Jacobi has, in coordinates, the simplest possible form. This analysis begun in [5] and has been completed in a very general setting in [29]. The "normal form" of the Jacobi equation defines a series of invariants of the sub-Riemannian structure along the given geodesic:

- A *canonical splitting* of the tangent space along the geodesic:

$$T_{\gamma(t)}M = \bigoplus_\alpha S^\alpha_{\gamma(t)},$$

  where α runs over a set of indices that depends on the germ of the sub-Riemannian structure along the geodesic.
- A *canonical curvature operator* $\mathfrak{R}_{\gamma(t)} : T_{\gamma(t)}M \to T_{\gamma(t)}M$.

The operator $\mathfrak{R}_{\gamma(t)}$ (and its partial traces) is the correct object to bound to obtain sectional-type (and Ricci-type) comparison theorems in the general sub-Riemannian setting, as it controls the evolution of the Jacobian of the exponential map (see for instance [9]).

The general formulation is complicated, since the very structure of the normal form depends on the type of sub-Riemannian structure. In Section 3 we give an ad-hoc presentation for the contact case. In particular, we prove the following (Theorem 6.1).

Theorem 1.6 (Canonical splitting)**.** *Let $\gamma(t)$ be a unit speed geodesic of a contact sub-Riemannian structure with initial covector λ. Then the canonical splitting is given by*

$$\begin{aligned} S^a_{\gamma(t)} &:= \operatorname{span}\{X_0 - 2Q(\dot\gamma,\dot\gamma) - h_0\dot\gamma\}, && \dim S^a_{\gamma(t)} = 1,\\ S^b_{\gamma(t)} &:= \operatorname{span}\{J\dot\gamma\}, && \dim S^b_{\gamma(t)} = 1,\\ S^c_{\gamma(t)} &:= J\dot\gamma^\perp \cap \mathscr{D}_{\gamma(t)}, && \dim S^c_{\gamma(t)} = 2d-1, \end{aligned}$$

where X_0 is the Reeb field, Q is the Tanno tensor, $h_0 = \langle\lambda, X_0\rangle$ and everything is computed along the extremal. Indeed $\mathscr{D}_{\gamma(t)} = S^b_{\gamma(t)} \oplus S^c_{\gamma(t)}$ and $\dot\gamma(t) \in S^c_{\gamma(t)}$.

The above theorem follows from the explicit computation of the canonical frame. Moreover, we obtain an explicit expression for $\mathfrak{R}_{\gamma(t)}$ (Theorems 6.2-6.3).

1.5. **Relation between the two approaches.** The main goal of this paper is to find the relation between the curvature-like objects introduced so far: on one hand, the canonical curvature operator $\mathfrak{R}_{\gamma(t)}$ that we have just defined, on the other hand, the invariants $\mathcal{Q}^{(i)}_\lambda$, for $i \geq 0$, defined by the asymptotics (1) (both associated with a given geodesic with initial covector λ).

Notice that $\mathfrak{R}$ is a canonical operator defined on a space of dimension n, while the operators $\mathcal{Q}^{(i)}$ are defined on a space of dimension k, equal to the dimension of the distribution. In the Riemannian case $k = n$, and a dimensional argument suggests that the first element $\mathcal{Q}^{(0)}$ should "contain" all the canonical curvatures. Indeed we have

$$\mathfrak{R}_{\gamma(t)} = 3\mathcal{Q}^{(0)}_{\lambda(t)} = \overline{\mathrm{R}}(\cdot,\dot\gamma(t))\dot\gamma(t).$$

In the sub-Riemannian setting the relation is much more complicated and in general the first element $\mathcal{Q}^{(0)}$ recovers only *a part* of the canonical curvature $\mathfrak{R}$. More precisely, as proved in [4]

$$\mathfrak{R}_{\gamma(t)}\big|_{\mathscr{D}_{\gamma(t)}} = 3\mathcal{Q}^{(0)}_{\lambda(t)}.$$

It turns out that, on contact structures, we recover the whole $\mathfrak{R}$ by computing the higher order invariants $\mathcal{Q}^{(0)}$, $\mathcal{Q}^{(1)}$ and $\mathcal{Q}^{(2)}$ (see Theorem 4.12 for the explicit relations).

1.6. **Comparison theorems.** The restriction of the curvature operator on the invariant subspaces $S^\alpha_{\gamma(t)}$ is denoted $\mathfrak{R}^{\alpha\beta}_{\gamma(t)}$, for $\alpha,\beta = a,b,c$. We have more than one Ricci curvature, one partial trace for each subspace

$$\mathfrak{Ric}^\alpha_{\gamma(t)} := \operatorname{tr}\left(\mathfrak{R}^{\alpha\alpha}_{\gamma(t)} : S^\alpha_{\gamma(t)} \to S^\alpha_{\gamma(t)}\right), \qquad \alpha = a,b,c.$$

In the Riemannian case, we only have one subspace (the whole tangent space) and only one average: the classical Ricci curvature.

In [9], under suitable conditions on the canonical curvature of a given sub-Riemannian geodesic, we obtained bounds on the first conjugate time along the geodesic and, in particular, Bonnet-Myers type results. In Sec. 5 we first apply the results of [9] to contact structures. It is interesting to express these conditions in terms of the classical tensors of the contact structure (Tanno's tensor, curvature and torsion). With the explicit expressions for the canonical curvature of Sec. 6, we obtain the following results (Theorem 5.4).

Theorem 1.7. *Consider a complete, contact structure of dimension $2d+1$, with $d > 1$. Assume that there exists constants $\kappa_1 > \kappa_2 \geq 0$ such that, for any horizontal unit vector X*

$$\operatorname{Ric}(X) - R(X,JX,JX,X) \geq (2d-2)\kappa_1, \qquad \|Q(X,X)\|^2 \leq (2d-2)\kappa_2. \tag{3}$$

Then the manifold is compact with sub-Riemannian diameter not greater than $\pi/\sqrt{\kappa_1-\kappa_2}$, and the fundamental group is finite.

We stress that the curvatures appearing in (3) are computed w.r.t. Tanno connection. This generalizes the results for Sasakian structures obtained in [18, 19].

Finally, we obtain the following corollary for (strongly pseudo-convex) CR manifolds (that is, for $Q = 0$). Notice that this condition is strictly weaker than Sasakian. Observe that in the CR case, Tanno's curvature coincides with the classical Tanaka-Webster curvature.

Corollary 1.8. *Consider a complete, (strongly pseudo-convex) CR structure of dimension* $2d+1$*, with* $d > 1$*, such that, for any horizontal unit vector* X

$$\mathrm{Ric}(X) - R(X, JX, JX, X) \geq (2d-2)\kappa > 0.$$

Then the manifold is compact with sub-Riemannian diameter not greater than $\pi/\sqrt{\kappa}$*, and the fundamental group is finite.*

Besides the above references, other Bonnet-Myers type results are found in the literature, proved with different techniques and for different sub-Riemannian structures. For example, with heat semigroup approaches: for Yang-Mills type structures with transverse symmetries [11] and Riemannian foliations with totally geodesic leaves [12]. With direct computation of the second variation formula: for 3D contact CR structures [27] and for general 3D contact ones [17]. Finally, with Riccati comparison techniques for 3D contact [6], 3-Sasakian [26] and for any sub-Riemannian structure [9].

A compactness result for contact structures is also obtained [13] by applying the classical Bonnet-Myers theorem to a suitable Riemannian extension of the metric.

1.7. **Final comments and open questions.** In the contact setting, the invariants $\mathcal{Q}^{(i)}$ for $i = 0, 1, 2$ recover the whole canonical curvature operator $\mathfrak{R}$. It is natural to conjecture that, in the general case, there exists $N \in \mathbb{N}$ (depending on the sub-Riemannian structure) such that the invariants $\mathcal{Q}^{(i)}$ for $i = 0, \dots, N$, recover the whole canonical curvature $\mathfrak{R}$. Already in the contact case the relation is complicated due to the high number of derivatives required.

Finally, the comparison results obtained above (and in the general sub-Riemannian setting in [9]) rely on the explicit computations of $\mathfrak{R}$ and its traces. In view of the relation we obtained between $\mathfrak{R}$ and the operators $\mathcal{Q}^{(i)}$, it is natural to ask whether it is possible to obtain comparison theorems in terms of suitable C^N-bounds on the geodesic cost (for some finite N depending on the sub-Riemannian structure).

Acknowledgements. 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. This research has also been supported by the European Research Council, ERC StG 2009 "GeCoMethods," contract number 239748 and by the iCODE institute, research project of the Idex Paris-Saclay. The second and third authors were supported by the Grant ANR-15-CE40-0018 of the ANR. The third author was supported by the SMAI project "BOUM". This research, benefited from the support of the "FMJH Program Gaspard Monge in optimization and operation research", and from the support to this program from EDF.

We thank the Chinese University of Hong Kong, where part of this project has been carried out. We are grateful to Paul W.Y. Lee for his kind invitation and his precious contribution. We thank Fabrice Baudoin for useful discussions on Yang-Mills structures. We also warmly thank Stefan Ivanov for pointing out a redundant assumption in a previous version of Theorem 1.7.

2. Preliminaries

We recall some basic facts in sub-Riemannian geometry. We refer to [3] for further details.

Let M be a smooth, connected manifold of dimension $n \geq 3$. A sub-Riemannian structure on M is a pair $(\mathscr{D}, g)$ where $\mathscr{D}$ is a smooth vector distribution of constant rank $k \leq n$ satisfying the *Hörmander condition* (i.e. $\mathrm{Lie}_x\mathscr{D} = T_xM$, $\forall x \in M$) and g is a smooth Riemannian metric on $\mathscr{D}$. A Lipschitz continuous curve $\gamma : [0, T] \to M$ is *horizontal* (or *admissible*) if $\dot\gamma(t) \in \mathscr{D}_{\gamma(t)}$ for a.e. $t \in [0, T]$. Given a horizontal curve $\gamma : [0, T] \to M$, the *length of* γ is

$$\ell(\gamma) = \int_0^T \|\dot\gamma(t)\| dt,$$

where $\|\cdot\|$ is the norm induced by g. The *sub-Riemannian distance* is the function

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

The Rashevsky-Chow theorem (see [15, 22]) guarantees the finiteness and the continuity of $\mathsf{d} : M \times M \to \mathbb{R}$ with respect to the topology of M. The space of vector fields (resp. horizontal vector fields) on M is denoted by $\Gamma(TM)$ (resp. $\Gamma(\mathscr{D})$).

Locally, the pair $(\mathscr{D}, g)$ can be given by assigning a set of k smooth vector fields that span $\mathscr{D}$, orthonormal for g. In this case, the set $\{X_1, \dots, X_k\}$ is called a *local orthonormal frame* for the sub-Riemannian structure.

A sub-Riemannian *geodesic* is an admissible curve $\gamma : [0, T] \to M$ such that $\|\dot\gamma(t)\|$ is constant and for every $t \in [0, T]$ there exists an interval $t_1 < t < t_2$ such that the restriction $\gamma|_{[t_1,t_2]}$ realizes the distance between its endpoints. The length of a geodesic is invariant by reparametrization of the latter. Geodesics with $\|\dot\gamma(t)\| = 1$ are called *length parametrized* (or of *unit speed*). A sub-Riemannian manifold is *complete* if (M, d) is complete as a metric space.

With a sub-Riemannian structure we associate the Hamiltonian function $H \in C^\infty(T^*M)$

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

for any local orthonormal frame $X_1, \dots, X_k$, where $\langle \lambda, \cdot\rangle$ denotes the action of the covector λ on vectors. Let σ be the canonical symplectic form on T^*M. For any function $a \in C^\infty(T^*M)$, the associated Hamiltonian vector field $\vec a$ is defined by the formula $da = \sigma(\cdot, \vec a)$. For $i = 1, \dots, k$ let $h_i \in C^\infty(T^*M)$ be the linear-on-fibers functions $h_i(\lambda) := \langle \lambda, X_i\rangle$. Clearly,

$$H = \frac{1}{2}\sum_{i=1}^{k} h_i^2, \qquad \text{and} \qquad \vec H = \sum_{i=1}^{k} h_i \vec h_i.$$

2.1. **Contact sub-Riemannian structures.** In this paper we consider contact structures defined as follows. Let M be a smooth manifold of odd dimension $\dim M = 2d + 1$, and let $\omega \in \Lambda^1 M$ a one-form such that $\omega \wedge (d\omega)^d \neq 0$. In particular $\omega \wedge (d\omega)^d$ defines a volume form and M is orientable. The *contact distribution* is $\mathscr{D} := \ker \omega$. Indeed $d\omega|_{\mathscr{D}}$ is non-degenerate. The choice of a scalar product g on $\mathscr{D}$ defines a sub-Riemannian structure on M. In this case we say that (M, ω, g) is a *contact sub-Riemannian manifold*.

Trajectories minimizing the distance between two points are solutions of first-order necessary conditions for optimality, given by a weak version of the Pontryagin Maximum Principle (see [21], or [3] for an elementary proof). We denote by $\pi : T^*M \to M$ the standard bundle projection.

Theorem 2.1. *Let M be a contact sub-Riemannian manifold and let $\gamma : [0, T] \to M$ be a sub-Riemannian geodesic. Then there exists a Lipschitz curve $\lambda : [0, T] \to T^*M$, such that $\pi \circ \lambda = \gamma$ and for all $t \in [0, T]$:*

$$\dot\lambda(t) = \vec H(\lambda(t)). \tag{4}$$

If $\lambda : [0, T] \to M$ is a curve satisfying (4), it is called a *normal extremal*. It is well known that if $\lambda(t)$ is a normal extremal, then $\lambda(t)$ is smooth and its projection $\gamma(t) := \pi(\lambda(t))$ is a smooth geodesic. Let us recall that this characterization is not complete on a general sub-Riemannian manifold, since also abnormal extremals can appear.

Let $\lambda(t) = e^{t\vec H}(\lambda_0)$ be the integral curve of the Hamiltonian vector field $\vec H$ starting from λ_0. The sub-Riemannian *exponential map* (from x_0) is

$$\mathcal{E}_{x_0} : T^*_{x_0}M \to M, \qquad \mathcal{E}_{x_0}(\lambda_0) := \pi(e^{\vec H}(\lambda_0)).$$

Unit speed geodesics correspond to initial covectors such that $H(\lambda_0) = 1/2$.

Definition 2.2. Let $\gamma(t) = \mathcal{E}_{x_0}(t\lambda_0)$ be a normal unit speed geodesic. We say that $t > 0$ is a *conjugate time* (along the geodesic) if $t\lambda_0$ is a critical point of $\mathcal{E}_{x_0}$. The *first conjugate time* is $t_*(\lambda_0) = \inf\{t > 0 \mid t\lambda_0 \text{ is a critical point of } \mathcal{E}_{x_0}\}$.

On a contact sub-Riemannian manifold or, in general, when there are no non-trivial abnormal extremals, the first conjugate time is separated from zero (see for instance [1, 3]) and, after its first conjugate time, geodesics lose local optimality.

2.2. **Contact geometry.** Given a contact manifold (M, ω), the *Reeb vector field* X_0 is the unique vector field satisfying $\omega(X_0) = 1$ and $d\omega(X_0, \cdot) = 0$. Clearly X_0 is transverse to $\mathscr{D}$. We can extend the sub-Riemannian metric g on $\mathscr{D}$ to a global Riemannian structure (that we denote with the same symbol g) by promoting X_0 to an unit vector orthogonal to $\mathscr{D}$.

We define the *contact endomorphism* $J : TM \to TM$ by:

$$g(X, JY) = d\omega(X, Y), \qquad \forall\, X, Y \in \Gamma(TM).$$

Clearly J is skew-symmetric w.r.t. to g. By a classical result (see [14, Thm. 4.4]) there always exists a choice of the metric g on $\mathscr{D}$ such that $J^2 = -\mathbb{I}$ on $\mathscr{D}$ and $J(X_0) = 0$, or equivalently

$$J^2 = -\mathbb{I} + \omega \otimes X_0.$$

In this case, g is said to be *compatible* with the contact structure and (M, ω, g, J) is usually referred to as a *contact metric structure* or a *contact Riemannian structure*. In this paper, we always assume the metric g to be compatible.

Theorem 2.3 (Tanno connection, [14, 28])**.** *There exists a unique linear connection ∇ on (M, ω, g, J) such that*

(i) $\nabla\omega = 0$,
(ii) $\nabla X_0 = 0$,
(iii) $\nabla g = 0$,
(iv) $T(X, Y) = d\omega(X, Y)X_0$ *for any* $X, Y \in \Gamma(\mathscr{D})$,
(v) $T(X_0, JX) = -JT(X_0, X)$ *for any vector field* $X \in \Gamma(TM)$,

where T is the torsion tensor of ∇.

The *Tanno's tensor* is the $(2, 1)$ tensor field defined by

$$Q(X, Y) := (\nabla_Y J)X, \qquad \forall\, X, Y \in \Gamma(TM).$$

A fundamental result due to Tanno is that (M, ω, g, J) is a (strongly pseudo-convex) CR manifold if and only if $Q = 0$. In this case, Tanno connection is Tanaka-Webster connection. Thus, Tanno connection is a natural generalisation of the Tanaka-Webster connection for contact structures that are not CR.

2.3. **K-type structures.** If X is an horizontal vector field, so is $T(X_0, X)$. As a consequence, if we define $\tau(X) = T(X_0, X)$, τ is a symmetric horizontal endomorphism which satisfies $\tau \circ J + J \circ \tau = 0$, by property (v). On CR manifolds, τ is called *pseudo-Hermitian torsion*. A contact structure is K-type iff X_0 is a Killing vector field or, equivalently, if $\tau = 0$.

2.4. **Yang-Mills structures.** We say that a contact sub-Riemannian structure is *Yang-Mills* if the torsion T of Tanno connection satisfies

$$\sum_{i=1}^{2d} (\nabla_{X_i} T)(X_i, Y) = 0, \qquad \forall\, Y \in \Gamma(TM), \tag{5}$$

for every orthonormal frame $X_1, \ldots, X_{2d}$ of $\mathscr{D}$. This definition coincides with the classical one given in [16, Sec. 5.1] for totally geodesic foliations with bundle like metric, and generalizes it to contact sub-Riemannian structures that are not foliations, e.g. when $\tau \neq 0$.

2.5. **Sasakian structures.** If (M, ω, g, J) is a contact sub-Riemannian manifold with Reeb vector X_0, consider the manifold $M \times \mathbb{R}$. We denote vector fields on $M \times \mathbb{R}$ by $(X, f\partial_t)$, where X is tangent to M and t is the coordinate on $\mathbb{R}$. Define the $(1, 1)$ tensor

$$\mathbf{J}(X, f\partial_t) = (JX - fX_0, \omega(X)\partial_t).$$

Indeed $\mathbf{J}^2 = -\mathbb{I}$ and thus defines an almost complex structure on $M \times \mathbb{R}$ (this clearly was not possible on the odd-dimensional manifold M). We say that the contact sub-Riemannian structure (M, ω, g, J) is *Sasakian* (or *normal*) if the almost complex structure $\mathbf{J}$ comes from a true complex structure. A celebrated theorem by Newlander and Nirenberg states that this

condition is equivalent to the vanishing of the Nijenhuis tensor of $\mathbf{J}$. For a $(1,1)$ tensor T, its Nijenhuis tensor $[T,T]$ is the $(2,1)$ tensor

$$[T,T](X,Y) := T^2[X,Y] + [TX,TY] - T[TX,Y] - T[X,TY].$$

In terms of the original structure, the integrability condition $[\mathbf{J},\mathbf{J}] = 0$ is equivalent to

$$[J,J](X,Y) + d\omega(X,Y)X_0 = 0.$$

Remark 2.4. A contact sub-Riemannian structure is Sasakian if and only if $Q = 0$ and $\tau = 0$, that is if and only if it is CR and K-type (see [14, Theorems 6.7, 6.3]). Notice that a three-dimensional contact sub-Riemannian structure is automatically CR (see [14, Cor. 6.4]), in particular, it is Sasakian if and only if it is K-type.

3. Jacobi fields revisited

Let $\lambda \in T^*M$ be the initial covector of a geodesic, projection of the extremal $\lambda(t) = e^{t\vec{H}}(\lambda)$. For any $\xi \in T_\lambda(T^*M)$ we define the field along the extremal $\lambda(t)$ as

$$X(t) := e_*^{t\vec{H}}\xi \in T_{\lambda(t)}(T^*M).$$

Definition 3.1. The set of vector fields obtained in this way is a $2n$-dimensional vector space, that we call *the space of Jacobi fields along the extremal.*

In the Riemannian case, the projection π_* is an isomorphisms between the space of Jacobi fields along the extremal and the classical space of Jacobi fields along the geodesic γ. Thus, this definition agrees with the standard one in Riemannian geometry and does not need curvature or connection.

In Riemannian geometry, the subspace of Jacobi fields vanishing at zero carries information about conjugate points along the given geodesic. This corresponds to the subspace of Jacobi fields along the extremal such that $\pi_* X(0) = 0$. This motivates the following construction.

For any $\lambda \in T^*M$, let $\mathcal{V}_\lambda := \ker \pi_*|_\lambda \subset T_\lambda(T^*M)$ be the *vertical subspace.* We define the family of Lagrangian subspaces along the extremal

$$\mathcal{L}(t) := e_*^{t\vec{H}}\mathcal{V}_\lambda \subset T_{\lambda(t)}(T^*M).$$

Remark 3.2. A time $t > 0$ is a conjugate time along γ if $\mathcal{L}(t) \cap \mathcal{V}_{\lambda(t)} \neq \{0\}$. The first conjugate time is the smallest conjugate time, namely $t_*(\gamma) = \inf\{t > 0 \mid \mathcal{L}(t) \cap \mathcal{V}_{\lambda(t)} \neq \{0\}\}$.

Notice that conjugate points correspond to the critical values of the sub-Riemannian exponential map with base at $\gamma(0)$. In other words, if $\gamma(t)$ is conjugate with $\gamma(0)$ along γ, there exists a one-parameter family of geodesics starting at $\gamma(0)$ and ending at $\gamma(t)$ at first order. Indeed, let $\xi \in \mathcal{V}_\lambda$ such that $\pi_* \circ e_*^{t\vec{H}}\xi = 0$, then the vector field $\tau \mapsto \pi_* \circ e_*^{\tau\vec{H}}\xi$ is precisely the vector field along $\gamma(\tau)$ of the aforementioned variation.

3.1. Linearized Hamiltonian.

For any vector field $X(t)$ along an integral line $\lambda(t)$ of the (sub-)Riemannian Hamiltonian flow, a dot denotes the Lie derivative in the direction of $\vec{H}$:

$$\dot{X}(t) := \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} e_*^{-\varepsilon\vec{H}} X(t+\varepsilon).$$

The space of Jacobi fields along the extremal $\lambda(t)$ coincides with the set of solutions of the *(sub-)Riemannian Jacobi equation*:

$$\dot{X} = 0.$$

We want to write the latter in a more standard way. Pick a Darboux frame $\{E_i(t), F_i(t)\}_{i=1}^n$ along $\lambda(t)$ (to fix ideas, one can think at the canonical basis $\{\partial_{p_i}|_{\lambda(t)}, \partial_{q_i}|_{\lambda(t)}\}$ induced by a choice of coordinates $(q_1,\dots,q_n)$ on M). In terms of this frame, $X(t)$ has components $(p(t),q(t)) \in \mathbb{R}^{2n}$:

$$X(t) = \sum_{i=1}^n p_i(t)E_i(t) + q_i(t)F_i(t).$$

The elements of the frame satisfy

$$\begin{pmatrix} \dot{E} \\ \dot{F} \end{pmatrix} = \begin{pmatrix} C_1(t) & -C_2(t) \\ R(t) & -C_1^*(t) \end{pmatrix} \begin{pmatrix} E \\ F \end{pmatrix},$$

for some smooth families of $n \times n$ matrices $C_1(t), C_2(t), R(t)$, where $C_2(t) = C_2(t)^*$ and $R(t) = R(t)^*$. The notation for these matrices will be clear in the following. We only stress here that the particular structure of the equations is implied solely by the fact that the frame is Darboux. Moreover, $C_2(t) \geq 0$ as a consequence of the non-negativity of the sub-Riemannian Hamiltonian (in the Riemannian case $C_2(t) > 0$). In turn, the Jacobi equation, written in terms of the components $p(t), q(t)$, becomes

$$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = \begin{pmatrix} -C_1(t)^* & -R(t) \\ C_2(t) & C_1(t) \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}. \tag{6}$$

3.2. **Canonical frame: Riemannian case.** In the Riemannian case one can choose a suitable frame (related with parallel transport) in such a way that, in Eq. (6), $C_1(t) = 0$, $C_2(t) = \mathbb{I}$ (in particular, they are constant), and the only remaining non-trivial block $R(t)$ is the curvature operator along the geodesic. The precise statement is as follows (see [9]).

Proposition 3.3. *Let $\lambda(t)$ be an integral line of a Riemannian Hamiltonian. There exists a smooth moving frame $\{E_i(t), F_i(t)\}_{i=1}^n$ along $\lambda(t)$ such that:*

(i) $\operatorname{span}\{E_1(t), \dots, E_n(t)\} = \mathcal{V}_{\lambda(t)}$.

(ii) *It is a Darboux basis, namely*

$$\sigma(E_i, E_j) = \sigma(F_i, F_j) = \sigma(E_i, F_j) - \delta_{ij} = 0, \qquad i, j = 1, \dots, n.$$

(iii) *The frame satisfies the structural equations*

$$\dot{E}_i = -F_i, \qquad \dot{F}_i = \sum_{j=1}^n R_{ij}(t) E_j,$$

for some smooth family of $n \times n$ symmetric matrices $R(t)$.

Moreover, the projections $f_i(t) := \pi_* F_i(t)$ are a parallel transported orthonormal frame along the geodesic $\gamma(t)$.

Properties (i)-(iii) uniquely define the moving frame up to orthogonal transformations: if $\{\widetilde{E}_i(t), \widetilde{F}_j(t)\}_{i=1}^n$ is another frame satisfying (i)-(iii), for some family $\widetilde{R}(t)$, then there exists a constant $n \times n$ orthogonal matrix O such that

$$\widetilde{E}_i(t) = \sum_{j=1}^n O_{ij} E_j(t), \qquad \widetilde{F}_i(t) = \sum_{j=1}^n O_{ij} F_j(t), \qquad \widetilde{R}(t) = O R(t) O^*. \tag{7}$$

A few remarks are in order. Property (ii) implies that $\operatorname{span}\{E_1, \dots, E_n\}$, $\operatorname{span}\{F_1, \dots, F_n\}$, evaluated at $\lambda(t)$, are Lagrangian subspaces of $T_{\lambda(t)}(T^*M)$. Eq. (7) reflects the fact that a parallel transported frame is defined up to a constant orthogonal transformation. In particular, one can use properties (i)-(iii) to *define* a parallel transported frame $\gamma(t)$ by $f_i(t) := \pi_* F_i(t)$. The symmetric matrix $R(t)$ induces a well defined operator $\mathfrak{R}_{\gamma(t)} : T_{\gamma(t)}M \to T_{\gamma(t)}M$

$$\mathfrak{R}_{\gamma(t)} f_i(t) := \sum_{j=1}^n R_{ij}(t) f_j(t).$$

Lemma 3.4. *Let $\overline{\mathrm{R}} : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$ the Riemannian curvature tensor w.r.t. the Levi-Civita connection. Then*

$$\mathfrak{R}_{\gamma(t)} v = \overline{\mathrm{R}}(v, \dot{\gamma}(t)) \dot{\gamma}(t), \qquad v \in T_{\gamma(t)}M.$$

In terms of the above parallel transported frame, the Jacobi equation (6) reduces to the classical one for Riemannian structures:

$$\ddot{q} + R(t) q = 0, \qquad \text{with} \qquad R_{ij}(t) = g\left(\overline{\mathrm{R}}(\dot{\gamma}(t), f_i(t)) f_j(t), \dot{\gamma}(t)\right).$$

3.3. **Canonical frame: sub-Riemannian case.** In the (general) sub-Riemannian setting, such a drastic simplification cannot be achieved, as the structure is more singular and more non-trivial invariants appear. Yet it is possible to simplify as much as possible the Jacobi equation (seen as a first order equation $\dot{X} = 0$) for a sufficiently regular extremal. This is achieved by the so-called *canonical Darboux frame*, introduced in [5, 20, 29].

In particular, $C_1(t)$ and $C_2(t)$ can be again put in a constant, normal form. However, in sharp contrast with the Riemannian case, their very structure may depend on the extremal. Besides, the remaining block $R(t)$ is still non-constant (in general) and *defines* a canonical curvature operator along $\gamma(t)$. As in the Riemannian case, this frame is unique up to constant, orthogonal transformations that preserve the "normal forms" of C_1 and C_2.

At this point, the discussion can be simplified as, for contact sub-Riemannian structures, C_1 and C_2 have only one possible form that is the same for all non-trivial extremal. We refer to the original paper [29] and the more recent [4, 9, 10] for a discussion of the general case and the relation between the normal forms and the invariants of the geodesic (the so-called *geodesic flag*).

Proposition 3.5. *Let $\lambda(t)$ be an integral line of a contact sub-Riemannian Hamiltonian. There exists a smooth moving frame along $\lambda(t)$*

$$E(t) = (E_0(t), E_1(t), \dots, E_{2d}(t))^*, \qquad F(t) = (F_0(t), F_1(t), \dots, F_{2d}(t))^*$$

such that the following hold true for any t:

(i) $\operatorname{span}\{E_0(t), \dots, E_{2d}(t)\} = \mathcal{V}_{\lambda(t)}$.

(ii) *It is a Darboux basis, namely*

$$\sigma(E_\mu, E_\nu) = \sigma(F_\mu, F_\nu) = \sigma(E_\mu, F_\nu) - \delta_{\mu\nu} = 0 \qquad \mu, \nu = 0, \dots, 2d.$$

(iii) *The frame satisfies the* structural equations

$$\dot{E}(t) = C_1 E(t) - C_2 F(t), \tag{8}$$

$$\dot{F}(t) = R(t) E(t) - C_1^* F(t), \tag{9}$$

where C_1, C_2 are $(2d+1) \times (2d+1)$ matrices defined by

$$C_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0_{2d-1} \end{pmatrix}, \qquad C_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathbb{I}_{2d-1} \end{pmatrix},$$

and $R(t)$ is a $(2d+1) \times (2d+1)$ smooth family of symmetric matrices of the form

$$R(t) = \begin{pmatrix} R_{00}(t) & 0 & R_{0j}(t) \\ 0 & R_{11}(t) & R_{1j}(t) \\ R_{i0}(t) & R_{i1}(t) & R_{ij}(t) \end{pmatrix}, \qquad i, j = 2, \dots, 2d.$$

Notice that C_1 is nilpotent, and C_2 is idempotent.

Moreover, the projections $f_i(t) := \pi_* F_i(t)$ for $i = 1, \dots, 2d$ are an orthonormal frame for $\mathscr{D}_{\gamma(t)}$ and $f_0(t) := \pi_* F_0(t)$ is transverse to $\mathscr{D}_{\gamma(t)}$.

If $\{\widetilde{E}(t), \widetilde{F}(t)\}$ is another frame that satisfies (i)-(iii) for some matrix $\widetilde{R}(t)$, then there exists a constant orthogonal matrix O that preserves the structural equations (i.e. $OC_iO^* = C_i$) and

$$\widetilde{E}(t) = OE(t), \qquad \widetilde{F}(t) = OF(t), \qquad \widetilde{R}(t) = OR(t)O^*.$$

3.4. **Invariant subspaces and curvature.** Let $f_\mu(t) = \pi_* F_\mu(t)$ a frame for $T_{\gamma(t)}M$ (here $\mu = 0, \dots, 2d$). The uniqueness part of Proposition 3.5 implies that this frame is unique up to a constant rotation of the form

$$O = \pm \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & U \end{pmatrix}, \qquad U \in \mathrm{O}(2d-1),$$

as one can readily check by imposing the conditions $OC_iO^* = C_i$. In particular, the following invariant subspaces of $T_{\gamma(t)}M$ are well defined

$$\begin{aligned} S^a_{\gamma(t)} &:= \mathrm{span}\{f_0\}, \\ S^b_{\gamma(t)} &:= \mathrm{span}\{f_1\}, \\ S^c_{\gamma(t)} &:= \mathrm{span}\{f_2, \dots, f_{2d}\}, \end{aligned}$$

and do not depend on the choice of the canonical frame (but solely on the geodesic $\gamma(t)$). The sets of indices $\{0\}$, $\{1\}$ and $\{2, \dots, 2d\}$ play distinguished roles and, in the following, we relabel these groups of indices as: $a = \{0\}$, $b = \{1\}$ and $c = \{2, \dots, 2d\}$. In particular we have:

$$T_{\gamma(t)}M = S^a_{\gamma(t)} \oplus S^b_{\gamma(t)} \oplus S^c_{\gamma(t)}.$$

With this notation, the curvature matrix $R(t)$ has the following block structure

$$R(t) = \begin{pmatrix} R_{aa}(t) & 0 & R_{ac}(t) \\ 0 & R_{bb}(t) & R_{bc}(t) \\ R_{ca}(t) & R_{cb}(t) & R_{cc}(t) \end{pmatrix}, \tag{10}$$

where $R_{aa}(t)$, $R_{bb}(t)$ are 1×1 matrices, $R_{ac}(t) = R_{ca}(t)^*, R_{bc}(t) = R_{cb}(t)^*$ are $1 \times (2d-1)$ matrices and R_{cc} is a $(2d-1) \times (2d-1)$ matrix.

Definition 3.6. The *canonical curvature* is the symmetric operator $\mathfrak{R}_{\gamma(t)} : T_{\gamma(t)}M \to T_{\gamma(t)}M$, that, in terms of the basis $f_a, f_b, f_{c_2}, \dots, f_{c_{2d}}$ is represented by the matrix $R(t)$.

The definition is well posed, in the sense that different canonical frames give rise to the same operator. For $\alpha, \beta = a, b, c$, we denote by $\mathfrak{R}^{\alpha\beta}_{\gamma(t)} : S^\alpha_{\gamma(t)} \to S^\beta_{\gamma(t)}$ the restrictions of the canonical curvature to the appropriate invariant subspace.

Definition 3.7. The *canonical Ricci curvatures* are the partial traces

$$\mathfrak{Ric}^\alpha_{\gamma(t)} := \mathrm{tr}\left(\mathfrak{R}^{\alpha\alpha}_{\gamma(t)} : S^\alpha_{\gamma(t)} \to S^\alpha_{\gamma(t)}\right), \qquad \alpha = a, b, c.$$

The canonical curvature $\mathfrak{R}(t)$ contains all the information on the germ of the structure along the geodesic. More precisely, consider two pairs (M, g, γ) and (M', g', γ') where (M, g) and (M', g') are two contact sub-Riemannian structures and γ, γ' two geodesics. Then $\mathfrak{R}_{\gamma(t)}$ is congruent to $\mathfrak{R}'_{\gamma'(t)}$ if and only if the linearizations of the respective geodesic flows (as flows on the cotangent bundle, along the respective extremals) are equivalent (i.e. symplectomorphic).

4. Geodesic cost and its asymptotic

We start by a characterization of smooth points of the squared distance d^2 on a contact manifold. Let $x_0 \in M$, and let $\Sigma_{x_0} \subset M$ be the set of points x such that there exists a unique length minimizer $\gamma : [0, 1] \to M$ joining x_0 with x, that is non-conjugate.

Theorem 4.1 (see [2])**.** *Let $x_0 \in M$ and set $\mathfrak{f} := \frac{1}{2}\mathsf{d}^2(x_0, \cdot)$. 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 extremal of the unique length minimizer $\gamma(t)$ joining x_0 to x in time 1.*

Remark 4.2. By homogeneity of the Hamiltonian, if in the statement of the previous theorem one considers a geodesic $\gamma : [0, T] \to M$ joining x_0 and x in time T, then $d_x\mathfrak{f} = T\lambda(T)$.

Remark 4.3. The statement of Theorem 4.1 is valid on a general sub-Riemannian manifold, with Σ_{x_0} defined as the set of points x such that there exists a unique minimizer joining x_0 with x, which must be not abnormal, and such that x is not conjugate with x_0 (see [2–4, 24]).

Remark 4.4. If M is a contact sub-Riemannian manifold, the function $\mathfrak{f} = \frac{1}{2}\mathsf{d}^2(x_0, \cdot)$ is Lipschitz on $M \setminus \{x_0\}$, due to the absence of abnormal minimizers (see for instance [3, 23]). In particular from this one can deduce that the set $M \setminus \Sigma_{x_0}$ where $\mathfrak{f}$ is not smooth has measure zero.

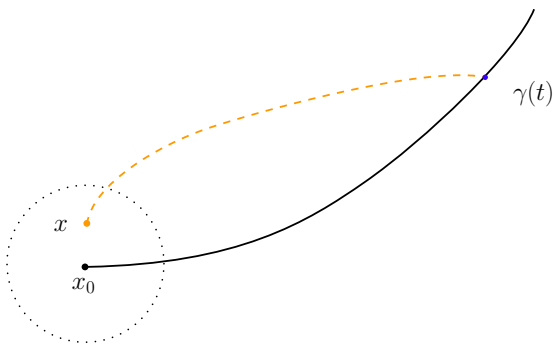

FIGURE 1. The geodesic cost function.

Definition 4.5. Let $x_0 \in M$ and consider a geodesic $\gamma : [0,T] \to M$ such that $\gamma(0) = x_0$. The *geodesic cost* associated with γ is the family of functions for $t > 0$

$$c_t(x) := -\frac{1}{2t}\mathsf{d}^2(x, \gamma(t)), \qquad x \in M.$$

From Theorem 4.1 one obtains smoothness properties of the geodesic cost.

Theorem 4.6. *Let $x_0 \in M$ and $\gamma(t) = \mathcal{E}_{x_0}(t\lambda_0)$ be a geodesic. Then there exists $\varepsilon > 0$ 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 function $(t,x) \mapsto c_t(x)$ is smooth on U,*
(iii) *for any $(t,x) \in U$, the covector $\lambda_x = d_x c_t$ is the initial covector of the unique geodesic connecting x with $\gamma(t)$ in time t.*

In particular $\lambda_0 = d_{x_0} c_t$ and x_0 is a critical point for the function $\dot{c}_t := \frac{d}{dt} c_t$ for every $t \in (0,\varepsilon)$.

Proof. There exists $\varepsilon > 0$ small enough such that for $t \in (0,\varepsilon)$, the curve $\gamma|_{[0,t]}$ is the unique minimizer joining $x_0 = \gamma(0)$ and $\gamma(t)$, and $\gamma(t)$ is non conjugate to $\gamma(0)$. As a direct consequence of Theorem 4.1 one gets $c_t(x)$ is smooth for fixed $t \in (0,\varepsilon)$ and x in a neighborhood of x_0. The fact that the function $c_t(x)$ is smooth on an open set U as a function of the two variables is proved in [4, Appendix A]. Let us prove (iii). Notice that

$$d_{x_0} c_t = -\frac{1}{t} d_{x_0}\left(\frac{1}{2}\mathsf{d}^2(\gamma(t), \cdot)\right) = -\frac{1}{t} d_{x_0}\mathfrak{f}_t.$$

where $\mathfrak{f}_t$ denotes one half of the squared distance from the point $\gamma(t)$. Observe that $x \in \Sigma_{\gamma(t)}$ is in the set of smooth points of the squared distance from $\gamma(t)$. Hence the differential $d_x\mathfrak{f}_t$ is the final covector of the unique geodesic joining $\gamma(t)$ with x in time 1. Thus, $-d_x\mathfrak{f}_t$ is the initial covector of the unique geodesic joining x with $\gamma(t)$ in time 1, and $d_x c_t = -\frac{1}{t} d_x\mathfrak{f}_t$ is the initial covector λ_0 of the unique geodesic connecting x with $\gamma(t)$ in time t. □

4.1. **A geometrical interpretation.** By Theorem 4.6, for each $(t,x) \in U$ the function $x \mapsto \dot{c}_t(x)$ has a critical point at x_0. This function will play a crucial role in the following, and has a nice geometrical interpretation. Let $W^t_{x,\gamma(t)} \in T^*_{\gamma(t)}M$ be the final tangent vector of the unique minimizer connecting x with $\gamma(t)$ in time t. We have:

$$\dot{c}_t(x) = \frac{1}{2}\|W^t_{x_0,\gamma(t)} - W^t_{x,\gamma(t)}\|^2 - \frac{1}{2}\|W^t_{x_0,\gamma(t)}\|^2. \tag{11}$$

This formula is a consequence of Theorem 4.1 and is proved in [4, Appendix I]. The second term $\|W^t_{x_0,\gamma(t)}\|$ is the speed of the geodesic γ and is then an inessential constant. Eq. (11) has also natural physical interpretation as follows. Suppose that A and B live at points x_A and x_B respectively (see Fig. 2). Then A chooses a geodesic $\gamma(t)$, starting from x_A, and tells B to meet at some point $\gamma(t)$ (at time t). Then B must choose carefully its geodesic in order to meet A at the point $\gamma(t)$ starting from x_B, following the curve for time t. When they meet at $\gamma(t)$ at time

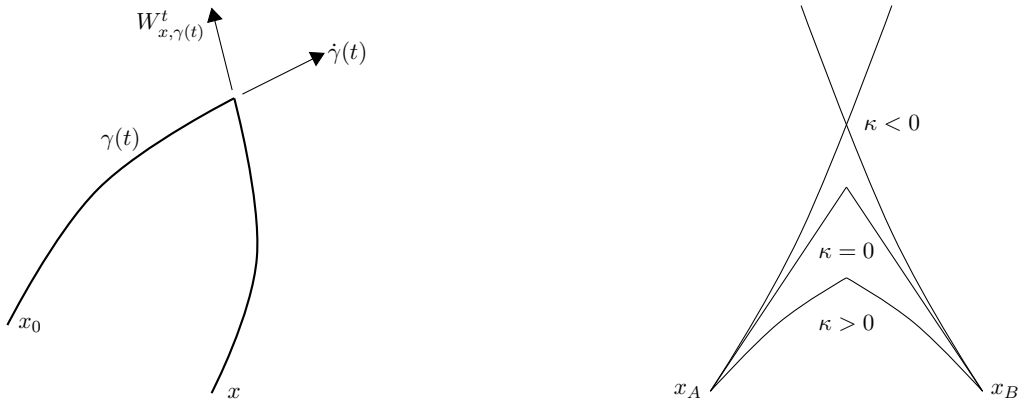

FIGURE 2. Geometric interpretation of $\dot c_t$. $W^t_{x,\gamma(t)}$ is the tangent vector, at time t, of the unique geodesic connecting x with $\gamma(t)$ in time t.

t, they compare their velocity by computing the squared norm (or *energy*) of the difference of the tangent vectors. This gives the value of the function $\dot c_t$, up to a constant.

The "curvature at x_0" is encoded in the behavior of this function for small t and x close to x_0. To support this statement, consider the Riemannian setting, in which Eq. (11) clearly remains true. If the Riemannian manifold is positively (resp. negatively) curved, then the two tangent vectors, compared at $\gamma(t)$, are more (resp. less) divergent w.r.t. the flat case (see Fig. 2).

Remark 4.7. We do not need parallel transport: A and B meet at $\gamma(t)$ and make *there* their comparison. We only used the concept of "optimal trajectory" and "difference of the cost". This interpretation indeed works for a general optimal control system, as in the general setting of [4].

4.2. **A family of operators.** Let $f : M \to \mathbb{R}$ be a smooth function. 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 is well defined only at a critical point, i.e. at those points x such that $d_x f = 0$. In this case the map

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

where V, W are vector fields such that $V(x) = v$ and $W(x) = w$, respectively, is a well defined symmetric bilinear form that does not depend on the choice of the extensions. The associated quadratic form, that we denote by the same symbol $d^2_x f : T_x M \to \mathbb{R}$, is defined by

$$d^2_x f(v) = \frac{d^2}{dt^2}\Big|_{t=0} f(\gamma(t)), \qquad \text{where} \quad \gamma(0) = x, \quad \dot\gamma(0) = v.$$

By Theorem 4.6, for small $t > 0$ the function $x \mapsto \dot c_t(x)$ has a critical point at x_0. Hence we can consider the family of quadratic forms restricted on the distribution

$$d^2_{x_0} \dot c_t\big|_{\mathscr{D}_{x_0}} : \mathscr{D}_{x_0} \to \mathbb{R}, \qquad \forall\, t > 0.$$

Using the sub-Riemannian scalar product on $\mathscr{D}_{x_0}$ we can associate with this family of quadratic forms the family of symmetric operators $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ defined for small $t > 0$ by

$$d^2_{x_0} \dot c_t(w) = g(\mathcal{Q}_\lambda(t) w, w), \qquad \forall\, w \in \mathscr{D}_{x_0}, \tag{12}$$

where $\lambda = d_{x_0} c_t$ is the initial covector of the fixed geodesic γ.

Theorem 4.8. *Let M be a contact sub-Riemannian manifold and $\gamma(t)$ be a non-trivial geodesic with initial covector $\lambda \in T^*_{x_0} M$. Let $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ be as in* (12).

The family of operators $t \mapsto t^2 \mathcal{Q}_\lambda(t)$ can be extended to a smooth family of operators on $\mathscr{D}_{x_0}$ for small $t \geq 0$, symmetric with respect to g. Moreover,

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

Then we have the following Laurent expansion at $t = 0$:

$$\mathcal{Q}_\lambda(t) = \frac{1}{t^2}\mathcal{I}_\lambda + \sum_{i=0}^{m} \mathcal{Q}_\lambda^{(i)} t^i + O(t^{m+1}).$$

In particular, Theorem 4.8 *defines* a sequence of operators $\mathcal{Q}_\lambda^{(i)}$, for $i \in \mathbb{N}$. These operators, together with $\mathcal{I}_\lambda$, contain all the information on the germ of the structure along the fixed geodesic (clearly any modification of the structure that coincides with the given one on a neighborhood of the geodesic gives the same family of operators).

Remark 4.9. Applying this construction to a Riemannian manifold, one finds that $\mathcal{I}_\lambda = \mathbb{I}$ for any geodesic, and the operator $\mathcal{Q}_\lambda^{(0)}$ is indeed the "directional" sectional curvature in the direction of the geodesic (see [4, Sect. 4.4.2]):

$$\mathcal{Q}_\lambda^{(0)} = \frac{1}{3}\overline{\mathrm{R}}(\cdot, \dot\gamma)\dot\gamma,$$

where $\dot\gamma$ is the initial velocity vector of the geodesic (that is the vector corresponding to λ in the canonical isomorphism $T_x^*M \simeq T_xM$ defined by the Riemannian metric).

The expansion is interesting in two directions. First, the singularity is controlled ($\mathcal{Q}_\lambda(t)$ has a second order pole at $t = 0$), even though the sub-Riemannian squared distance is not regular on the diagonal. Second, the Laurent polynomial of second order of $\mathcal{Q}_\lambda(t)$ recovers the whole canonical curvature operator $\mathfrak{R}$ at the point. The next two theorems describe in detail the Laurent expansion of the operator $\mathcal{Q}_\lambda(t)$.

Theorem 4.10. *The symmetric operator $\mathcal{I}_\lambda : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ satisfies*

(i) $\operatorname{spec}\mathcal{I}_\lambda = \{1, 4\}$,
(ii) $\operatorname{tr}\mathcal{I}_\lambda = 2d + 3$.

The vector subspaces of $\mathscr{D}_{\gamma(0)}$

$$S^b_{\gamma(0)} := \operatorname{span}\{J\dot\gamma(0)\}, \qquad S^c_{\gamma(0)} := J\dot\gamma(0)^\perp \cap \mathscr{D}_{\gamma(0)}$$

are the eigenspaces of $\mathcal{I}_\lambda$ corresponding to eigenvalues 4 and 1, respectively. In particular the eigenvalue 4 has multiplicity 1 while the eigenvalue 1 has multiplicity $2d - 1$.

Remark 4.11. Let $K_{\gamma(0)} \subset \mathscr{D}_{\gamma(0)}$ be the hyperplane $d\omega$-orthogonal to $\dot\gamma(0)$, that is

$$K_{\gamma(0)} := \{v \in \mathscr{D}_{\gamma(0)} \mid d\omega(v, \dot\gamma(0)) = 0\}.$$

Then $S^c_{\gamma(0)} = K_{\gamma(0)}$ and $S^b_{\gamma(0)} = K^\perp_{\gamma(0)} \cap \mathscr{D}_{\gamma(0)}$.

Theorem 4.12. *The asymptotics of $\mathcal{Q}_\lambda(t)$ recovers the canonical curvature at $\gamma(0)$. More precisely, according to the orthogonal decomposition $\mathscr{D}_{\gamma(0)} = S^b_{\gamma(0)} \oplus S^c_{\gamma(0)}$ we have*

$$\mathcal{Q}_\lambda^{(0)} = \begin{pmatrix} \frac{2}{15}R_{bb} & \frac{1}{12}R_{bc} \\ \frac{1}{12}R_{cb} & \frac{1}{3}R_{cc} \end{pmatrix},$$

$$\mathcal{Q}_\lambda^{(1)} = \begin{pmatrix} \frac{1}{15}\dot R_{bb} & \frac{1}{10}R_{ac} - \frac{1}{30}\dot R_{bc} \\ \frac{1}{10}R_{ca} - \frac{1}{30}\dot R_{cb} & \frac{1}{6}\dot R_{cc} \end{pmatrix},$$

$$\mathcal{Q}_\lambda^{(2)} = \frac{1}{240}\begin{pmatrix} \frac{1}{35}\left(240R_{aa} + 44R_{bb}^2 + 65R_{bc}R_{cb} + 240\ddot R_{bb}\right) & R_{bb}R_{bc} - 2R_{bc}R_{cc} + 12\dot R_{ac} - 6\ddot R_{bc} \\ R_{cb}R_{bb} - 2R_{cc}R_{cb} + 12\dot R_{ca} - 6\ddot R_{cb} & 16R_{cc}^2 + R_{bc}R_{cb} + 12\ddot R_{cc} \end{pmatrix},$$

where the dot denotes the derivative w.r.t. t and here $R_{\alpha\beta} = R_{\alpha\beta}(0)$ for $\alpha, \beta = a, b, c$ are the components of the canonical curvature operator $\mathfrak{R}_{\gamma(0)} : T_{x_0}M \to T_{x_0}M$.

Remark 4.13. To recover the operator $\mathfrak{R}_{\gamma(s)} : T_{\gamma(s)}M \to T_{\gamma(s)}M$ for all s on the geodesic, one needs the Laurent expansion of the operator $\mathcal{Q}_{\lambda(s)}(t)$ for all the points $\lambda(s)$ on the extremal.

Proof of Theorems. 4.8, 4.10 and 4.12. For a m-tuple $v = (v_1, \dots, v_m)$ of vectors and a linear map $L : V \to W$, the notation $L(v)$ denotes the m-tuple $(Lv_1, \dots, Lv_m)$. Consider a canonical frame $E(t), F(t)$ from Proposition 3.5. Accordingly, $f(t) := \pi_* F(t) \in T_{\gamma(t)}M$ is the $(2d+1)$-tuple

$$f(t) = (f_0(t), \underbrace{f_1(t), \dots, f_{2d}(t)}_{\text{o.n. basis for } \mathscr{D}_{\gamma(t)}})^*.$$

The geodesic cost function c_t is regular at x_0, and thus its second differential is not well defined as a quadratic form on $T_{x_0}M$. Still it is well defined as a linear map as follows. Consider the differential $dc_t : M \to T^*M$. Taking again the differential at x_0, we get a map $d^2_{x_0}c_t : T_{x_0}M \to T_\lambda(T^*M)$, where $\lambda = d_{x_0}c_t$ is the initial covector associated with the geodesic.

For fixed $t > 0$ and x in a neighborhood U_t of x_0, we clearly have $\pi(d_x c_t) = x$. Thus $d^2_{x_0}c_t$ has maximal rank. Similarly, $\pi(e^{t\vec{H}} d_x c_t) = \gamma(t)$ by Theorem 4.6. It follows that $e_*^{t\vec{H}} d^2_{x_0}c_t(T_{x_0}M) = \mathcal{V}_{\lambda(t)}$. Then there exists a smooth family of $(2d+1) \times (2d+1)$ matrices $\Theta(t)$ such that

$$e_*^{t\vec{H}} d^2_{x_0}c_t(f(0)) = \Theta(t)E(t). \tag{13}$$

We pull back the canonical frame $E(t)$ through the Hamiltonian flow, and express this in terms of the canonical frame at time zero:

$$e_*^{-t\vec{H}} E(t) = A(t)E(0) + B(t)F(0), \qquad t \geq 0. \tag{14}$$

for some smooth families of $(2d+1) \times (2d+1)$ matrices $A(t), B(t)$ (in particular $A(0) = \mathbb{I}$ and $B(0) = 0$). As we already noticed, $\pi(d_x c_t) = x$ (here $t > 0$), then

$$\begin{aligned} f(0) &= \pi_* d^2_{x_0}c_t(f(0)) \\ &= \Theta(t)\pi_* e_*^{-t\vec{H}} E(t) && \text{(by Eq. (13))} \\ &= \Theta(t)\pi_*(A(t)E(0) + B(t)F(0)) && \text{(by Eq. (14))} \\ &= \Theta(t)B(t)f(0) && \text{(by definition of } f \text{ and verticality of } E) \end{aligned}$$

Thus $\Theta(t)B(t) = \mathbb{I}$ (in particular, $B(t)$ is not-degenerate for small $t > 0$). Then

$$d^2_{x_0}c_t(f(0)) = B(t)^{-1}A(t)E(0) + F(0), \qquad t > 0.$$

Let $S(t) := A(t)^{-1}B(t)$. We get, taking a derivative

$$d^2_{x_0}\dot{c}_t(f(0)) = \frac{d}{dt}\left[S(t)^{-1}\right] E(0).$$

In particular, $d^2_{x_0}\dot{c}_t$ maps $T_{x_0}M$ to the vertical subspace $\mathcal{V}_\lambda = T_\lambda(T^*_{x_0}M) \simeq T^*_{x_0}M$. By the latter identification, we recover the standard second differential (the Hessian) $d^2_{x_0}\dot{c}_t : T_{x_0}M \to T^*_{x_0}M$ at the critical point x_0. The associated quadratic form is, for $t > 0$,

$$\langle d^2_{x_0}\dot{c}_t(f(0)), f(0)\rangle = \frac{d}{dt}\left[S(t)^{-1}\right]. \tag{15}$$

According to (12), $\mathcal{Q}_\lambda(t) : \mathscr{D}_{x_0} \to \mathscr{D}_{x_0}$ is the operator associated with the quadratic form $d^2_{x_0}\dot{c}_t : T_{x_0}M \to T^*_{x_0}M$, restricted on $\mathscr{D}_{x_0}$, through the sub-Riemannian product. From Proposition 3.5, that $f = (f_0, f_1, \dots, f_{2d})$ where $f_1, \dots, f_{2d}$ is an orthonormal frame for $\mathscr{D}$. Then, in terms of the basis $f_1, \dots, f_{2d}$ of $\mathscr{D}_{x_0}$ (here we suppressed explicit evaluation at $t = 0$) we have, from (15):

$$\mathcal{Q}_\lambda(t) = \frac{d}{dt}\left[S(t)^{-1}\right]_\square.$$

where the notation $M_\square$ denotes the bottom right $2d \times 2d$ block of a matrix M.

We are left to compute the asymptotics of the matrix $S(t)^{-1} = B(t)^{-1}A(t)$, where $A(t)$ and $B(t)$ are defined by Eq. (14). In order to do that we study a more general problem. We introduce smooth matrices $C(t)$, $D(t)$ such that (here $t \geq 0$)

$$\begin{aligned} e_*^{-t\vec{H}} E(t) &= A(t)E(0) + B(t)F(0), \\ e_*^{-t\vec{H}} F(t) &= C(t)E(0) + D(t)F(0). \end{aligned}$$

Since $(E(t), F(t))$ satisfies the structural equations (8)-(9), we find that A, B, C, D are the solutions of the following Cauchy problem:

$$\frac{d}{dt}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} C_1 & -C_2 \\ R(t) & -C_1^* \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \qquad \begin{pmatrix} A(0) & B(0) \\ C(0) & D(0) \end{pmatrix} = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix}. \tag{16}$$

From (16) one can compute an arbitrarily high order Taylor polynomial of $A(t)$ and $B(t)$, needed for the computation of $S(t)^{-1}$.

Notice that $A(0) = \mathbb{I}$, but $B(0) = 0$ (but $B(t)$ is non-singular as soon as $t > 0$). We say that $B(t)$ has order r (at $t = 0$) if r is the smallest integer such that the r-th order Taylor polynomial of $B(t)$ (at $t = 0$) is non-singular. It follows from explicit computation that $B(t)$ has order 3:

$$B(t) = -t\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mathbb{I}_{2d-1} \end{pmatrix} + \frac{t^2}{2}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{t^3}{6}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + O(t^4).$$

In particular, for $t > 0$, the Laurent polynomial of $B(t)^{-1}$ has poles of order ≥ 3. Thus, to compute the Laurent polynomial of $S(t)^{-1} = B(t)^{-1}A(t)$ at order N, one needs to compute $B(t)$ and $A(t)$ at order $N + 6$ and $N + 3$, respectively. After explicit and long computations, one finds

$$S(t)^{-1} = \frac{1}{t^3}\begin{pmatrix} 12 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{t^2}\begin{pmatrix} 0 & -6 & 0 \\ -6 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{t}\begin{pmatrix} \frac{6}{5}R_{bb} & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -\mathbb{I}_{2d-1} \end{pmatrix} + O(1).$$

In particular, for the restriction we obtain

$$\left[S(t)^{-1}\right]_\square = -\frac{1}{t}\begin{pmatrix} 4 & 0 \\ 0 & \mathbb{I}_{2d-1} \end{pmatrix} + O(1).$$

The restriction to the distribution has the effect of taking the "least degenerate" block. Taking the derivative, we obtain

$$\frac{d}{dt}\left[S(t)^{-1}\right]_\square = \frac{1}{t^2}\begin{pmatrix} 4 & 0 \\ 0 & \mathbb{I}_{2d-1} \end{pmatrix} + O(t). \tag{17}$$

This proves Theorems 4.8 and 4.10, as $\mathcal{I}_\lambda = \lim_{t\to 0^+} t^2\mathcal{Q}_\lambda(t)$ is finite and $\geq \mathbb{I}$.

For Theorem 4.12, we need the Laurent polynomial of $\mathcal{Q}_\lambda(t) = \frac{d}{dt}\left[S(t)^{-1}\right]_\square$ at order 2. That is, we need $S(t)^{-1}$ at order $N = 3$. According to the discussion above, this requires the solution of Cauchy problem (16) at order $N+6 = 9$. This is achieved by long computations (made easier by the properties of C_1, C_2: $C_1^2 = 0$, $C_2^2 = C_2$, $C_1C_2 = C_1$ and $C_2C_1 = 0$). One checks that

$$\frac{d}{dt}\left[S(t)^{-1}\right]_\square = \frac{1}{t^2}\mathcal{I}_\lambda + \mathcal{Q}_\lambda^{(0)} + t\mathcal{Q}_\lambda^{(1)} + t^2\mathcal{Q}_\lambda^{(2)} + O(t^3),$$

where $\mathcal{I}_\lambda$ was computed in Eq. (17) while

$$\mathcal{Q}_\lambda^{(0)} = \begin{pmatrix} \frac{2}{15}R_{bb} & \frac{1}{12}R_{bc} \\ \frac{1}{12}R_{cb} & \frac{1}{3}R_{cc} \end{pmatrix},$$

$$\mathcal{Q}_\lambda^{(1)} = \begin{pmatrix} \frac{1}{15}\dot{R}_{bb} & \frac{1}{10}R_{ac} - \frac{1}{30}\dot{R}_{bc} \\ \frac{1}{10}R_{ca} - \frac{1}{30}\dot{R}_{cb} & \frac{1}{6}\dot{R}_{cc} \end{pmatrix},$$

$$\mathcal{Q}_\lambda^{(2)} = \frac{1}{240}\begin{pmatrix} \frac{1}{35}\left(240R_{aa} + 44R_{bb}^2 + 65R_{bc}R_{cb} + 240\ddot{R}_{bb}\right) & R_{bb}R_{bc} - 2R_{bc}R_{cc} + 12\dot{R}_{ac} - 6\ddot{R}_{bc} \\ R_{cb}R_{bb} - 2R_{cc}R_{cb} + 12\dot{R}_{ca} - 6\ddot{R}_{cb} & 16R_{cc}^2 + R_{bc}R_{cb} + 12\ddot{R}_{cc} \end{pmatrix}.$$

where we labeled the indices of $R(t)$ as in (10), that is $a = \{0\}$, $b = \{1\}$ and $c = \{2, \dots, 2d\}$.

This computes the asymptotics of $\mathcal{Q}_\lambda(t)$ in terms of the orthonormal frame $f_1(0), \dots, f_{2d}(0)$ obtained as a projection of the canonical frame. By definition (see Sec. 3.4) $S^b_{\gamma(0)} = \mathrm{span}\{f_1(0)\}$ and $S^c_{\gamma(0)} = \mathrm{span}\{f_2(0), \dots, f_{2d}(0)\}$ and thus concludes the proofs of statements (i)-(ii) of

Theorem 4.10 and Theorem 4.12. The explicit form of $S^b_{\gamma(0)}$ and $S^c_{\gamma(0)}$ follows from computation of the canonical frame in Theorem 6.1. □

4.3. **Reparametrization and homogeneity.** The operators $\mathcal{I}_\lambda$ and $\mathcal{Q}^{(i)}_\lambda$ for $i \geq 0$, are well defined for each nontrivial geodesic on the contact manifold M, i.e. for each covector $\lambda \in T^*_x M$ such that $H(\lambda) \neq 0$.

By homogeneity of the Hamiltonian, if $H(\lambda) \neq 0$, then also $H(\alpha\lambda) = \alpha^2 H(\lambda) \neq 0$ for $\alpha \neq 0$.

Proposition 4.14. *For all λ such that $H(\lambda) \neq 0$, the operators $\mathcal{I}_\lambda, \mathcal{Q}^{(i)}_\lambda : \mathscr{D}_x \to \mathscr{D}_x$ satisfy the following homogeneity properties:*

$$\mathcal{I}_{\alpha\lambda} = \mathcal{I}_\lambda, \qquad \mathcal{Q}^{(i)}_{\alpha\lambda} = \alpha^{2+i}\mathcal{Q}^{(i)}_\lambda. \qquad i \in \mathbb{N},\ \alpha > 0.$$

Proof. Here we denote by $x \mapsto c^\lambda_t(x)$ the geodesic cost function associated with $\lambda \in T^*_x M$. For $\alpha > 0$, if $\mathcal{E}_x(t\lambda) = \gamma(t)$ then $\mathcal{E}_x(t\alpha\lambda) = \mathcal{E}_x(\alpha t\lambda) = \gamma(\alpha t)$. Hence

$$c^{\alpha\lambda}_t(x) = -\frac{1}{2t}\mathsf{d}^2(x, \gamma(\alpha t)) = -\frac{\alpha}{2\alpha t}\mathsf{d}^2(x, \gamma(\alpha t)) = \alpha c^\lambda_{\alpha t}(x).$$

This implies that $\dot{c}^{\alpha\lambda}_t = \alpha^2 \dot{c}^\lambda_{\alpha t}$ and $d^2_x \dot{c}^{\alpha\lambda}_t = \alpha^2 d^2_x \dot{c}^\lambda_{\alpha t}$. If we restrict these quadratic forms to the distribution $\mathscr{D}_x$, we get the identity

$$\mathcal{Q}_{\alpha\lambda}(t) = \alpha^2 \mathcal{Q}_\lambda(\alpha t). \tag{18}$$

Applying the expansion of Theorem 4.8 to (18) we get

$$\frac{1}{t^2}\mathcal{I}_{\alpha\lambda} + \sum_{i=0}^m \mathcal{Q}^{(i)}_{\alpha\lambda} t^i + O(t^{m+1}) = \alpha^2 \left(\frac{1}{\alpha^2 t^2}\mathcal{I}_\lambda + \sum_{i=0}^m \alpha^i \mathcal{Q}^{(i)}_\lambda t^i + O(t^{m+1}) \right),$$

which gives the desired identities. □

5. Bonnet-Myers type theorems

In [9], under suitable bounds on the canonical curvature of a given sub-Riemannian geodesic, we obtained bounds on the first conjugate time. Then, with global conditions on $\mathfrak{R}$, one obtains Bonnet-Myers type results. We first specify the results of [9, Sec. 6] to contact structures (in the contact case all the necessary hypotheses apply). Then, using the results of Sec. 6, we express them in terms of known tensors (Tanno's tensor, curvature and torsion).

Theorem 5.1. *Assume that for every unit speed geodesic $\gamma(t)$ of a complete contact sub-Riemannian structure of dimension $2d+1$, with $d > 1$, we have*

$$\mathfrak{Ric}^c_{\gamma(t)} \geq (2d-2)\kappa_c, \qquad \forall t \geq 0,$$

for some $\kappa_c > 0$. Then the manifold is compact, with sub-Riemannian diameter not greater than $\pi/\sqrt{\kappa_c}$, and the fundamental group is finite.

Theorem 5.2. *Assume that for every unit speed geodesic $\gamma(t)$ of a complete contact sub-Riemannian structure of dimension $2d+1$, with $d \geq 1$, we have*

$$\mathfrak{Ric}^a_{\gamma(t)} \geq \kappa_a, \quad \mathfrak{Ric}^b_{\gamma(t)} \geq \kappa_b, \qquad \forall t \geq 0,$$

for some $\kappa_a, \kappa_b \in \mathbb{R}$ such that

$$\begin{cases} \kappa_b > 0, \\ 4\kappa_a > -\kappa_b^2, \end{cases} \qquad \text{or} \qquad \begin{cases} \kappa_b \leq 0, \\ \kappa_a > 0. \end{cases} \tag{19}$$

Then the manifold is compact, with sub-Riemannian diameter not greater than $t_*(\kappa_a, \kappa_b) < +\infty$, and the fundamental group is finite.

Remark 5.3. The value of $t_*(\kappa_a, \kappa_b)$ in Theorem 5.2 is the first conjugate time of a particular variational problem (a so-called *LQ optimal control problem*) which, in the framework of [9], plays the role of a constant curvature model. These are *synthetic models*, that are not sub-Riemannian manifolds but on which the concept of conjugate time t_* is well defined. It can be explicitly computed as the smallest positive root of a transcendental equation. In this specific case, $t_*(\kappa_a, \kappa_b)$ is the first positive solution of:

$$\frac{4\kappa_a[1-\cos(\lambda_+ t)\cosh(\lambda_- t)] - 2\kappa_b\lambda_+\lambda_-\sin(t\lambda_+)\sinh(t\lambda_-)}{2\kappa_a(4\kappa_a+\kappa_b^2)} = 0, \quad \lambda_\pm = \sqrt{\frac{\pm\kappa_b+\sqrt{4\kappa_a+\kappa_b^2}}{2}}. \tag{20}$$

This equation admits positive roots (actually, an infinite number of them) precisely if and only if (19) are satisfied. This result, that already in this simple case is quite cumbersome to check, comes from an general theorem on the existence of conjugate times for LQ optimal control problems, obtained in [7]. For $\kappa_a = 0$, $\kappa_b > 0$, (20) must be taken in the limit $\kappa_a \to 0$ and reduces to $2 - 2\cos(\sqrt{\kappa_b}t) - \sin(\sqrt{\kappa_b}t) = 0$ and $t_*(0, \kappa_b) = 2\pi/\sqrt{\kappa_b}$.

With the explicit expressions for the canonical curvature of Theorem 6.3, we obtain, from Theorem 5.1, the following result.

Theorem 5.4. *Consider a complete, contact structure of dimension $2d+1$, with $d > 1$. Assume that there exist constants $\kappa_1 > \kappa_2 \geq 0$ such that, for any horizontal unit vector X*

$$\mathrm{Ric}(X) - R(X, JX, JX, X) \geq (2d-2)\kappa_1, \qquad \|Q(X,X)\|^2 \leq (2d-2)\kappa_2. \tag{21}$$

Then the manifold is compact with sub-Riemannian diameter not greater than $\pi/\sqrt{\kappa_1-\kappa_2}$, and the fundamental group is finite.

Proof. From the results of Theorem 6.3, we have for any unit speed geodesic $\gamma(t)$

$$\mathfrak{Ric}^c_{\gamma(t)} = \mathrm{Ric}(\mathsf{T}) - R(\mathsf{T}, J\mathsf{T}, J\mathsf{T}, \mathsf{T}) + \frac{1}{4}h_0^2(2d-2) - \|Q(\mathsf{T},\mathsf{T})\|^2,$$

where $\mathsf{T} = \dot\gamma(t)$. Under our assumptions

$$\mathfrak{Ric}^c_{\gamma(t)} \geq (2d-2)(\kappa_1-\kappa_2), \qquad \forall\, t \geq 0.$$

The result then follows from Theorem 5.1. □

Remark 5.5. The Ricci tensor appearing in Eq. (21) is the trace of Tanno's curvature, that is

$$\mathrm{Ric}(X) = \sum_{i=1}^{2d} R(X, X_i, X_i, X) + R(X, X_0, X_0, X),$$

for any orthonormal basis $X_1, \ldots, X_{2d}$ of the sub-Riemannian structure. One can check that $R(X, X_0, X_0, X) = 0$ (since $\nabla X_0 = 0$ for Tanno connection). Therefore the l.h.s. of the first formula of Eq. (21) is the partial trace on the $2d-2$ dimensional subspace $\mathrm{span}\{X \oplus JX\}^\perp \cap \mathscr{D}$.

We directly obtain the following corollary for strongly pseudo-convex CR manifolds (that is, for $Q = 0$). Notice that this condition is strictly weaker than Sasakian.

Corollary 5.6. *Consider a complete, strongly pseudo-convex CR structure of dimension $2d+1$, with $d > 1$, such that, for any horizontal unit vector X*

$$\mathrm{Ric}(X) - R(X, JX, JX, X) \geq (2d-2)\kappa > 0.$$

Then the manifold is compact with sub-Riemannian diameter not greater than $\pi/\sqrt{\kappa}$, and the fundamental group is finite.

Example 5.7. The classical example is the Hopf fibration $S^1 \hookrightarrow \mathbb{S}^{2d+1} \xrightarrow{\pi} \mathbb{CP}^d$, where the sub-Riemannian structure on $\mathbb{S}^{2n+1}$ is given by the restriction of the round metric to the orthogonal complement $\mathscr{D}_x := (\ker d_x\pi)^\perp$. In this case $Q = 0$ and Corollary 5.6 applies with $\kappa = 1$. The bound on the sub-Riemannian diameter (equal to π) is then sharp (see also [19]).

In the 3D case (i.e. $d = 1$) $Q = 0$ automatically and $\mathrm{Ric}(X) = R(X, JX, JX, X)$. Then Theorem 5.4 never applies as $\kappa_1 = \kappa_2 = 0$. One can still obtain a Bonnet-Myers result from Theorem 5.2. However, as one can see from Theorem 6.2, the explicit statement in terms of standard tensors results in complicated expressions. This simplifies for Sasakian structures (in any dimension), where also $\tau = 0$. In this case $\mathfrak{Ric}^a = 0$, we apply Theorem 5.2 and we obtain the following result (already proved in [6] for the 3D case and [19] for the general case).

Corollary 5.8. *Consider a complete, Sasakian structure of dimension $2d+1$, with $d \geq 1$, such that, for any horizontal unit vector X*

$$R(X, JX, JX, X) \geq \kappa > 0.$$

Then the manifold is compact with sub-Riemannian diameter not greater than $2\pi/\sqrt{\kappa}$, and the fundamental group is finite.

For the Hopf fibration $R(X, JX, JX, X) = 4$ and we still obtain the sharp bound on the sub-Riemannian diameter (equal to π), which agrees with Example 5.7.

6. Computation of the canonical curvature

We are ready to compute the canonical frame for an extremal $\lambda(t)$ of a contact sub-Riemannian structure. Recall that $\dim M = 2d + 1$. Fix a normalized covector such that $2H(\lambda) = 1$. Let $\gamma_\lambda(t)$ the associated unit speed geodesic, with tangent vector $\mathsf{T} = \dot{\gamma}_\lambda$. X_0 is the Reeb vector field and $h_0(\lambda) = \langle \lambda, X_0 \rangle$.

Theorem 6.1 (Canonical splitting)**.** *Let $\gamma_\lambda(t)$ be a unit speed geodesic of a contact sub-Riemannian structure. Then the canonical splitting is*

$$T_{\gamma(t)}M = S^a_{\gamma(t)} \oplus S^b_{\gamma(t)} \oplus S^c_{\gamma(t)},$$

where, defining $X_a := X_0 - 2Q(\mathsf{T}, \mathsf{T}) + h_0\mathsf{T}$ and $X_b := -J\mathsf{T}$, we have:

$$\begin{aligned}
S^a_{\gamma(t)} &:= \mathrm{span}\{X_a\}, & \dim S^a_{\gamma(t)} &= 1,\\
S^b_{\gamma(t)} &:= \mathrm{span}\{X_b\}, & \dim S^b_{\gamma(t)} &= 1,\\
S^c_{\gamma(t)} &:= J\mathsf{T}^\perp \cap \mathscr{D}_{\gamma(t)}, & \dim S^c_{\gamma(t)} &= 2d - 1.
\end{aligned}$$

where everything is computed along the geodesic. In particular, $\mathscr{D}_{\gamma(t)} = S^b_{\gamma(t)} \oplus S^c_{\gamma(t)}$ and $\mathsf{T}|_{\gamma(t)} \in S^c_{\gamma(t)}$.

The directional curvature is a symmetric operator $\mathfrak{R}_{\gamma(t)} : T_{\gamma(t)}M \to T_{\gamma(t)}M$. As in Sec. 3.4, for $\alpha = a, b, c$, we denote by $\mathfrak{R}^{\alpha\beta}_{\gamma(t)} : S^\alpha_{\gamma(t)} \to S^\beta_{\gamma(t)}$ the restrictions of the canonical curvature to the appropriate invariant subspace. We suppress the explicit dependence on $\gamma(t)$ from now on.

Let X_a and X_b defined as above, and $\{X_{c_j}\}_{j=2}^{2d}$ be an orthonormal frame for S^c. Without loss of generality we assume $X_{c_{2d}} = \mathsf{T}$. Then $S^c = \mathrm{span}\{X_{c_2}, \ldots, X_{c_{2d-1}}\} \oplus \mathrm{span}\{\mathsf{T}\}$.

Theorem 6.2 (Canonical curvature, case $Q = 0$)**.** *In terms of the above frame,*

$$\begin{aligned}
\mathfrak{R}^{aa} &= g((\nabla^2_{\mathsf{T}}\tau)(\mathsf{T}), J\mathsf{T}) - 5h_0 g((\nabla_{\mathsf{T}}\tau)(\mathsf{T}), \mathsf{T}) - 2g(\tau(\mathsf{T}), \mathsf{T})^2 - 6h_0^2 g(\tau(\mathsf{T}), J\mathsf{T})\\
&\quad - 2\|\tau(\mathsf{T})\|^2 - g((\nabla_{X_0}\tau)(\mathsf{T}), \mathsf{T}),\\
\mathfrak{R}^{ab} &= 0,\\
\mathfrak{R}^{ac} &= R(\mathsf{T}, X_0, X_j, \mathsf{T}) + g((\nabla_{\mathsf{T}}\tau)(\mathsf{T}), X_j) + 2h_0 g(\tau(\mathsf{T}), JX_j)\\
&\quad - [g((\nabla_{\mathsf{T}}\tau)(\mathsf{T}), \mathsf{T}) + 2h_0 g(\tau(\mathsf{T}), J\mathsf{T})] g(\mathsf{T}, X_j),\\
\mathfrak{R}^{bb} &= R(\mathsf{T}, J\mathsf{T}, J\mathsf{T}, \mathsf{T}) - 3g(\tau(\mathsf{T}), J\mathsf{T}) + h_0^2,\\
\mathfrak{R}^{bc} &= -R(\mathsf{T}, J\mathsf{T}, X_j, \mathsf{T}) + g(\tau(\mathsf{T}), X_j) - g(\tau(\mathsf{T}), \mathsf{T}) g(\mathsf{T}, X_j),\\
\mathfrak{R}^{cc} &= \mathcal{S}[R(\mathsf{T}, X_i, X_j, \mathsf{T})] + \frac{1}{4}h_0^2 \left[g(X_i, X_j) - g(X_i, \mathsf{T}) g(X_j, \mathsf{T})\right],
\end{aligned}$$

where $i,j=2,\dots,2d$ and $\mathcal{S}$ denotes symmetrisation. The canonical Ricci curvatures are

$$\begin{aligned}\mathfrak{Ric}^a &= g((\nabla^2_{\mathsf{T}}\tau)(\mathsf{T}),J\mathsf{T}) - 5h_0 g((\nabla_{\mathsf{T}}\tau)(\mathsf{T}),\mathsf{T}) - 2g(\tau(\mathsf{T}),\mathsf{T})^2 - 6h_0^2 g(\tau(\mathsf{T}),J\mathsf{T})\\ &\quad - 2\|\tau(\mathsf{T})\|^2 - g((\nabla_{X_0}\tau)(\mathsf{T}),\mathsf{T}),\\ \mathfrak{Ric}^b &= R(\mathsf{T},J\mathsf{T},J\mathsf{T},\mathsf{T}) - 3g(\tau(\mathsf{T}),J\mathsf{T}) + h_0^2,\\ \mathfrak{Ric}^c &= \mathrm{Ric}(\mathsf{T}) - R(\mathsf{T},J\mathsf{T},J\mathsf{T},\mathsf{T}) + \frac{1}{4}h_0^2(2d-2).\end{aligned}$$

Theorem 6.3 (Canonical curvature, general Q)**.** *In terms of above frame*

$$\begin{aligned}\mathfrak{R}^{bb} &= R(\mathsf{T},J\mathsf{T},J\mathsf{T},\mathsf{T}) + 3\|Q(\mathsf{T},\mathsf{T})\|^2 - 3g(\tau(\mathsf{T}),J\mathsf{T}) + h_0^2,\\ \mathfrak{R}^{bc} &= -R(\mathsf{T},J\mathsf{T},X_j,\mathsf{T}) + g(\tau(\mathsf{T}),X_j) - g(\tau(\mathsf{T}),\mathsf{T})g(\mathsf{T},X_j) + 3g((\nabla_{\mathsf{T}}Q)(\mathsf{T},\mathsf{T}),X_j)\\ &\quad + 8h_0 g(Q(\mathsf{T},\mathsf{T}),JX_j),\\ \mathfrak{R}^{cc} &= \mathcal{S}[R(\mathsf{T},X_i,X_j,\mathsf{T})] + h_0\mathcal{S}[g(\mathsf{T},Q(X_j,X_i))] + \frac{1}{4}h_0^2 g(X_i,X_j)\\ &\quad - \frac{1}{4}g(X_i,h_0\mathsf{T}-2Q(\mathsf{T},\mathsf{T}))g(X_j,h_0\mathsf{T}-2Q(\mathsf{T},\mathsf{T})),\end{aligned}$$

where $i,j=2,\dots,2d$ and $\mathcal{S}$ denotes symmetrisation. The canonical Ricci curvatures are

$$\begin{aligned}\mathfrak{Ric}^b =& R(\mathsf{T},J\mathsf{T},J\mathsf{T},\mathsf{T}) + 3\|Q(\mathsf{T},\mathsf{T})\|^2 - 3g(\tau(\mathsf{T}),J\mathsf{T}) + h_0^2,\\ \mathfrak{Ric}^c =& \mathrm{Ric}(\mathsf{T}) - R(\mathsf{T},J\mathsf{T},J\mathsf{T},\mathsf{T}) + \frac{1}{4}h_0^2(2d-2) - \|Q(\mathsf{T},\mathsf{T})\|^2.\end{aligned}$$

Remark 6.4. As in the Riemannian setting, there is no curvature in the direction of the geodesic, that is $\mathfrak{R}\mathsf{T}=0$.

The rest of the section is devoted to the proofs of Theorems 6.1, 6.2 and 6.3.

6.1. **Lifted frame.** Only in this subsection, M is an n-dimensional manifold, since the construction is general. We define a local frame on T^*M, associated with the choice of a local frame $X_1,\dots,X_n$ on M. All our considerations are local, then we assume that the frame is globally defined. For $\alpha=1,\dots,n$ let $h_\alpha : T^*M \to \mathbb{R}$ be the linear-on-fibers function defined by $h_\alpha(\lambda) := \langle\lambda, X_\alpha\rangle$. The action of 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 $Y \in \Gamma(TM)$, $g \in C^\infty(M)$. Then, we define the *lift of a field* $X \in \Gamma(TM)$ as the field $\widetilde{X} \in \Gamma(T(T^*M))$ such that $\widetilde{X}(h_\alpha) = 0$ for $\alpha=1,\dots,n$ and $\widetilde{X}(\pi^*g) = X(g)$. This, together with Leibniz's rule, characterize the action of $\widetilde{X}$ on affine functions, and completely define $\widetilde{X}$. Observe that $\pi_*\widetilde{X} = X$. On the other hand, we define the (vertical) fields ∂_{h_α} such that $\partial_{h_\alpha}(\pi^*g) = 0$, and $\partial_{h_\alpha}(h_\beta) = \delta_{\alpha\beta}$. It is easy to check that $\{\partial_{h_\alpha}, \widetilde{X}_\alpha\}_{\alpha=1}^n$ is a local frame on T^*M. We call such a frame the *lifted frame.*

Remark 6.5. Let $(q_1,\dots,q_n)$ be local coordinates on M, and let $X_\alpha = \partial_{q_\alpha}$. In this case, our construction gives the usual frame $\{\partial_{q_\alpha}, \partial_{p_\alpha}\}_{\alpha=1}^n$ on T^*M associated with the dual coordinates $(q_1,\dots,q_n,p_1,\dots,p_n)$ on T^*M. The construction above is then a generalization of this classical construction to non-commuting local frames.

Any fixed lifted frame defines a splitting

$$T_\lambda(T^*M) = \mathrm{span}_\lambda\{\widetilde{X}_1,\dots,\widetilde{X}_n\} \oplus \mathrm{span}_\lambda\{\partial_{h_1},\dots,\partial_{h_n}\}.$$

Since $\mathrm{span}_\lambda\{\partial_{h_1},\dots,\partial_{h_n}\} = \ker \pi_*|_\lambda$ is the vertical subspace of $T_\lambda(T^*M)$, we associate, with any vector $\xi \in T_\lambda(T^*M)$ its *horizontal part* $\xi^h \in T_xM$ and its *vertical part* $\xi^v \in T_x^*M$:

$$\xi^h := \pi_*\xi, \qquad \xi^v := \xi - \widetilde{\pi_*\xi},$$

where here, as we will always do in the following, we used the canonical identification $\ker \pi_*|_\lambda = T_\lambda(T_x^*M) \simeq T_x^*M$, where $x = \pi(\lambda)$. We stress that the concept of vertical part introduced here makes sense w.r.t. a fixed frame, has no invariant meaning, and is only a tool for computations.

Let $\nu_1, \dots, \nu_n$ be the dual frame of $X_1, \dots, X_n$, i.e. $\nu_\alpha(X_\beta) = \delta_{\alpha\beta}$. In the following, we use the notation $\partial_\alpha := \partial_{h_\alpha}$ and we suppress the tilde from $\widetilde{X}$; the meaning will be clear from the context. Then the symplectic form is written as

$$\sigma = \sum_{\alpha=1}^{n} dh_\alpha \wedge \pi^* \nu_\alpha + h_\alpha \pi^* d\nu_\alpha. \tag{22}$$

We say that a vector $V_\lambda \in T_\lambda(T^*M)$ is *vertical* if $\pi_* V = 0$. In this case $V_\lambda \in T_\lambda(T^*_{\pi(\lambda)}M) \simeq T^*_{\pi(\lambda)}M$ can be identified with a covector. We have the following useful lemma.

Lemma 6.6. *Let $V, W \in T_\lambda(T^*M)$. If V is vertical, then*

$$\sigma_\lambda(V, W) = \langle V, \pi_* W \rangle.$$

Using Eq. (22), the Hamiltonian vector field associated with h_α has the form

$$\vec{h}_\alpha = X_\alpha + \sum_{\delta=1}^{n} c^\delta_{\alpha\beta} h_\delta \partial_\beta, \qquad \alpha, \beta = 1, \dots, n.$$

Since $H = \frac{1}{2}\sum_{i=1}^k h_i^2$, then $\vec{H} = \sum_{i=1}^k h_i \vec{h}_i$, where $k = \operatorname{rank} \mathscr{D}$.

6.2. **Some useful formulas.** In what follows ∇ will always denote Tanno connection. Let $X_1, \dots, X_{2d}$ be an orthonormal frame for $\mathscr{D}$ and X_0 be the Reeb field. The *structural functions* are defined by

$$[X_\alpha, X_\beta] = \sum_{\delta=0}^{2d} c^\delta_{\alpha\beta} X_\delta, \qquad \alpha, \beta = 0, \dots, 2d.$$

Notice that $c^0_{\alpha 0} = \omega([X_\alpha, X_0]) = d\omega(X_\alpha, X_0) = 0$. We define "horizontal" Christoffel symbols:

$$\Gamma^k_{ij} := \frac{1}{2}\left(c^k_{ij} + c^j_{ki} + c^i_{kj}\right), \qquad i, j, k = 1, \dots, 2d.$$

Notice that $\Gamma^k_{ij} + \Gamma^j_{ik} = 0$ and that one can recover some of the structural functions with the relation

$$c^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}.$$

In terms of the structural functions, we have, for $i, j, k = 1, \dots, 2d$

$$\nabla_{X_i} X_j = \sum_{k=1}^{2d} \Gamma^k_{ij} X_k, \qquad \nabla_{X_0} X_i = \frac{1}{2}\sum_{k=1}^{2d} (c^i_{k0} - c^k_{i0}) X_k,$$

$$\tau(X_i) = \frac{1}{2}\sum_{k=1}^{2d} (c^i_{k0} + c^k_{i0}) X_k, \qquad T(X_j, X_k) = -c^0_{jk} X_0, \qquad JX_i = \sum_{j=1}^{2d} c^0_{ij} X_j.$$

Lemma 6.7. *Let $\lambda \in T^*M$ an initial covector, associated with the geodesic $\gamma_\lambda(t) = \pi(\lambda(t))$. In terms of Tanno covariant derivative, along the geodesic $\gamma_\lambda(t)$, we have*

$$\begin{aligned} \nabla_{\mathsf{T}} \mathsf{T} &= h_0 J \mathsf{T}, \\ \nabla_{\mathsf{T}} (J\mathsf{T}) &= -h_0 \mathsf{T} + Q(\mathsf{T}, \mathsf{T}), \\ \nabla_{\mathsf{T}} h_0 &= g(\tau(\mathsf{T}), \mathsf{T}), \end{aligned}$$

where $h_0 = h_0(\lambda(t)) = \langle \lambda(t), X_0 \rangle$ and $\mathsf{T} := \dot{\gamma}_\lambda$.

Proof. Along the geodesic, $\mathsf{T}|_{\gamma(t)} = \sum_{i=1}^{2d} h_i(t) X_i|_{\gamma(t)}$, where $h_i(t) = h_i(\lambda(t)) = \langle \lambda(t), X_i \rangle$. Then

$$\begin{aligned}\nabla_{\mathsf{T}}\mathsf{T} &= \sum_{i=1}^{2d} \dot{h}_i X_i + \sum_{i,j=1}^{2d} h_j h_i \nabla_{X_i} X_j = \sum_{i=1}^{2d} \{H, h_i\} X_i + \sum_{i,j,k=1}^{2d} h_j h_i \Gamma_{ij}^k X_k \\ &= \sum_{\alpha=0}^{2d} \sum_{i,j=1}^{2d} h_j c_{ji}^{\alpha} h_\alpha X_i + \sum_{i,j,k=1}^{2d} h_j h_i \Gamma_{ij}^k X_k = h_0 \sum_{j=1}^{2d} h_j c_{ji}^0 X_i = \sum_{j=1}^{2d} h_0 h_j J X_j = h_0 J \mathsf{T}.\end{aligned}$$

where we used the Hamilton's equation for normal geodesics. The second identity follows from the first and the definition of Q. For the third identity, we use again Hamilton's equations:

$$\nabla_{\mathsf{T}} h_0 = \dot{h}_0 = \{H, h_0\} = \sum_{i,j=1}^{2d} h_i c_{i0}^j h_j = \frac{1}{2} \sum_{i,j=1}^{2d} h_i h_j (c_{i0}^j + c_{j0}^i) = g(\tau(\mathsf{T}), \mathsf{T}). \qquad \square$$

Lemma 6.8. *Let X, Y, Z be vector fields. Then*

a) $Q(X, Y)$ *is horizontal,*
b) $Q(JX, Y) = -JQ(X, Y)$,
c) $Q(X, Y) \perp X$,
d) $g(X, Q(Y, Z)) = -g(Y, Q(X, Z))$ *(skew-symmetry w.r.t. the first two arguments),*
e) $Q(X, Y) \perp JX$,
f) $\mathfrak{S} g(T(X, Y), JZ) = 0$, *where $\mathfrak{S}$ denotes the cyclic sum,*
g) $\mathfrak{S} g(Y, Q(Z, X)) = 0$,
h) $Q(Y, JY) = Q(JY, Y) = -JQ(Y, Y)$,
i) $\operatorname{tr} Q = 0$.

Proof. To prove a), observe that, since $JX_0 = 0$, it follows that $Q(X_0, Y) = 0$ for any Y. Then we can assume w.l.o.g. that X is horizontal. Clearly JX is horizontal too. Then $\omega(JX) = 0$. Therefore, since ω is parallel for Tanno connection, $0 = \omega(\nabla_Y(J)X + J\nabla_Y X) = \omega(Q(X, Y))$, where we used the fact that $\nabla_Y X$ is horizontal. To prove b), observe first that if $X = X_0$ the identity is trivially true. Then, w.l.o.g. we assume that X is horizontal. Consider the covariant derivative of the identity $J^2 X = -X$. Then we obtain $(\nabla_Y J)(JX) + J(\nabla_Y J)X + J^2 \nabla_Y X = -\nabla_Y X$, which implies b), by definition of Q. To prove c), observe that $X \perp JX$ by definition. Then $g(JX, X) = 0$. By taking the covariant derivative we obtain $g(Q(X, Y), X) + g(J\nabla_Y X, X) + g(JX, \nabla_Y X) = 0$, which implies c). To prove d), observe that the identity is trivially true when $X = X_0$ or $Y = X_0$. Then we may assume w.l.o.g. that X, Y are horizontal. Then we take the covariant derivative in the direction Z of the identity $g(JX, Y) = -g(X, JY)$, and we obtain the result. Point e) follows from d) and b). To prove f), we have

$$\mathfrak{S} g(T(X, Y), JZ) = g(T(X, Y), JZ) + g(T(Z, X), JY) + g(T(Y, Z), JX).$$

If $X, Y, Z \in \mathscr{D}$, then all the terms are trivially zero. Then assume w.l.o.g. that $X = X_0$. Then

$$\begin{aligned}\mathfrak{S} g(T(X_0, Y), JZ) &= g(T(X_0, JY), Z) - g(T(X_0, JZ), Y) + g(T(Y, Z), \cancel{JX_0}) \\ &= g(\tau(JY), Z) - g(\tau(JZ), Y) = 0,\end{aligned}$$

where we used the fact that τ is symmetric and $\tau \circ J + J \circ \tau = 0$. To prove g), consider the following identity for the differential of a 2-forms α:

$$d\alpha(X, Y, Z) = \mathfrak{S}(X(\alpha(Y, Z)) - \mathfrak{S}\alpha([X, Y], Z).$$

We apply it to the two form $\alpha = d\omega$. Since $d\alpha = d^2\omega = 0$, we have

$$\begin{aligned}0 = d^2\omega &= \mathfrak{S}\left[X(d\omega(Y, Z)) - d\omega([X, Y], Z)\right] = \mathfrak{S}\left[X(g(Y, JZ)) - g([X, Y], JZ)\right] \\ &= \mathfrak{S}\left[\cancel{g(\nabla_X Y, JZ)} + g(Y, Q(Z, X)) + g(Y, J\nabla_X Z) - g(\cancel{\nabla_X Y} - \nabla_Y X, JZ)\right] \\ &\quad + \cancel{\mathfrak{S} g(T(X, Y), JZ)} \\ &= \mathfrak{S}\left[g(Y, Q(Z, X)) - \cancel{g(\nabla_X Z), JY)} + \cancel{g(\nabla_Y X, JZ)}\right].\end{aligned}$$

where we used the definition of J and property f). Point h) follows from g), e), d), and b). For i) observe that $Q(X_0, X_0) = 0$. Then, for every orthonormal frame $X_1, \dots, X_{2d}$ of $\mathscr{D}$, we have

$$\operatorname{tr} Q = \sum_{i=1}^{2d} Q(X_i, X_i) = -\sum_{i=1}^{2d} J^2 Q(X_i, X_i) = \sum_{i=1}^{2d} JQ(JX_i, X_i) = -\sum_{i=1}^{2d} Q(JX_i, JX_i) = -\operatorname{tr} Q,$$

where we used properties a), h) and b). □

Tanno connection has torsion. Then the curvature tensor

$$R(X, Y, Z, W) := g(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, W)$$

is not symmetric w.r.t. to exchange of the first and second pair of arguments.

Lemma 6.9. *For any $X, Y, Z \in \mathscr{D}$ we have*

$$\begin{aligned} R(X, Z, Y, X) = R(X, Y, Z, X) &+ g(X, JZ) g(X, \tau(Y)) + g(Z, JY) g(X, \tau(X)) \\ &+ g(Y, JX) g(X, \tau(Z)). \end{aligned} \tag{24}$$

In particular, if $Y, Z \in JX^\perp \cap \mathscr{D}$ we have

$$R(X, Z, Y, X) = R(X, Y, Z, X) + g(Z, JY) g(X, \tau(X)).$$

Moreover, for Tanno's tensor we have

$$\mathcal{S} g(X, Q(X_i, X_j)) = \frac{1}{2} g(X_j, Q(X_i, X)) - g(X_j, Q(X, X_i)), \tag{25}$$

where $\mathcal{S}$ denotes the symmetrization w.r.t. the displayed indices.

Proof. The first identity is a consequence of first Bianchi identity for connections with torsion:

$$\mathfrak{S}(R(X, Y)Z) = \mathfrak{S}(T(T(X, Y), Z)) + (\nabla_X T)(Y, Z),$$

where $\mathfrak{S}$ denotes the cyclic sum. If $X, Y, Z \in \mathscr{D}$, for Tanno connection we obtain

$$\begin{aligned} \mathfrak{S}(R(X, Y)Z)) &= \mathfrak{S}\left[g(X, JY)\tau(Z) + \nabla_X (T(Y, Z)) - T(\nabla_X Y, Z) - T(Y, \nabla_X Z)\right] \\ &= \mathfrak{S}\left[g(X, JY)\tau(Z) + \nabla_X (g(Y, JZ)) X_0 - g(\nabla_X Y, JZ) X_0 - g(Y, J\nabla_X Z) X_0\right] \\ &= \mathfrak{S}\left[g(X, JY)\tau(Z)\right] + \cancel{\mathfrak{S} g(Y, Q(Z, X))}, \end{aligned}$$

where we used property g) of Tanno's tensor. If we take the scalar product with X, we get

$$R(X, Y, Z, X) + R(Y, Z, X, X) + R(Z, X, Y, X) = g(X, \mathfrak{S} g(X, JY)\tau(Z)).$$

Tanno's curvature is still skew symmetric in the first and last pairs of indices, and we get

$$\begin{aligned} R(X, Y, Z, X) - R(X, Z, Y, X) = g(X, JY) g(X, \tau(Z)) &+ g(Y, JZ) g(X, \tau(X)) \\ &+ g(Z, JX) g(X, \tau(Y)). \end{aligned}$$

This proves (24). For (25), using property g) of Tanno's tensor, we get

$$\frac{1}{2} g(X_j, Q(X_i, X)) - g(X_j, Q(X, X_i)) = \frac{1}{2}\left[g(X, Q(X_i, X_j)) + g(X, Q(X_j, X_i))\right]. \qquad \square$$

6.3. **Computation of the curvatures.** We choose a conveniently adapted frame along the geodesic, of which we will consider the lift.

Definition 6.10. A local frame $X_0, X_1, \dots, X_{2d}$ is *adapted* if

- X_0 is Reeb vector field,
- X_{2d} is an horizontal extension of $\dot\gamma_\lambda(t)$, namely $X_{2d} \in \mathscr{D}$ and $\dot\gamma_\lambda(t) = X_{2d}(\gamma_\lambda(t))$,
- $X_1 = -JX_{2d}$,
- $X_2, \dots, X_{2d-1} \in \operatorname{span}(X_{2d}, JX_{2d})^\perp$ are orthonormal.

In particular, $X_0, X_1, \dots, X_{2d}$ is an orthonormal frame for the Riemannian extension of g. From now on the last index $2d$ plays a different role, since it is constrained to be an horizontal extension of the tangent vector of the fixed geodesic.

Definition 6.11. An adapted frame $X_0, X_1, \dots, X_{2d}$ is *parallel transported* if

$$\nabla_{\mathsf{T}} X_j = \frac{1}{2}(h_0 J X_j + a_j J X_{2d}), \qquad j = 2, \dots, 2d \tag{26}$$

along the geodesic $\gamma_\lambda(t)$, where

$$a_j := g(h_0 X_{2d} - 2Q(X_{2d}, X_{2d}), X_j).$$

As one can readily check, an adapted, parallel transported frame is unique up to constant rotation of $X_2, \dots, X_{2d-1}$. Notice that a_j and h_0 are well defined functions on the curve $\gamma_\lambda(t)$ once the initial covector λ is fixed.

6.3.1. *The algorithm.* To compute the canonical frame and curvatures, we follow the general algorithm developed in [29]. The elements of the canonical frame and the curvature matrices will be computed in the following order:

$$E_a \to E_b \to F_b \to E_c \to F_c \to R_{bb}, R_{cc}, R_{bc} \to F_a \to R_{aa}, R_{ac}.$$

We choose an adapted parallel transported frame $X_0, X_1, \dots, X_{2d}$ as defined above, and the associated lifted frame $\{X_\alpha, \partial_\alpha\}_{\alpha=0}^{2d}$. Along the extremal, we have

$$h_i = \langle \lambda, X_i \rangle = g(X_{2d}, X_i) = \delta_{i,2d}, \qquad i = 1, \dots, 2d.$$

We have the following simplifications for some structural functions, for $\alpha = 0, \dots, 2d$:

$$\begin{aligned} c^0_{1,\alpha} &= \omega([X_1, X_\alpha]) = -d\omega(X_1, X_\alpha) = -g(X_1, JX_\alpha) = g(JX_{2d}, JX_\alpha) = \delta_{2d,\alpha}, \\ c^0_{2d,\alpha} &= \omega([X_{2d}, X_\alpha]) = -d\omega(X_{2d}, X_\alpha) = -g(X_{2d}, JX_\alpha) = -g(X_1, X_\alpha) = \delta_{1,\alpha}. \end{aligned}$$

From here, repeated Latin indices are summed from 1 to $2d$, and Greek ones from 0 to $2d$.

6.3.2. *Preliminary computations.* The Lie derivative w.r.t. $\vec{H}$ is denoted with a dot, namely $\dot{E}_\alpha = [\vec{H}, E_\alpha]$. We compute it for the basic elements of the lifted frame $\{X_\alpha, \partial_\alpha\}_{\alpha=0}^{2d}$. Here, we use the shorthand $\nabla_\alpha = \nabla_{X_\alpha}$ for $\alpha = 0, \dots, 2d$.

Lemma 6.12. *For any adapted frame (not necessarily parallel transported), we have*

$$\begin{aligned} \dot{\partial}_0 &= \partial_1, \\ \dot{\partial}_i &= -X_i - 2g(\tau(X_{2d}), X_i)\partial_0 + h_0 g(JX_j, X_i)\partial_j - g(\nabla_{2d} X_j, X_i)\partial_j - g(\nabla_i X_j, X_{2d})\partial_j, \\ \dot{X}_i &= \nabla_{2d} X_i - \nabla_i X_{2d} - \delta_{i1} X_0 - \nabla_i g(\tau(X_{2d}), X_{2d})\partial_0 - \nabla_i g(X_{2d}, \nabla_{2d} X_j)\partial_j, \\ \dot{X}_0 &= -\nabla_0 X_{2d} + \tau(X_{2d}) - \nabla_0 g(\tau(X_{2d}), X_{2d})\partial_0 - \nabla_0 g(X_{2d}, \nabla_{2d} X_j)\partial_j. \end{aligned}$$

Proof. The proof is a routine computation, using the properties of the adapted frame.

$$\dot{\partial}_0 = [\vec{H}, \partial_0] = [h_i \vec{h}_i, \partial_0] = h_i[X_i + c^\beta_{i\alpha} h_\beta \partial_\alpha, \partial_0] = c^0_{2d,\alpha}\partial_\alpha = \partial_1.$$

For $i = 1, \dots, 2d$, we have

$$\begin{aligned} \dot{\partial}_i = [\vec{H}, \partial_i] = [h_j \vec{h}_j, \partial_i] &= -\vec{h}_i - c^i_{2d,\alpha}\partial_\alpha \\ &= -X_i - c^\beta_{i\alpha} h_\beta \partial_\alpha - c^i_{2d,\alpha}\partial_\alpha \\ &= -X_i - c^{2d}_{i0}\partial_0 - c^{2d}_{ij}\partial_j - c^0_{ij} h_0 \partial_j - c^i_{2d,0}\partial_0 - c^i_{2d,j}\partial_j \\ &= -X_i - (c^{2d}_{i0} + c^i_{2d,0})\partial_0 - c^0_{ij} h_0 \partial_j - (c^i_{2d,j} + c^{2d}_{ij})\partial_j \\ &= -X_i - 2g(\tau(X_{2d}), X_i)\partial_0 - h_0 g(X_j, JX_i)\partial_j - (\Gamma^i_{2d,j} + \Gamma^{2d}_{ij})\partial_j \\ &= -X_i - 2g(\tau(X_{2d}), X_i)\partial_0 + h_0 g(JX_j, X_i)\partial_j - g(\nabla_{2d} X_j, X_i)\partial_j - g(\nabla_i X_j, X_{2d})\partial_j. \end{aligned}$$

Moreover, for $\mu = 0, \dots, 2d$

$$\begin{aligned}
\dot{X}_\mu = [\vec{H}, X_\mu] &= [h_i \vec{h}_i, X_\mu] = [\vec{h}_{2d}, X_\mu] \\
&= [X_{2d} + c^{\beta}_{2d,\alpha} h_\beta \partial_\alpha, X_\mu] \\
&= [X_{2d}, X_\mu] - X_\mu(c^{\beta}_{2d,\alpha}) h_\beta \partial_\alpha \\
&= [X_{2d}, X_\mu] - X_\mu(c^{2d}_{2d,0})\partial_0 - X_\mu(c^{2d}_{2d,j})\partial_j - X_\mu(c^{0}_{2d,j}) h_0 \partial_j \\
&= \nabla_{2d} X_\mu - \nabla_\mu X_{2d} - T(X_{2d}, X_\mu) - X_\mu(g(\tau(X_{2d}), X_{2d}))\partial_0 \\
&\quad - X_\mu(g(\nabla_{2d} X_j, X_{2d}))\partial_j,
\end{aligned}$$

and we obtain the result using the properties of the torsion of Tanno connection. □

Step 1: E_a. The element E_a is uniquely determined by the following conditions:

(i) $\pi_* E_a = 0$,
(ii) $\pi_* \dot{E}_a = 0$,
(iii) $2H(\dot{E}_a) = 1$ (equivalent to $\sigma(\ddot{E}_a, \dot{E}_a) = 1$).

The first condition implies $E_a = v_\alpha \partial_\alpha$ for some smooth function v_α along the extremal. The second condition, using Lemma 6.12, implies $E_a = v_0 \partial_0$. To apply the third condition, observe that

$$\dot{E}_a = \dot{v}_0 \partial_0 + v_0 \dot{\partial}_0 = \dot{v}_0 \partial_0 + v_0 \partial_1.$$

By standard identifications, $2H(\dot{E}_a) = v_0^2 \|X_1\|^2 = v_0^2$, and $v_0 = \pm 1$ (we choose the "+" sign).

Step 2: E_b. E_a directly determines E_b through the structural equation:

$$E_b = \dot{E}_a = \dot{\partial}_0 = \partial_1.$$

Step 3: F_b. E_b directly determines F_b through the structural equation:

$$\begin{aligned}
F_b = \dot{E}_b &= -\dot{\partial}_1 \\
&= X_1 + 2g(\tau(X_{2d}), X_1)\partial_0 + h_0 g(JX_1, X_j)\partial_j - g(\nabla_{2d} X_1, X_j)\partial_j - g(\nabla_1 X_{2d}, X_j)\partial_j \\
&= X_1 + 2g(\tau(X_{2d}), X_1)\partial_0 + \cancel{h_0 g(X_{2d}, X_j)\partial_j} + g(Q(X_{2d}, X_{2d}) - \cancel{h_0 X_{2d}}, X_j)\partial_j \\
&\quad - g(\nabla_1 X_{2d}, X_j)\partial_j \\
&= X_1 + 2g(\tau(X_{2d}), X_1)\partial_0 + g(Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}, X_j)\partial_j.
\end{aligned}$$

Where we used the equation $\nabla_{2d} X_{2d} = h_0 J X_{2d}$.

Intermediate step: $\dot{F}_b$. We now compute $\dot{F}_b$. We use Leibniz's rule, Lemma 6.12 and we obtain:

$$\begin{aligned}
\dot{F}_b &= \dot{X}_1 + 2\nabla_{2d} g(\tau(X_{2d}), X_1)\partial_0 + 2g(\tau(X_{2d}), X_1)\dot{\partial}_0 + \nabla_{2d} g(Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}, X_j)\partial_j \\
&\quad + g(Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}, X_j)\dot{\partial}_j \\
&= \nabla_{2d} X_1 - \nabla_1 X_{2d} - X_0 - \nabla_1 g(\tau(X_{2d}), X_{2d})\partial_0 - \nabla_1 g(X_{2d}, \nabla_{2d} X_j)\partial_j \\
&\quad + 2\nabla_{2d} g(\tau(X_{2d}), X_1)\partial_0 + 2g(\tau(X_{2d}), X_1)\partial_1 + \nabla_{2d} g(Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}, X_j)\partial_j \\
&\quad - Q(X_{2d}, X_{2d}) + \nabla_1 X_{2d} - 2g(\tau(X_{2d}), Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d})\partial_0 \\
&\quad + h_0 g(JX_j, Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d})\partial_j - g(\nabla_{2d} X_j, Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d})\partial_j \\
&\quad - g(\nabla_{Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}} X_j, X_{2d})\partial_j \\
&= \nabla_{2d} X_1 - X_0 - Q(X_{2d}, X_{2d}) + 2\nabla_{2d} g(\tau(X_{2d}), X_1)\partial_0 - \nabla_1 g(\tau(X_{2d}), X_{2d})\partial_0 \\
&\quad - 2g(\tau(X_{2d}), Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d})\partial_0 + 2g(\tau(X_{2d}), X_1)\partial_1 - \nabla_1 g(X_{2d}, \nabla_{2d} X_j)\partial_j \\
&\quad + \nabla_{2d} g(Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}, X_j)\partial_j + h_0 g(JX_j, Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d})\partial_j \\
&\quad - g(\nabla_{2d} X_j, Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d})\partial_j - g(\nabla_{Q(X_{2d}, X_{2d}) - \nabla_1 X_{2d}} X_j, X_{2d})\partial_j.
\end{aligned}$$

Step 4: E_{c_j}. The elements E_{c_j}, for $j = 2, \dots, 2d$ are uniquely determined by the conditions:

(i) They generate the skew-orthogonal complement of $\mathrm{span}\{E_a, E_b, F_b, \dot{F}_b\}$,
(ii) Darboux property: $\sigma(\dot{E}_{c_i}, E_{c_j}) = \delta_{ij}$,
(iii) Their derivative $F_{c_j} = -\dot{E}_{c_j}$ generate an isotropic subspace (or, equivalently, assuming the two points above, $\pi_* \ddot{E}_{c_j} = 0$).

Since $\pi_* E_{c_j} = 0$ we have

$$E_{c_j} = \alpha_{ij}\partial_i + \beta_j \partial_0, \qquad j = 2, \dots, 2d,$$

for some smooth functions α_{ij} and β_j along the extremal. Observe that α is a $2d \times (2d-1)$ matrix. In order to apply condition (i) we compute $\pi_* \dot{F}_b$. From the previous step we get

$$\pi_* \dot{F}_b = \nabla_{2d} X_1 - X_0 - Q(X_{2d}, X_{2d}) = -2Q(X_{2d}, X_{2d}) + h_0 X_{2d} - X_0.$$

Then we apply (i) and we obtain:

$$0 = \sigma(F_b, E_{c_j}) = -\langle E_{c_j}, \pi_* F_b \rangle = -\langle E_{c_j}, X_1 \rangle = -\alpha_{1j}.$$

Thus the first row of α is zero. Moreover

$$\begin{aligned} 0 = \sigma(\dot{F}_b, E_{c_j}) = -\langle E_{c_j}, \pi_* \dot{F}_b \rangle &= \langle E_{c_j}, X_0 + 2Q(X_{2d}, X_{2d}) - h_0 X_{2d} \rangle \\ &= \beta_j + \alpha_{ij} g(2Q(X_{2d}, X_{2d}) - h_0 X_{2d}, X_i). \end{aligned}$$

Denoting $a_i := g(h_0 X_{2d} - 2Q(X_{2d}, X_{2d}), X_i)$, for all $i = 1, \dots, 2d$ we get

$$E_{c_j} = \alpha_{ij}(\partial_i + a_i \partial_0), \qquad j = 2, \dots, 2d.$$

Condition (iii) gives

$$\begin{aligned} 0 = \pi_* \ddot{E}_{c_j} &= \alpha_{ij} \pi_* \left(\ddot{\partial}_i + \ddot{a}_i \partial_0 + 2\dot{a}_i \dot{\partial}_0 + a_i \ddot{\partial}_0 \right) + 2\dot{\alpha}_{ij} \pi_* \left(\dot{\partial}_i + \dot{a}_i \partial_0 + a_i \dot{\partial}_0 \right) \\ &= \alpha_{ij} \left(-\nabla_{2d} X_i + \cancel{\nabla_i X_{2d}} + h_0 J X_i - \nabla_{2d} X_i - \cancel{\nabla_i X_{2d}} - a_i X_1 \right) - 2\dot{\alpha}_{ij} X_i \\ &= -2\dot{\alpha}_{ij} X_i, \end{aligned}$$

where in the last line we used we used our choice of a parallel transported frame. Then α_{ij} is a constant $2d \times (2d-1)$ matrix. Since the first row is zero, we can consider α as a $(2d-1) \times (2d-1)$ constant matrix. Condition (ii) implies that this matrix is orthogonal. Without loss of generality, we may assume that $\alpha_{ij}(0) = \delta_{ij}$. Thus

$$E_{c_j} = \partial_j + a_j \partial_0, \qquad j = 2, \dots, 2d.$$

Step 5: F_{c_j}. For $j = 2, \dots, 2d$, the element F_{c_j} is obtained from E_{c_j} by the structural equations:

$$\begin{aligned} F_{c_j} = -\dot{E}_{c_j} &= -\dot{\partial}_j - \dot{a}_j \partial_0 - a_j \dot{\partial}_0 \\ &= X_j + 2g(\tau(X_{2d}), X_j)\partial_0 - h_0 g(JX_i, X_j)\partial_i + g(\nabla_{2d} X_i, X_j)\partial_i + g(\nabla_j X_i, X_{2d})\partial_i \\ &\quad - \dot{a}_j \partial_0 - a_j \partial_1 \\ &= X_j + \left(2g(\tau(X_{2d}), X_j) - \dot{a}_j\right)\partial_0 - h_0 g(JX_i, X_j)\partial_i - g(X_i, \nabla_{2d} X_j)\partial_i \\ &\quad + g(\nabla_j X_i, X_{2d})\partial_i - a_j \partial_1 \\ &= X_j + \left(2g(\tau(X_{2d}), X_j) - \dot{a}_j\right)\partial_0 - h_0 g(JX_i, X_j)\partial_i - \frac{1}{2} g(X_i, h_0 J X_j - a_j X_1)\partial_i \\ &\quad + g(\nabla_j X_i, X_{2d})\partial_i - a_j \partial_1 \\ &= X_j + \left(2g(\tau(X_{2d}), X_j) - \dot{a}_j\right)\partial_0 + g(\nabla_j X_i, X_{2d})\partial_i + \frac{1}{2} h_0 g(X_i, JX_j)\partial_i - \frac{1}{2} a_j \partial_1, \end{aligned}$$

where we used again the parallel transported frame to eliminate $\nabla_{2d} X_j$ for $j = 2, \dots, 2d$.

The computations so far prove Theorem 6.1. In fact the canonical subspaces are defined by

$$S^a := \mathrm{span}\{\pi_* F_a\}, \qquad S^b := \mathrm{span}\{\pi_* F_b\}, \qquad S^c := \mathrm{span}\{\pi_* F_{c_2}, \dots, \pi_* F_{c_{2d}}\},$$

(see Sec. 3.4) and from the computations above we have explicitly:

$$\begin{aligned} X_a &:= \pi_* F_a = -\pi_* \dot{F}_b = X_0 + 2Q(\mathsf{T},\mathsf{T}) - h_0 T = X_0 - a_j X_j, \\ X_b &:= \pi_* F_b = -J\mathsf{T}, \\ X_{c_j} &:= \pi_* F_{c_j} = X_j, \qquad j = 2,\dots,2d. \end{aligned}$$

Step 6: $\mathfrak{R}^{bb}$, $\mathfrak{R}^{bc}$ and $\mathfrak{R}^{cc}$. To compute the curvature we need to compute symplectic products.

Lemma 6.13. *Let $X_0,\dots,X_{2d}$ an adapted frame. Then, along the extremal*

$$\begin{aligned} \sigma(X_\mu, X_\nu) &= h_0\omega(T(X_\mu,X_\nu)) - g(X_{2d}, \nabla_\mu X_\nu - \nabla_\nu X_\mu - T(X_\mu,X_\nu)), \\ \sigma(\partial_\mu, X_\nu) &= \delta_{\mu\nu}, \\ \sigma(\partial_\mu,\partial_\nu) &= 0. \end{aligned}$$

for all $\mu,\nu = 0,\dots,2d$.

Proof. From Eq. (22), we obtain, along the extremal

$$\begin{aligned} \sigma(X_\mu,X_\nu) &= h_\alpha d\nu^\alpha(X_\mu,X_\nu) = -h_\alpha \nu^\alpha([X_\mu,X_\nu]) \\ &= -h_\alpha\nu^\alpha(\nabla_\mu X_\nu - \nabla_\nu X_\mu - T(X_\mu,X_\nu)) \\ &= h_0\omega(T(X_\mu,X_\nu)) - g(X_{2d}, \nabla_\mu X_\nu - \nabla_\nu X_\mu - T(X_\mu,X_\nu)). \end{aligned}$$

where, in the last line, we used the fact that the frame is adapted. The second and third identities follow with analogous but shorter computations starting from Eq. (22). □

Now we can compute $\mathfrak{R}^{bb}$. It is convenient to split the computation as follows.

$$\mathfrak{R}^{bb} = \sigma(\dot{F}_b, F_b) = \sigma(\dot{F}_b^h, F_b^h) + \sigma(\dot{F}_b^h, F_b^v) + \sigma(\dot{F}_b^v, F_b^h).$$

We compute the three pieces, using Lemma 6.13. We obtain, after some computations

$$\begin{aligned} \sigma(\dot{F}_b^h, F_b^h) &= -g(\tau(X_{2d}), X_1) + h_0^2 + g(X_{2d}, \nabla_{2Q(X_{2d},X_{2d}) - h_0X_{2d} + X_0} X_1) \\ &\quad + g(\nabla_1 X_{2d}, 2Q(X_{2d},X_{2d})), \\ \sigma(\dot{F}_b^v, F_b^h) &= 2g(\tau(X_{2d}), X_1) + g(\nabla_{2d}\nabla_1 X_1 - \nabla_1\nabla_{2d}X_1 . X_{2d}) + \|Q(X_{2d},X_{2d})\|^2 \\ &\quad - g(\nabla_{Q(X_{2d},X_{2d}) - \nabla_1 X_{2d}} X_1, X_{2d}), \\ \sigma(\dot{F}_b^h, F_b^v) &= 2g(\tau(X_{2d}), X_1) + 2\|Q(X_{2d},X_{2d})\|^2 - g(\nabla_1 X_{2d}, 2Q(X_{2d},X_{2d})). \end{aligned}$$

Taking in account that $Q(X_{2d},X_{2d}) - h_0X_{2d} + \nabla_1 X_{2d} + X_0 = -[X_{2d},X_1]$ (easily proved using the properties of Tanno connection) we obtain immediately

$$\mathfrak{R}^{bb} = R(\mathsf{T}, J\mathsf{T}, J\mathsf{T}, \mathsf{T}) + 3\|Q(\mathsf{T},\mathsf{T})\|^2 - 3g(\tau(\mathsf{T}), J\mathsf{T}) + h_0^2.$$

where we replaced the explicit tangent vector $\mathsf{T} = \dot{\gamma}_\lambda$. We proceed with $\mathfrak{R}^{cc}$. Again we split:

$$\mathfrak{R}^{cc}_{ij} = \sigma(\dot{F}_{c_i}, F_{c_j}) = \sigma(\dot{F}_{c_i}^h, F_{c_j}^h) + \sigma(\dot{F}_{c_i}^v, F_{c_j}^h) + \sigma(\dot{F}_{c_i}^h, F_{c_j}^v) = \sigma(\dot{F}_{c_i}^v, F_{c_j}^h),$$

where we used the fact that, by construction, $\dot{F}_{c_i}$ is vertical hence $\dot{F}_{c_i}^h = 0$. We use Leibniz's rule, Lemma 6.12, Eq. (26) for the parallel transported frame and the identity $[X_{2d}, X_i] = \nabla_{2d}X_i - \nabla_i X_{2d}$, valid for $i = 2,\dots,2d$. We obtain, after some computations

$$\begin{aligned} \sigma(\dot{F}_{c_i}^v, F_{c_j}^h) &= R(X_{2d}, X_i, X_j, X_{2d}) + \frac{1}{2} g(\tau(X_{2d}), X_{2d}) g(X_j, JX_i) \\ &\quad + \frac{1}{2} h_0 g(X_j, Q(X_i, X_{2d})) - h_0 g(X_j, Q(X_{2d}, X_i)) + \frac{1}{4} h_0^2 g(X_i, X_j) - \frac{1}{4} a_i a_j. \end{aligned}$$

for $i,j = 2,\dots,2d$. This expression is symmetric w.r.t. to i and j as a consequence of Lemma 6.9. Restoring the tangent vector $\mathsf{T} = \dot{\gamma}_\lambda$ and the functions a_j we get, for $i,j = 2,\dots,2d$

$$\begin{aligned} \mathfrak{R}^{cc}_{ij} &= \mathcal{S}[R(\mathsf{T}, X_i, X_j, \mathsf{T})] + h_0 \mathcal{S}[g(\mathsf{T}, Q(X_j, X_i))] + \frac{1}{4} h_0^2 g(X_i, X_j) \\ &\quad - \frac{1}{4} g(X_i, h_0\mathsf{T} - 2Q(\mathsf{T},\mathsf{T})) g(X_j, h_0\mathsf{T} - 2Q(\mathsf{T},\mathsf{T})). \end{aligned}$$

The trace over the subspace $S^c = \operatorname{span}\{X_2, \dots, X_{2d}\}$ is

$$\begin{aligned}\mathfrak{Ric}^c = \sum_{i=2}^{2d} \mathfrak{R}^{cc}_{ij} &= \operatorname{Ric}(\mathsf{T}) - R(\mathsf{T}, J\mathsf{T}, J\mathsf{T}, \mathsf{T}) + h_0 g(\mathsf{T}, \operatorname{tr} Q) + \frac{1}{4} h_0^2 (2d-2) - \|Q(\mathsf{T}, \mathsf{T})\|^2 \\ &= \operatorname{Ric}(\mathsf{T}) - R(\mathsf{T}, J\mathsf{T}, J\mathsf{T}, \mathsf{T}) + \frac{1}{4} h_0^2 (2d-2) - \|Q(\mathsf{T}, \mathsf{T})\|^2.\end{aligned}$$

where we used property c) and i) of Tanno's tensor.

Now we compute $\mathfrak{R}^{cb} : S^c \to S^b$ (this is a $(2d-1) \times 1$ matrix). As usual we split

$$\mathfrak{R}^{cb}_j = \sigma(\dot{F}_{c_j}, F_b) = \sigma(\dot{F}^v_{c_j}, F^h_b).$$

We use Leibniz's rule, Lemma 6.12, Eq. (26) for the parallel transported frame and the identity $[X_{2d}, X_i] = \nabla_{2d} X_i - \nabla_i X_{2d}$, valid for $i = 2, \dots, 2d$. Moreover, we use also properties b), d) and h) of Tanno's tensor. After some computations, and restoring the notation $\mathsf{T} = \dot{\gamma}$ we get

$$\begin{aligned}\mathfrak{R}^{cb}_j = &-R(\mathsf{T}, X_j, J\mathsf{T}, T) + 2g(\tau(\mathsf{T}), X_j) - 2g(\tau(\mathsf{T}), \mathsf{T}) g(\mathsf{T}, X_j) + 3g((\nabla_{\mathsf{T}} Q)(\mathsf{T}, \mathsf{T}), X_j) \\ &+ 8h_0 g(Q(\mathsf{T}, \mathsf{T}), JX_j).\end{aligned}$$

By construction, $\mathfrak{R}^{bc} = (\mathfrak{R}^{cb})^*$. We can write, in a more symmetric fashion using Eq. (24):

$$\begin{aligned}\mathfrak{R}^{bc}_j = &-R(\mathsf{T}, J\mathsf{T}, X_j, \mathsf{T}) + g(\tau(\mathsf{T}), X_j) - g(\tau(\mathsf{T}), \mathsf{T}) g(\mathsf{T}, X_j) + 3g((\nabla_{\mathsf{T}} Q)(\mathsf{T}, \mathsf{T}), X_j) \\ &+ 8h_0 g(Q(\mathsf{T}, \mathsf{T}), JX_j).\end{aligned}$$

Step 7: F_a. The structural equations give:

$$F_a = -\dot{F}_b + \mathfrak{R}^{bb} E_b + \sum_{j=2}^{2d} \mathfrak{R}^{bc}_j E_{c_j}.$$

After some computations, we obtain

$$F_a = X_0 - a_i X_i + \phi_i \partial_i + \phi_0 \partial_0,$$

where

$$\begin{aligned}a_i X_i &= h_0 X_{2d} - 2Q(X_{2d}, X_{2d}), \\ \phi_i X_i &= -\nabla_{X_0 + 2Q(X_{2d}, X_{2d})} X_{2d} + \tau(X_{2d}) - g(\tau(X_{2d}), X_{2d}) X_{2d} \\ &\quad - 4h_0 JQ(X_{2d}, X_{2d}) + 2(\nabla_{2d} Q)(X_{2d}, X_{2d}) - 3\|Q(X_{2d}, X_{2d})\|^2 JX_{2d}, \\ \phi_0 &= 2g((\nabla_{2d}\tau)(X_{2d}), JX_{2d}) - 4h_0 g(\tau(X_{2d}), X_{2d}) + g((\nabla_1 \tau)(X_{2d}), X_{2d}) \\ &\quad + 2R(X_{2d}, JX_{2d}, Q(X_{2d}, X_{2d}), X_{2d}) + 2g(\tau(X_{2d}), Q(X_{2d}, X_{2d})) \\ &\quad - 3\nabla_{2d} \|Q(X_{2d}, X_{2d})\|^2.\end{aligned}$$

The vector $\phi := \phi_i X_i$ is orthogonal to X_{2d} (or, equivalently, $\phi_{2d} = 0$). Clearly ϕ_0 is a well defined, smooth function along the geodesic. In particular, it makes sense to take the derivative of ϕ_0 in the direction of the geodesic or, in terms of our adapted frame, $\nabla_{2d}\phi_0$.

To complete the computation, we assume $Q = 0$.

Step 8: $\mathfrak{R}^{aa}$ (with $Q = 0$). Many simplification occur. In particular

$$\begin{aligned}a_i X_i &= h_0 X_{2d}, \\ \phi_i X_i &= -\nabla_0 X_{2d} + \tau(X_{2d}) - g(\tau(X_{2d}), X_{2d}) X_{2d}, \\ \phi_0 &= 2g((\nabla_{2d}\tau)(X_{2d}), JX_{2d}) - 4h_0 g(\tau(X_{2d}), X_{2d}) + g((\nabla_1 \tau)(X_{2d}), X_{2d}).\end{aligned}$$

As usual, we split

$$\mathfrak{R}^{aa} = \sigma(\dot{F}_a, F_a) = \sigma(\dot{F}^v_a, F^h_a) = \langle \dot{F}^v_a, F^h_a \rangle = \langle \dot{F}^v_a, X_0 - h_0 X_{2d} \rangle = \dot{F}^v_a(h_0) - h_0 \dot{F}^v_a(h_{2d}),$$

where we used the fact that $\dot{F}_a$ is vertical by construction. By Lemma 6.12, we obtain

$$\mathfrak{R}^{aa} = g((\nabla_{h_0 \mathsf{T} - X_0} \tau)(\mathsf{T}), \mathsf{T}) + 2h_0^2 g(\tau(\mathsf{T}), J\mathsf{T}) - 2\|\tau(\mathsf{T})\|^2 + 2g(\tau(\mathsf{T}), \mathsf{T})^2 + \nabla_{\mathsf{T}} \phi_0.$$

It is not convenient to explicitly compute the derivative $\nabla_{2d}\phi_0$ since no simplifications occur.

Remark 6.14. Since $\pi_* F_a = X_0 - h_0 \mathsf{T}$, one would expect a term of the form $R(\mathsf{T}, X_0 - h_0\mathsf{T}, X_0 - h_0\mathsf{T}, \mathsf{T})$ in the expression for $\mathfrak{R}(X_a, X_a)$. However, using the symmetries of the Riemann tensor (with torsion) and the fact that $\nabla X_0 = 0$, we obtain that $R(\mathsf{T}, X_0 - h_0\mathsf{T}, X_0 - h_0\mathsf{T}, \mathsf{T}) = 0$.

Step 9: $\mathfrak{R}^{ac}$ (with $Q = 0$). The term $\mathfrak{R}^{ac} : S^a \to S^c$ is a $1 \times (2d-1)$ matrix. For $j = 2, \dots, 2d$

$$\mathfrak{R}^{ac}_j = \sigma(\dot{F}_a, F_{c_j}) = \sigma(\dot{F}^v_a, F^h_{c_j}) = \langle \dot{F}^v_a, X_j \rangle.$$

With long computations using Leibniz's rule, the identity $[X_{2d}, X_0] = -\nabla_0 X_{2d} + \tau(X_{2d})$ and Lemma 6.12, we obtain for $j = 2, \dots, 2d-1$

$$\mathfrak{R}^{ac}_j = R(\mathsf{T}, X_0, X_j, \mathsf{T}) + g((\nabla_{\mathsf{T}}\tau)(\mathsf{T}), X_j) + 2h_0 g(\tau(J\mathsf{T}), X_j),$$

and $\mathfrak{R}^{ac}_{2d} = 0$.

Appendix A. On the Yang-Mills condition

We give here an equivalent characterization of the Yang-Mills condition for contact manifolds.

Proposition A.1. *The Yang-Mills condition* (5) *is equivalent to*

$$\sum_{i=1}^{2d} (\nabla_{X_i}\tau)(X_i) = 0.$$

for every orthonormal frame $X_1, \dots, X_{2d}$ of $\mathscr{D}$.

Proof. Let $X \in \Gamma(\mathscr{D})$. For $Y \in \Gamma(\mathscr{D})$, using the properties of Tanno connection

$$\begin{aligned}
(\nabla_X T)(X, Y) &= X(T(X, Y)) - T(\nabla_X X, Y) - T(X, \nabla_X Y) \\
&= X(d\omega(X, Y)X_0) - d\omega(\nabla_X X, Y)X_0 - d\omega(X, \nabla_X Y)X_0 \\
&= [X(g(X, JY)) - g(\nabla_X X, JY) - g(X, J\nabla_X Y)]X_0 \\
&= [-X(g(JX, Y)) + g(J\nabla_X X, Y) + g(JX, \nabla_X Y)]X_0 \\
&= -g((\nabla_X J)X, Y)X_0 = -g(Q(X, X), Y)X_0.
\end{aligned}$$

On the other hand, for $Y = X_0$ we get

$$\begin{aligned}
(\nabla_X T)(X, X_0) &= X(T(X, X_0)) - T(\nabla_X X, X_0) - T(X, \nabla_X X_0) \\
&= -X(\tau(X)) + \tau(\nabla_X X) = -(\nabla_X \tau)(X).
\end{aligned}$$

Then, for any $X \in \Gamma(\mathscr{D})$ and $Y \in \Gamma(TM)$, we have the following identity

$$(\nabla_X T)(X, Y) = -g(Q(X, X), Y)X_0 - \omega(Y)(\nabla_X \tau)(X).$$

Taking the trace, and using that $\operatorname{tr} Q = 0$ by item i) of Lemma 6.8, we get the result. □

References

[1] A. A. Agrachev. Exponential mappings for contact sub-Riemannian structures. *J. Dynam. Control Systems*, 2(3):321–358, 1996.

[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://webusers.imj-prg.fr/ davide.barilari/notes.php. 2012.

[4] A. A. Agrachev, D. Barilari, and L. Rizzi. The curvature: a variational approach. *Memoirs of the AMS (to appear)*, June 2013.

[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 and P. W. Y. Lee. Generalized Ricci curvature bounds for three dimensional contact sub-riemannian manifolds. *Math. Ann.*, 360(1-2):209–253, 2014.

[7] A. A. Agrachev, L. Rizzi, and P. Silveira. On conjugate times of LQ optimal control problems. *Journal of Dynamical and Control Systems*, pages 1–17, 2014.

[8] A. A. Agrachev and I. Zelenko. Geometry of Jacobi curves. I. *J. Dynam. Control Systems*, 8(1):93–140, 2002.

[9] D. Barilari and L. Rizzi. Comparison theorems for conjugate points in sub-Riemannian geometry. *ESAIM: Control, Optimisation and Calculus of Variations (to appear)*, 2015.

[10] D. Barilari and L. Rizzi. On Jacobi fields and canonical connection in sub-Riemannian geometry. *ArXiv e-prints*, June 2015.
[11] F. Baudoin and N. Garofalo. Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries. *Journal of the EMS (to appear)*, Jan. 2011.
[12] F. Baudoin, B. Kim, and J. Wang. Transverse Weitzenböck formulas and curvature dimension inequalities on Riemannian foliations with totally geodesic leaves. *Comm. Anal. Geom. (to appear)*, Aug. 2014.
[13] F. Baudoin and J. Wang. Curvature dimension inequalities and subelliptic heat kernel gradient bounds on contact manifolds. *Potential Anal.*, 40(2):163–193, 2014.
[14] D. E. Blair. *Riemannian geometry of contact and symplectic manifolds*, volume 203 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, second edition, 2010.
[15] W.-L. Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–105, 1939.
[16] M. Falcitelli, S. Ianus, and A. M. Pastore. *Riemannian submersions and related topics*. World Scientific Publishing Co., Inc., River Edge, NJ, 2004.
[17] W. K. Hughen. *The sub-Riemannian geometry of three-manifolds*. ProQuest LLC, Ann Arbor, MI, 1995. Thesis (Ph.D.)–Duke University.
[18] P. W. Y. Lee and C. Li. Bishop and Laplacian comparison theorems on Sasakian manifolds. *ArXiv e-prints*, Oct. 2013.
[19] P. W. Y. Lee, C. Li, and I. Zelenko. Ricci curvature type lower bounds for sub-Riemannian structures on Sasakian manifolds. *Discrete Contin. Dyn. Syst.*, 36(1):303–321, 2016.
[20] 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.
[21] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The mathematical theory of optimal processes*. Translated from the Russian by K. N. Trirogoff; edited by L. W. Neustadt. Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.
[22] P. Rashevsky. Any two points of a totally nonholonomic space may be connected by an admissible line. *Uch. Zap. Ped Inst. im. Liebknechta*, 2:83–84, 1938.
[23] L. Rifford. *Sub-Riemannian geometry and optimal transport*. Springer Briefs in Mathematics. Springer, Cham, 2014.
[24] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.
[25] L. Rizzi. Measure contraction properties of Carnot groups. *ArXiv e-prints*, Oct. 2015.
[26] L. Rizzi and P. Silveira. Sub-Riemannian Ricci curvatures and universal diameter bounds for 3-Sasakian manifolds. *ArXiv e-prints*, Sept. 2015.
[27] M. Rumin. Formes différentielles sur les variétés de contact. *J. Differential Geom.*, 39(2):281–330, 1994.
[28] S. Tanno. Variational problems on contact Riemannian manifolds. *Trans. Amer. Math. Soc.*, 314(1):349–379, 1989.
[29] I. Zelenko and C. Li. Differential geometry of curves in Lagrange Grassmannians with given Young diagram. *Differential Geom. Appl.*, 27(6):723–742, 2009.