ESAIM: COCV 22 (2016) 786–810
DOI: 10.1051/cocv/2015026

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

MINIMUM-TIME STRONG OPTIMALITY OF A SINGULAR ARC: THE MULTI-INPUT NON INVOLUTIVE CASE *

Francesca Chittaro[1,2] and Gianna Stefani[3]

Abstract. We consider the minimum-time problem for a multi-input control-affine system, where we assume that the controlled vector fields generate a non-involutive distribution of constant dimension, and where we do not assume *a priori* bounds for the controls. We use Hamiltonian methods to prove that the coercivity of a suitable second variation associated to a Pontryagin singular arc is sufficient to prove its strong-local optimality. We provide an application of the result to a generalization of Dubins problem.

Mathematics Subject Classification. 49J15, 49J30, 49K15, 49K30.

Received April 29, 2014.
Published online May 18, 2016.

1. Introduction

In this paper we are concerned with the minimum-time problem associated with a control-affine system with several controls:

$$\min T \tag{1.1}$$

subject to

$$\begin{cases} \dot{\xi} = (f_0 + \sum_{i=1}^m u_i f_i) \circ \xi(t) \\ \xi(0) \in N_0, \quad \xi(T) \in N_f \\ \mathbf{u} = (u_1, \dots, u_m) \in U \subset \mathbb{R}^m. \end{cases} \tag{1.2}$$

The state q belongs to a smooth n-dimensional manifold M, $f_0, f_1, \dots, f_m$, are smooth vector fields on M, N_0 and N_f are smooth submanifolds of M and the control functions belong to $L^\infty([0,T], U)$. We remark that for smooth we mean C^∞.

We are interested in sufficient conditions for the strong-local optimality of *singular* Pontryagin extremals of problem (1.1)−(1.2), were *strong* means *with respect to the C^0-norm* of the trajectories $\xi(\cdot)$, and *singular* means

Keywords and phrases. Control-affine systems, singular extremals, minimum-time problem, sufficient optimality conditions, second variation, Hamiltonian methods.

* *This work was supported by Digiteo grant Congeo; by the ANR project GCM, program "Blanche", project number NT09_504490; by the European Research Council, ERC StG 2009 "GeCoMethods", contract number 239748; by PRIN 200894484E_002; and by the research fund CARTT, IUT Toulon–La Garde.*

[1] Aix Marseille Université, CNRS, ENSAM, LSIS UMR 7296, 13397 Marseille, France

[2] Université de Toulon, CNRS, LSIS UMR 7296, 83957 La Garde, France. francesca-carlotta.chittaro@univ-tln.fr

[3] DIMAI, via S. Marta 3 − 50137 Firenze, Italy. gianna.stefani@unifi.it

Article published by EDP Sciences

© EDP Sciences, SMAI 2016

that $\mathbf{u} \in \operatorname{int} U$. More precisely, we assume that there exists a candidate Pontryagin extremal $\widehat{\lambda} : [0, \widehat{T}] \to T^*M$ with associated control function $\widehat{\mathbf{u}}(\cdot) \in L^\infty([0, \widehat{T}], \operatorname{int} U)$ that satisfies $\pi\widehat{\lambda}(0) \in N_0$ and $\pi\widehat{\lambda}(\widehat{T}) \in N_f$, and we look for sufficient conditions that guarantee the strong-local optimality of the trajectory $\widehat{\xi} = \pi\widehat{\lambda}$, according to the following definition:

Definition 1.1. The trajectory $\widehat{\xi}$ is a strong-local minimizer of the above considered problem, if there exist a neighborhood V of its graph in $\mathbb{R} \times M$ and $\epsilon > 0$ such that $\widehat{\xi}$ is a minimizer among the admissible trajectories whose graph is contained in V and whose final time is greater than $\widehat{T} - \epsilon$, independently on the values of the associated controls.

This notion has been called *time-state-local optimality* in [25], where also a stronger version of optimality is considered.

The only assumption we do on the control set U is that it has non-empty interior; although by Filippov's Theorem [1] we know that the existence of the minimum is guaranteed when U is compact and convex, here the existence of a candidate minimizer is taken as assumption.

A classical approach to sufficient optimality conditions is to consider the second variation (see for instance [1–3, 14, 16, 28] and references therein). In particular, in [1, 2] the authors propose the definitions of an *intrinsically defined* second variation, which is invariant for coordinate changes, and therefore suitable to study optimal control problems defined on smooth manifolds.

A peculiarity of control-affine minimum-time problems is that the second variation does not contain the Legendre term (that is, the term which is quadratic in the variations of the control), thus turning out to be singular. A tool widely used to overcome this problem is the so-called Goh transformation [17]. Thanks to this transformation, performed in a coordinate-free way, we are able to convert the second variation proposed in [2] into another functional which is no more singular, and thus it can be asked to be coercive with respect to the L^2-norm of the new control variable. This approach, both for classical and intrinsically-defined second variations, has been widely used in the analysis of sufficient optimality conditions (see for instance [3, 15, 25, 27], and references therein).

In optimal-control problems, a classical method to prove the optimality of a Pontryagin extremal is to cover a neighborhood of the reference trajectory with other admissible trajectories, to lift them to the cotangent bundle, and compare the costs evaluated along each trajectory. In the standard theory, the trajectories to be lifted are obtained by projecting suitable solutions of the Hamiltonian system associated with the maximized Hamiltonian $F_{\max}$, see for example [1, 2]. This *Hamiltonian method* is particularly effective, since it allows us to compare trajectories that belong to a C^0-neighborhood of the reference trajectory, independently on the value of the control.

When the extremal is singular $F_{\max}$ cannot be used (see [25]), then to construct the lifted trajectories we consider the solutions of a system governed by a Hamiltonian greater than or equal to $F_{\max}$, as suggested by the approach used in [25, 27].

Ultimately, the paradigm to get sufficient optimality condition for singular extremals combines an approach based on the coercivity of the second variation with the Hamiltonian approach. It relies on the following facts.

- Under some regularity conditions, it is possible to define a smooth super-Hamiltonian whose flow is tangent to all singular extremals.
- The derivative of the super-Hamiltonian flow is, up to an isomorphism, the Hamiltonian flow associated with the linear-quadratic problem given by the second variation.
- If the second variation is coercive, it is possible to transform the linear-quadratic problem associated with the original one into a problem with free initial point, whose second variation is still coercive. In particular, this implies that the space of initial constraints for the linear-quadratic problem remains horizontal (that is, it projects bijectively on M) under the action of the associated Hamiltonian flow.
- The previous points imply that the projection on M of the super-Hamiltonian flow emanating from the Lagrangian manifold associated with the initial conditions of the new problem is locally invertible. As a

result we get that it is possible to lift the trajectories to the cotangent bundle, in order to apply the Hamiltonian method.

In the single-input case, problem (1.1)−(1.2) has already been studied in [25], where it has been shown that the coercivity of the second variation is a sufficient condition for the strong-local optimality of singular Pontryagin extremals. In [11] the authors studied the multi-input problem under the assumptions that the controlled vector fields generate an involutive distribution. In this paper we relax this condition, that is we allow the controlled vector field to generate a non-involutive distribution.

Our assumptions on the reference extremal are stronger than the usual ones; in particular, we ask that the High Order Goh condition (HOGC, Assumption 3) holds true. These requirements are justified by the fact that our result remains true even if $U = \mathbb{R}^m$; indeed, we prove that in this case the HOGC is a necessary optimality condition, see Appendix A. This phenomenon is not pointed out when the Lie algebra generated by the controlled vector fields is involutive, in particular when the system is single-input. Indeed, in these cases the High Order Goh condition is automatically satisfied under Goh condition.

We believe that this result, applied to the case where U is an unbounded set, could be of help in the study of the infimum-time problem where the "optimal" trajectories may contain jumps, as in [9, 10, 18], where integral costs are considered.

The structure of the paper is the following: we state the regularity assumptions in Section 2; in Section 3 we define the second variation and investigate the implications of its coercivity; the Hamiltonian method is exposed in Sections 4 and 5, where we state and prove the main result; in Section 6 we provide an example, based on a high dimensional version of Dubins problem. In the Appendices there are technical details on some results stated in the paper.

2. Notations and regularity assumptions

In this section we clarify the notation we will use throughout the paper, and we state the regularity assumptions on the system.

Let f be a vector field on the manifold M and $\varphi : M \to \mathbb{R}$ be a smooth function. The action of f on φ (directional derivative or Lie derivative) evaluated at a point q is denoted with the two expressions

$$L_f\varphi(q) = \langle d\varphi(q), f(q)\rangle.$$

The Lie brackets of two vector fields f, g are denoted as commonly with $[f, g]$. When dealing with vector fields labeled by indexes, we will use the following notations to denote their Lie brackets:

$$f_{ij}(q) = [f_i, f_j](q), \qquad f_{ijk}(q) = [f_i, [f_j, f_k]](q).$$

We call $\mathfrak{f}$ the set of the controlled vector fields of the control system (1.2), that is $\mathfrak{f} = \{f_1, \dots, f_m\}$, and $\mathrm{Lie}(\mathfrak{f})$ the Lie algebra generated by the set $\mathfrak{f}$. We denote $\mathrm{Lie}_q(\mathfrak{f}) = \mathrm{span}\{f(q) : f \in \mathrm{Lie}(\mathfrak{f})\}$. In the following, for every $q \in M$, we call $\mathcal{I}_q$ the integral manifold of the distribution $\mathrm{Lie}(\mathfrak{f})$ passing through q. The first assumption of this paper concerns the regularity of $\mathrm{Lie}(\mathfrak{f})$.

Assumption 1. The controlled vector field $f_1, \dots, f_m$ are linearly independent and the Lie algebra $\mathrm{Lie}(\mathfrak{f})$ has constant dimension R, that is, $\dim \mathrm{Lie}_q(\mathfrak{f}) = R$ for every $q \in M$.

Let us consider the cotangent bundle T^*M of M, and let π denote the canonical projection on M. It is well known that T^*M possesses a canonically defined symplectic structure, given by the symplectic form $\sigma_\ell = d\varsigma(\ell)$, where ℓ denotes an element of T^*M and ς is the Liouville canonic 1-form $\varsigma(\ell) = \ell \circ \pi_*$.

We denote with the corresponding capital letter the Hamiltonian function associated with every vector field on M, that is $F(\ell) = \langle \ell, f(\pi\ell)\rangle$.

Remark 2.1. Let us recall that the following relation between the Lie brackets of two vector fields f, g and the Poisson brackets of their associated Hamiltonian functions holds:

$$\langle \ell, [f,g](\pi\ell)\rangle = \{F,G\}(\ell).$$

As above, we denote

$$F_{ij}(\ell) = \{F_i, F_j\}(\ell) \qquad F_{ijk}(\ell) = \{F_i, \{F_j, F_k\}\}(\ell).$$

We recall that the symplectic structure allows us to associate, with each Hamiltonian function F, the Hamiltonian vector field $\boldsymbol{F}$ on T^*M defined by the action

$$\langle dF(\ell), \cdot\rangle = \sigma_\ell(\cdot, \boldsymbol{F}(\ell)).$$

In the following, we consider some special Hamiltonian functions associated with the optimal control problem (1.1)−(1.2). The (time-dependent) *reference Hamiltonian* is defined by

$$\widehat{F}_t(\ell) = F_0(\ell) + \sum_{i=1}^{m} \widehat{u}_i(t) F_i(\ell), \tag{2.1}$$

where $\widehat{\mathbf{u}}(\cdot)$ is the *reference control*. The maximized Hamiltonian is defined by

$$F_{\max}(\ell) = \sup_{\mathbf{u}\in U} \Big(F_0(\ell) + \sum_{i=1}^{m} u_i F_i(\ell) \Big).$$

We point out that, when U is unbounded, $F_{\max}(\ell)$ may be $+\infty$. Nevertheless, the Pontryagin Maximum Principle assure that it is finite along the reference extremal, see the arguments below.

The Hamiltonian flow from time 0 to time t associated with the reference Hamiltonian, that is the solution of the equation $\dot{\ell} = \widehat{\boldsymbol{F}}_t(\ell)$, is denoted with $\widehat{\mathcal{F}}_t$.

Let us consider an *admissible triple* $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ for the problem (1.2), that is a solution of the control system; let us assume that $\widehat{\mathbf{u}} \in \operatorname{int} U$, and let us set $\widehat{q}_0 = \widehat{\xi}(0)$ and $\widehat{q}_f = \widehat{\xi}(\widehat{T})$. We study the strong-local optimality of the triple $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$, that in the following we call *reference triple*, among all solutions of (1.2) with $N_0 \subset \mathcal{I}_{\widehat{q}_0}$ and $N_f \subset \mathcal{I}_{\widehat{q}_f}$. In particular, Assumption 1 can be asked to hold only in a neighborhood of the reference trajectory.

A classical necessary condition for the optimality of the reference triple $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ is the Pontryagin Maximum Principle (PMP), that we recall here stated in its Hamiltonian form (see [1]). PMP states that if a reference trajectory $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ satisfying $\widehat{\mathbf{u}} \in \operatorname{int} U$ is time-optimal, then there exist a Lipschitzian curve $\widehat{\lambda} : [0, \widehat{T}] \to T^*M$ and $p_0 \in \{0, 1\}$ that satisfy the following equations:

$$\widehat{\lambda}(t) \neq 0 \qquad \forall\, t \in [0, \widehat{T}] \tag{2.2}$$

$$\pi\widehat{\lambda}(t) = \widehat{\xi}(t) \qquad \forall\, t \in [0, \widehat{T}] \tag{2.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\lambda}(t) = \widehat{\boldsymbol{F}}_{\boldsymbol{t}}(\widehat{\lambda}(t)) \qquad \forall\, t \in [0, \widehat{T}] \tag{2.4}$$

$$F_i(\widehat{\lambda}(t)) = 0 \qquad \forall\, i = 1, \dots, m \quad \forall\, t \in [0, \widehat{T}] \tag{2.5}$$

$$\widehat{F}_t(\widehat{\lambda}(t)) = F_0(\widehat{\lambda}(t)) = p_0 \qquad \forall\, t \in [0, \widehat{T}] \tag{2.6}$$

$$\widehat{\lambda}(0)|_{T_{\widehat{q}_0} N_0} = 0 \qquad \widehat{\lambda}(\widehat{T})|_{T_{\widehat{q}_f} N_f} = 0. \tag{2.7}$$

The Lipschitzian curves that satisfy equations (2.2)–(2.7) are called *extremals*. If $p_0 = 1$ we say the the extremal $\widehat{\lambda}$ is *normal*, while in the other case we say that it is *abnormal*.

Assumption 2. We assume that the reference triple satisfies the PMP, and that the extremal associated with the reference triple is normal. We call such an extremal *reference extremal* and we denote it with $\widehat{\lambda}$.

By differentiating equation (2.6) with respect to time, we obtain the following condition:

$$F_{0i}(\widehat{\lambda}(t)) + \sum_{j=1}^{m} \widehat{u}_j(t) F_{ji}(\widehat{\lambda}(t)) = 0 \qquad i = 1, \dots, m, \quad \text{a.e. } t \in [0, \widehat{T}]. \tag{2.8}$$

In literature additional necessary conditions for the optimality of a singular extremal are known (see [1]). Namely, if the reference triple is optimal, then there exists an extremal λ associated with the reference triple that satisfies the following conditions:

(Goh condition)

$$F_{ij}(\lambda(t)) = 0 \qquad \forall\, i, j = 1, \dots, m,\ t \in [0, \widehat{T}].$$

(Generalized Legendre Condition) the quadratic form

$$\mathbb{L}_{\lambda(t)} : \mathbf{v} \mapsto \sum_{i,j=1}^{m} \mathrm{v}_i \mathrm{v}_j F_{ij0}(\lambda(t)) + \sum_{i,j,k=1}^{m} \mathrm{v}_i \mathrm{v}_j \widehat{u}_k(t) F_{ijk}(\lambda(t)) \tag{2.9}$$

is non-positive for any $\mathbf{v} = (\mathrm{v}_1, \dots, \mathrm{v}_m) \in \mathbb{R}^m$ and for a.e. $t \in [0, \widehat{T}]$.

Remark 2.2. Notice that the matrix $\mathbb{L}_{\lambda(t)}$ is symmetric by (2.8) and Jacobi identity.

We strengthen the two necessary conditions above defined.

Assumption 3 (High Order Goh Condition)**.** We assume that the reference extremal $\widehat{\lambda}$ satisfies the following equations

$$\langle \widehat{\lambda}(t), f(\widehat{\xi}(t)) \rangle = 0 \qquad \forall\, f \in \mathrm{Lie}(\mathfrak{f}),\ t \in [0, \widehat{T}].$$

The HOGC is a stronger condition than the usual one, but in our case the optimality of the singular extremal is proved also when $U = \mathbb{R}^m$; in Appendix A we show that, for $U = \mathbb{R}^m$, the HOGC is a necessary optimality condition. As a matter of fact, if the Lie algebra generated by the controlled vector fields is 2-step bracket generating, then the HOGC coincides with Goh condition.

Remark moreover that, under Assumption 3, the quadratic form $\mathbb{L}_{\widehat{\lambda}(t)}$ is given by

$$\mathbb{L}_{\widehat{\lambda}(t)} : \mathbf{v} \mapsto \sum_{i,j=1}^{m} v_i v_j F_{ij0}(\widehat{\lambda}(t)), \tag{2.10}$$

so that it is continuous as a function of time.

Assumption 4 (Strengthened Generalized Legendre Condition)**.** There exists a constant $c > 0$ such that

$$\mathbb{L}_{\widehat{\lambda}(t)}[\mathbf{v}]^2 \leq -c|\mathbf{v}|^2 \tag{2.11}$$

for any $\mathbf{v} = (\mathrm{v}_1, \dots, \mathrm{v}_m) \in \mathbb{R}^m$ and for every $t \in [0, \widehat{T}]$.

As a consequence of Assumptions 1–3 and equation (2.8), we get that

$$\langle \widehat{\lambda}(t), [f_0, f](\widehat{\xi}(t)) \rangle = 0 \qquad \forall\, f \in \mathrm{Lie}(\mathfrak{f}), \quad \forall\, t \in [0, \widehat{T}] \tag{2.12}$$

$$\sum_{j=1}^{m} (\mathbb{L}_{\widehat{\lambda}(t)})_{ij} \widehat{u}_j(t) = F_{00i}(\widehat{\lambda}(t)) \qquad i = 1, \dots, m \quad \text{a.e. } t \in [0, \widehat{T}]. \tag{2.13}$$

From (2.13) and Assumption 4 we can recover the reference control as feedback on the cotangent bundle and, by induction, we can prove that it is smooth.

From now on we restrict to a (full-measure) neighborhood $\mathfrak{U}$ of $\widehat{\lambda}([0,\widehat{T}])$ in T^*M where the SGLC is satisfied, that is, where the quadratic form $\mathbb{L}_\ell$ is negative-definite. We define two subsets of $\mathfrak{U}$ which are crucial for our construction:

$$\Sigma = \{\ell \in \mathfrak{U} : \langle \ell, f(\pi\ell)\rangle = 0 \ \forall\, f \in \mathrm{Lie}(\mathfrak{f})\} \tag{2.14}$$

$$\mathcal{S} = \{\ell \in \Sigma : \langle \ell, [f_0, f](\pi\ell)\rangle = 0 \ \forall\, f \in \mathrm{Lie}(\mathfrak{f})\}. \tag{2.15}$$

By Assumption 1, possibly restricting $\mathfrak{U}$, Σ is an embedded manifold of codimension R. Moreover every singular extremal that satisfies the HOGC is contained in $\mathcal{S}$. We set the following regularity assumption on $\mathcal{S}$, which requires that it is a submanifold of maximal dimension (see the property (P2) below).

Assumption 5 (Regularity of $\mathcal{S}$)**.** $\mathcal{S}$ is a submanifold of Σ of codimension m.

Thanks to regularity assumptions, the manifolds Σ and $\mathcal{S}$ have the following properties. The proofs can be obtained adapting those in [11].

(P1) It is easy to see that the vector space $\mathrm{Lie}_\ell(\boldsymbol{F}_1, \dots, \boldsymbol{F}_m)$ has dimension R for every $\ell \in \Sigma$. Moreover, every vector field $X \in \mathrm{Lie}(\boldsymbol{F}_1, \dots, \boldsymbol{F}_m)$ is tangent to Σ.

(P2) It is not difficult to prove that the SGLC implies that, for $\ell \in \Sigma$, $\boldsymbol{F}_{01}(\ell), \dots, \boldsymbol{F}_{0m}(\ell)$ are linearly independent and that

$$\mathrm{span}\,(\boldsymbol{F}_{01}(\ell), \dots, \boldsymbol{F}_{0m}(\ell)) \cap T_\ell\Sigma = \{0\}.$$

Since the linear independence of $\boldsymbol{F}_{01}(\ell), \dots, \boldsymbol{F}_{0m}(\ell)$ implies that also $dF_{01}(\ell), \dots, dF_{0m}(\ell)$ are linearly independent, then $\mathcal{S}$ has codimension at least m in Σ. As a consequence, Assumption 5 states then that $\mathcal{S}$ has the maximal dimension. The arguments above prove also that $\mathcal{S}$ can be characterized by

$$\mathcal{S} = \{\ell \in \Sigma : F_{0i}(\ell) = 0 \ \forall\, i = 1, \dots, m\}.$$

Notice that the existence of a normal singular extremal satisfying the HOGC implies that $f_0(q)$ is never contained in $\mathrm{Lie}_q(\mathfrak{f})$, for q belonging to a neighborhood of the reference trajectory on M. Moreover, Assumption 5 is equivalent to the following one:

$$[f_0, f] \in \mathrm{Lie}(\mathfrak{f}) + \mathrm{span}(\{f_{01}, \dots, f_{0m}\}) \qquad \forall\, f \in \mathrm{Lie}(\mathfrak{f}).$$

(P3) Similar arguments show that $\mathrm{span}\,(\boldsymbol{F}_1(\ell), \dots, \boldsymbol{F}_m(\ell)) \cap T_\ell\mathcal{S} = \{0\}$, $\ell \in \mathcal{S}$.

(P4) $\boldsymbol{F}_0$ is tangent to Σ in $\mathcal{S}$.

(P5) Our assumptions guarantee the existence of a Hamiltonian vector field tangent to all singular extremals. Indeed, setting for every $\ell \in \mathfrak{U}$

$$\boldsymbol{\nu}(\ell) = \mathbb{L}_\ell^{-1} \begin{pmatrix} F_{001}(\ell) \\ \vdots \\ F_{00m}(\ell) \end{pmatrix}, \tag{2.16}$$

we get that the vector field $\boldsymbol{F}_{\mathcal{S}} = \boldsymbol{F}_0 + \sum_{i=1}^m \nu_i \boldsymbol{F}_i$ is tangent to $\mathcal{S}$, and the reference extremal $\widehat{\lambda}(\cdot)$ is an integral curve of $\boldsymbol{F}_{\mathcal{S}}$. Indeed every singular extremal associated with our dynamics is an integral curve of $\boldsymbol{F}_{\mathcal{S}}$.

3. Second variation

In this section we define the second variation for the problem under study, and we investigate the consequences of the coercivity of the second variation. The computations can be recovered by adapting those present in [25,27].

3.1. Construction of the second variation

We consider the second variation associated with the sub-problem of (1.1)−(1.2) with fixed final point, that is we add the constraint $\xi(T) = \widehat{q}_f$. To compute this second variation, we transform such minimum-time problem into a Mayer problem on the fixed time interval $[0, \widehat{T}]$ and on the state space $\mathbb{R} \times M$. Namely, putting u_0 as a new constant control with positive values, we reparametrize the time as $u_0 t$, and we set $\boldsymbol{q} = (q^0, q) \in \mathbb{R} \times M$, $\boldsymbol{f}_0(\boldsymbol{q}) = f_0(q) + \frac{\partial}{\partial q^0}$ and $\boldsymbol{f}_i(\boldsymbol{q}) = f_i(q)$, $i = 1, \dots, m$. Then the minimum-time problem between N_0 and $\widehat{q}_f$ is equivalent to the Mayer problem on $\mathbb{R} \times M$ described below.

$$\min \xi^0(\widehat{T}) \tag{3.1}$$

subject to

$$\begin{cases} \dot{\boldsymbol{\xi}}(t) = u_0 \boldsymbol{f}_0(\boldsymbol{\xi}(t)) + \sum_{i=1}^m u_0 u_i(t) \boldsymbol{f}_i(\boldsymbol{\xi}(t)) \quad t \in [0, \widehat{T}] \\ \boldsymbol{\xi}(0) \in \{0\} \times N_0, \quad \boldsymbol{\xi}(\widehat{T}) \in \mathbb{R} \times \{\widehat{q}_f\} \\ (u_0, \mathbf{u}) \in (0, +\infty) \times L^\infty([0, \widehat{T}], U) \end{cases} \tag{3.2}$$

where $\boldsymbol{\xi} = (\xi^0, \xi)$. It is not difficult to see that the trajectory $\widehat{\boldsymbol{\xi}}(t) = (t, \widehat{\xi}(t))$, associated with the controls $u_0 = 1$ and $\mathbf{u} = \widehat{\mathbf{u}}$, is an extremal with associated adjoint covector $\widehat{\boldsymbol{\lambda}} : s \mapsto ((-1, t), \widehat{\lambda}(t)) \in \mathbb{R}^* \times T^*M$.

For $t \in [0, \widehat{T}]$, we define the evolution map $\widehat{S}_t : M \to M$ by its action $\widehat{S}_t : x_0 \mapsto \xi(t)$, where ξ is the solution of the equation $\dot{\xi} = f_0(\xi) + \sum_{i=1}^m \widehat{u}_i f_i(\xi)$ with initial condition $\xi(0) = q_0$. In particular, $\widehat{S}_t(\widehat{q}_0) = \widehat{\xi}(t)$. We locally define around $\widehat{q}_0$ the *pull-back vector fields*

$$g_t^i = \widehat{S}_{t*}^{-1} f_i \circ \widehat{S}_t, \qquad i = 0, \dots, m.$$

Analogously, for the Mayer problem we define the evolution $\widehat{\mathbf{S}}_t : \mathbb{R} \times M \to \mathbb{R} \times M$ as $\widehat{\mathbf{S}}_t : (q^0, q) \mapsto (q^0 + t, \widehat{S}_t(q))$, and the pull-back system of (3.2) corresponding to the reference control $\widehat{\mathbf{u}}$ as

$$\boldsymbol{\eta}(t) = \widehat{\mathbf{S}}_t^{-1} \circ \boldsymbol{\xi}(t).$$

The Mayer problem (3.1)−(3.2) is then equivalent to the following one:

$$\min \eta^0(\widehat{T})$$

subject to the control system

$$\begin{cases} \dot{\eta}^0(t) = u_0 - 1 \\ \dot{\eta}(t) = (u_0 - 1) g_t^0(\eta(t)) + \sum_{i=1}^m (u_0 u_i(t) - \widehat{u}_i(t)) g_t^i(\eta(t)) \\ \boldsymbol{\eta}(0) \in \{0\} \times N_0 \qquad \boldsymbol{\eta}(\widehat{T}) \in \mathbb{R} \times \{\widehat{q}_0\}. \end{cases} \tag{3.3}$$

Let us now consider variations $(\delta u_0, \delta x, \delta u) \in \mathbb{R} \times T_{\widehat{q}_0} N_0 \times L^\infty([0, \widehat{T}], \mathbb{R}^m)$ around the reference trajectory, and let us evaluate the coordinate-free second variation of the Mayer problem, following [2]. We choose any two smooth functions $\boldsymbol{\alpha}, \boldsymbol{\beta} : \mathbb{R} \times M \to \mathbb{R}$ that satisfy the following constraints:

$$\boldsymbol{\alpha}(q^0, q) = \alpha(q) - q^0, \qquad \alpha|_{N_0} \equiv 0, \qquad d\alpha(\widehat{q}_0) = \widehat{\lambda}(0), \tag{3.4}$$

$$\boldsymbol{\beta}(q^0, q) = q^0 + \beta(q), \qquad d\beta(\widehat{q}_0) = -\widehat{\lambda}(0), \tag{3.5}$$

for two suitable smooth functions $\alpha, \beta : M \to \mathbb{R}$. Thanks to the High Order Goh Conditions, we can choose the function α in such a way that it satisfies the constraint $\alpha|_{\mathcal{I}_{\widehat{q}_0}} \equiv 0$, where $\mathcal{I}_{\widehat{q}_0}$ is the integral manifold of the distribution $\mathrm{Lie}(\mathfrak{f})$ passing through $\widehat{q}_0$. Moreover, we can choose $\beta = -\alpha$, since the second variation does not depend on the particular choice of α and β with the properties (3.4) and (3.5) (see [2]).

The second variation is given by

$$J''[(\delta x, \delta u_0, \delta u)]^2 = \int_0^{\widehat{T}} \delta u_0 L_{\delta\boldsymbol{\eta}(t)} L_{\boldsymbol{g}_t^0} \widehat{\boldsymbol{\beta}}(\widehat{T}, \widehat{q}_0) + \sum_{i=1}^m (\delta u_0 \widehat{u}_i(t) + \delta u_i(t)) L_{\delta\boldsymbol{\eta}(t)} L_{\boldsymbol{g}_t^i} \widehat{\boldsymbol{\beta}}(\widehat{T}, \widehat{q}_0)\, \mathrm{d}t, \tag{3.6}$$

where $\delta\boldsymbol{\eta}(t) \in \mathbb{R} \times T_{\widehat{q}_0} M$ is the linearization of $\boldsymbol{\eta}(t)$ and satisfies the following system:

$$\begin{cases} \delta\dot{\eta}^0(t) = \delta u_0 \\ \delta\dot{\eta}(t) = \delta u_0 g_t^0(\widehat{q}_0) + \sum_{i=1}^m (\delta u_0 \widehat{u}_i(t) + \delta u_i(t)) g_t^i(\widehat{q}_0) \\ \delta\boldsymbol{\eta}(0) = (0, \delta x) \in \{0\} \times T_{\widehat{q}_0} N_0, \quad \delta\boldsymbol{\eta}(\widehat{T}) \in \mathbb{R} \times \{0\}. \end{cases} \tag{3.7}$$

Remark 3.1. If $\delta\boldsymbol{\eta}$ satisfies the system (3.7), then the value of the second variation does not depend on the particular choice of α and β, provided that they satisfy properties (3.4)−(3.5) (see [2]). Then J'' is well defined and coordinate free.

Since we are interested only in the so-called time-state local optimality, we restrict us to the subproblem with $\delta u_0 = 0$, and, proceeding as in [25], we define $\boldsymbol{w}(\cdot)$ and $\boldsymbol{\epsilon}$ by

$$w_i(t) = \int_t^{\widehat{T}} \delta u_i(s)\, \mathrm{d}s \tag{3.8}$$

$$\epsilon_i = w_i(0), \tag{3.9}$$

for $i = 1, \ldots, m$. In this way, the control variation δu is embedded as the pair $(\boldsymbol{\epsilon}, \boldsymbol{w}(\cdot))$ in the space $\mathbb{R}^m \times L^2([0, \widehat{T}], \mathbb{R}^m)$. We remark that this embedding is continuous and it has dense image. Then the second variation defined by (3.6)−(3.7) writes as

$$\begin{aligned} J''[(\delta x, \boldsymbol{\epsilon}, \boldsymbol{w}(\cdot))]^2 = {} & \frac{1}{2} \sum_{i,j=1}^m \left(L_{\epsilon_i f_i} L_{\epsilon_j f_j} \beta(\widehat{q}_0) + \int_0^{\widehat{T}} w_i(t) w_j(t) L_{[\dot{g}_t^i, g_t^j]} \beta(\widehat{q}_0)\, \mathrm{d}t \right) \\ & + \sum_{i=1}^m \left(L_{\delta x} L_{\epsilon_i f_i} \beta(\widehat{q}_0) + \int_0^{\widehat{T}} w_i(t) L_{\zeta(t)} L_{\dot{g}_t^i} \beta(\widehat{q}_0)\, \mathrm{d}t \right), \end{aligned} \tag{3.10}$$

where the function $\zeta : [0, \widehat{T}] \to T_{\widehat{q}_0} M$ is the solution of the equation

$$\dot{\zeta}(t) = \sum_{i=1}^m w_i(t) \dot{g}_t^i(\widehat{q}_0), \tag{3.11}$$

with boundary conditions

$$\zeta(0) = \delta x + \sum_{i=1}^m \epsilon_i f_i(\widehat{q}_0), \qquad \zeta(\widehat{T}) = 0. \tag{3.12}$$

Let us observe that the second variation is realized as a linear-quadratic control problem in the state-variable ζ, with control $\boldsymbol{w}$ (see [2, 27, 28]). Notice moreover that

$$\dot{g}_t^i(\widehat{q}_0) = \widehat{S}_{t*}^{-1}[f_0, f_i] \circ \widehat{\xi}(t) \quad \text{and} \quad L_{[\dot{g}_t^i, g_t^j]} \beta(\widehat{q}_0) = -F_{ij0}(\widehat{\lambda}(t)).$$

Finally, thanks to the choice of β, the finite-dimensional term in (3.10) is null.

It is clear that, if $T_{\widehat{q}_0} N_0 \cap \operatorname{span}(\{f_1(\widehat{q}_0), \ldots, f_m(\widehat{q}_0)\}) \neq 0$, then the above defined quadratic form cannot be coercive. On the other hand, the paradigm exposed in the introduction requires that the flow of the super-Hamiltonian emanating from Σ remains contained in Σ, as Theorem 5.1 in Section 5 describes (see also [25,27]);

in particular, the horizontal Lagrangian sub-manifold of the initial constraints must be contained in Σ. These arguments suggest to require coercivity of (3.10) allowing all vectors in $\mathrm{Lie}_{\widehat{q}_0}(\mathfrak{f})$ to be initial conditions for the state ζ; see Section 3.2 and the proof of Theorem 5.2 in Section 5. In particular, this means to consider the second variation for the minimum time problem from $\mathcal{I}_{\widehat{q}_0}$ to $\widehat{q}_f$.

This is not surprising, since our result holds true also for unbounded controls; indeed, if $U = \mathbb{R}^m$, for every two points $q_1, q_2 \in \mathcal{I}_{\widehat{q}_0}$, there exist a sequence of times $t_k \to 0$ and a sequence of controls $\mathbf{u}_k$ such that the sequence $t_k \mapsto \{\xi_k(q_1, \mathbf{u}_k, t_k)\}_k$ of the solutions at time t_k of the system (1.2) starting from the point q_1 and relative to the control $\mathbf{u}_k$ tends to q_2 (see [4], Lem. 4.1, Cor. 4.1, Rem. 4.1). As a consequence, if the controls are not bounded, the infimum of the time for moving inside $\mathcal{I}_{\widehat{q}_0}$ is zero, therefore we can think that, in some sense, if T is the minimum time for joining $\widehat{q}_0$ with $\widehat{q}_f$, then it is also the infimum of the time for reaching $\widehat{q}_f$ from $\mathcal{I}_{\widehat{q}_0}$.

This digression suggests us the suitable coercivity assumption for this problem.

Assumption 6. For $(\boldsymbol{\epsilon}, \boldsymbol{w}(\cdot)) \in \mathbb{R}^R \times L^2([0, \widehat{T}], \mathbb{R}^m)$, let $\zeta(\cdot)$ be the solution of the control system

$$\begin{cases} \dot{\zeta}(t) = \sum_{i=1}^m w_i(t) \dot{g}_t^i(\widehat{q}_0) \\ \zeta(0) = \sum_{i=1}^R \epsilon_i f_i(\widehat{q}_0), \end{cases} \tag{3.13}$$

where $f_{m+1}, \dots, f_R$ are some locally defined vector fields chosen to complete the basis for $\mathrm{Lie}(\mathfrak{f})$ in a neighborhood of $\widehat{q}_0$. The quadratic form

$$J''[(\boldsymbol{\epsilon}, \boldsymbol{w}(\cdot))]^2 = \frac{1}{2} \sum_{i=1}^m \left(\int_0^{\widehat{T}} 2 w_i(t) L_{\zeta(t)} L_{\dot{g}_t^i} \beta(\widehat{q}_0) + \sum_{j=1}^m w_i(t) w_j(t) L_{[\dot{g}_t^i, g_t^j]} \beta(\widehat{q}_0) \, \mathrm{d}t \right) \tag{3.14}$$

is coercive on the subspace $\mathcal{W}$ of $\mathbb{R}^R \times L^2([0, \widehat{T}], \mathbb{R}^m)$ defined by the constraint $\zeta(\widehat{T}) = 0$.

3.2. Consequences of the coercivity assumption

Let us now introduce a special coordinate frame in a neighborhood of $\widehat{q}_0$, completing the set $\{f_1, \dots, f_m\}$ with $n-m$ locally defined vector fields $f_{m+1}, \dots, f_n$ such that $\{f_1(q), \dots, f_n(q)\}$ is a basis for $T_q M$, and $\{f_1, \dots, f_R\}$ is a basis for $\mathrm{Lie}(\mathfrak{f})$, in a neighborhood of $\widehat{q}_0$. The coordinate frame is the inverse of the map $\Upsilon : \mathbb{R}^n \to M$ defined as

$$\Upsilon(x_1, \dots, x_n) = \exp(x_1 f_1) \circ \exp(x_2 f_2) \circ \cdots \circ \exp(x_n f_n)(\widehat{q}_0). \tag{3.15}$$

In particular, $\Upsilon^{-1}(\widehat{q}_0) = (0, \dots, 0)$, for $j = 1, \dots, n$ we have

$$\left(\Upsilon_* \frac{\partial}{\partial x_j} \right) \Big|_{(0,\dots,0)} = f_j(\widehat{q}_0)$$

and

$$L_f x_i \equiv 0 \,, \ \forall f \in \mathrm{Lie}(\mathfrak{f}) \,, \ i = R+1, \dots, n. \tag{3.16}$$

If we denote with $(\widehat{p}_1, \dots, \widehat{p}_n)$ the coefficients of $\widehat{\ell}_0$ in this coordinate frame, then it is easy to see that $\widehat{\ell}_0 = \sum_{i=R+1}^n \widehat{p}_i \mathrm{d}x_i$, since $\widehat{\ell}_0$ belongs to $\mathcal{S}$ and satisfies the HOGC.

Define the symmetric 2-form

$$\Omega = \frac{1}{2} \sum_{i=R+1}^n \mathrm{d}x_i \otimes \mathrm{d}x_i,$$

and extend it on the whole $T_{\widehat{q}_0} M \times L^2([0, \widehat{T}], \mathbb{R}^m)$ putting $\Omega[(\delta x, \boldsymbol{w}(\cdot))]^2 = \frac{1}{2} \sum_{i=R+1}^n (\delta x_i)^2$.

Then, if Assumption 6 holds, we can apply ([19], Thm. 13.2) to conclude that there exists a $\rho > 0$ such that the form

$$\boldsymbol{J}''_\rho = J'' + \rho\Omega \tag{3.17}$$

is coercive on the subspace $\widetilde{\mathcal{W}} \in T_{\widehat{q}_0}M \times L^2([0,\widehat{T}],\mathbb{R}^m)$ of the variations such that the solutions of the system (3.11) satisfy

$$\zeta(0) \in T_{\widehat{q}_0}M, \qquad \zeta(\widehat{T}) = 0. \tag{3.18}$$

Namely, $\boldsymbol{J}''_\rho$ represents the second variation of the linear-quadratic problem associated with the variable ζ, with *free* initial condition and fixed final condition.

The Hamiltonian $H''_t : T^*_{\widehat{q}_0}M \times T_{\widehat{q}_0}M \to \mathbb{R}$ associated with this linear-quadratic problem is given by

$$H''_t(\omega,\delta x) = \frac{1}{2}\mathbb{L}^{-1}_{\widehat{\lambda}(t)}\left[\begin{pmatrix} \langle\omega,\dot{g}^1_t\rangle + L_{\delta x}L_{\dot{g}^1_t}\beta(\widehat{q}_0) \\ \vdots \\ \langle\omega,\dot{g}^m_t\rangle + L_{\delta x}L_{\dot{g}^m_t}\beta(\widehat{q}_0)\end{pmatrix}\right]^2, \tag{3.19}$$

where and $\mathbb{L}^{-1}_{\widehat{\lambda}(t)}$ has to be thought of as a quadratic form on $\mathbb{R}^m$, for every t (see [2,11]).

Set

$$L'' = \{(-2\rho\Omega(\delta x,\cdot),\delta x) : \delta x \in T_{\widehat{q}_0}M\}. \tag{3.20}$$

The quadratic form (3.17) is coercive on the space $\widetilde{\mathcal{W}}$ if and only if

$$\ker \pi\mathcal{H}''_t|_{L''} = \{0\} \tag{3.21}$$

for every $t \in [0,\widehat{T}]$ (see [28]).

4. Geometry near the reference extremal

In this section we state the geometric properties of the vector fields and the Hamiltonians linked to our system and we define the super-Hamiltonian, stating its properties.

In a neighborhood of the reference extremal, Σ can be described as $\mathcal{S} \times [-\epsilon,\epsilon]^m$, for some $\epsilon > 0$. Indeed, for $\boldsymbol{t} = (t_1,\dots,t_m) \in \mathbb{R}^m$, let us denote with $\boldsymbol{tF}$ the vector field $\sum_{i=1}^m t_i\boldsymbol{F}_i$, and let us consider the map $\ell \mapsto \exp(\boldsymbol{tF})(\ell)$ (that is, the solution at time $t = 1$ of the equation $\dot{\ell} = \sum_{i=1}^m t_i\boldsymbol{F}_i(\ell)$). For a sufficiently small $\epsilon > 0$, the map

$$\psi : (\ell,\boldsymbol{t}) \in \mathfrak{U} \times [-\epsilon,\epsilon]^m \mapsto \exp(\boldsymbol{tF})(\ell) \in T^*M$$

is well defined and it is easy to prove the following result:

Proposition 4.1. *Possibly restricting $\mathfrak{U}$, there exists an $\epsilon > 0$ such that $\psi : \mathcal{S} \times [-\epsilon,\epsilon]^m \to \Sigma$ is a diffeomorphism.*

Proof. Since the whole Lie algebra generated by the $\boldsymbol{F}_i$ is tangent to Σ, then the range of the map ψ restricted to $\mathcal{S} \times [-\epsilon,\epsilon]$ is contained in Σ. The thesis follows by compactness of the interval $[0,\widehat{T}]$, since $D\psi(\ell,\mathbf{0}) = \mathrm{id} \times (\boldsymbol{F}_1(\ell),\dots,\boldsymbol{F}_m(\ell))$ has maximal rank (see property (P3) at the end of Sect. 2). □

Remark 4.2. It is easy to see that

$$\partial_{t_i}\psi(\ell,\mathbf{0}) = \boldsymbol{F}_i(\ell), \qquad i = 1,\dots,m \tag{4.1}$$

and that for each function F defined on $\mathfrak{U}$

$$\partial^2_{t_it_j}(F\circ\psi)(\ell,\mathbf{0}) = \frac{1}{2}\left(L_{\boldsymbol{F}_i}L_{\boldsymbol{F}_j} + L_{\boldsymbol{F}_j}L_{\boldsymbol{F}_i}\right)F(\ell), \qquad i,j = 1,\dots,m. \tag{4.2}$$

We remark also that ψ maps $\Sigma \times [-\epsilon,\epsilon]^m$ into Σ.

For singular extremals, the maximized Hamiltonian is well-defined and coincides with F_0, but its associated Hamiltonian vector field is multi-valued: indeed, all the Hamiltonians of the form $F_0 + \sum_{i=1}^m u_i F_i$, $\mathbf{u} \in U$, coincide and realize the maximum along any extremal contained in Σ. Moreover, no selection of such multi-valued Hamiltonian vector fields is suitable to construct the field of non-intersecting state-extremals that we will use to compare the costs associated with the candidate trajectories. For an insight in the single-input case (see [25], Sect. 4). Then, as already done in [11,25,26], we substitute the maximized Hamiltonian $F_{\max}$ with a the time-dependent super-Hamiltonian $H_t = H_0 + \sum_{i=1}^m \widehat{u}_i(t) F_i$, where H_0 is defined as described below.

The first step is to define a suitable map which turns out to project Σ onto $\mathcal{S}$.

Lemma 4.3. *Possibly restricting $\mathfrak{U}$, there exist m smooth functions $\vartheta_i : \mathfrak{U} \to \mathbb{R}$, $i = 1, \dots, m$, such that, denoting with $\boldsymbol{\vartheta F}(\ell)$ the vector field $\sum_{i=1}^m \vartheta_i \boldsymbol{F}_i$, the map*

$$\phi : \ell \in \mathfrak{U} \mapsto \exp(\boldsymbol{\vartheta F})(\ell) \in \mathfrak{U}$$

satisfies

$$\phi(\Sigma) \subset \mathcal{S}. \tag{4.3}$$

Moreover, for every $\ell \in \mathcal{S}$, $\delta\ell \in T_\ell T^*M$, it holds

$$\begin{pmatrix} \langle d\vartheta_1(\ell), \delta\ell \rangle \\ \vdots \\ \langle d\vartheta_m(\ell), \delta\ell \rangle \end{pmatrix} = \mathbb{L}_\ell^{-1} \begin{pmatrix} \langle dF_{01}(\ell), \delta\ell \rangle \\ \vdots \\ \langle dF_{0m}(\ell), \delta\ell \rangle \end{pmatrix}. \tag{4.4}$$

Proof. Let us consider the function $\Phi : \mathfrak{U} \times \mathbb{R}^m \to \mathbb{R}^m$ defined by

$$\Phi(\ell, \boldsymbol{t}) = \begin{pmatrix} F_{01} \circ \psi(\ell, \boldsymbol{t}) \\ \vdots \\ F_{0m} \circ \psi(\ell, \boldsymbol{t}) \end{pmatrix}.$$

Notice that for every $\ell \in \mathcal{S}$ we have $\psi(\ell, \mathbf{0}) = \ell$ and then $\Phi(\ell, \mathbf{0}) = \mathbf{0}$. Moreover, using (4.1), it is easy to show that $\partial_{\boldsymbol{t}} \Phi(\ell, \mathbf{0}) = -\mathbb{L}_\ell$, for all $\ell \in \mathcal{S}$, and then it has rank m. The implicit function theorem and the compactness of the interval $[0, \widehat{T}]$ ensure the existence of m smooth functions $\vartheta_1, \dots, \vartheta_m$, defined in a neighborhood of the reference extremal, such that

$$F_{0i} \circ \phi(\ell) \equiv 0\,, \ i = 1, \dots, m. \tag{4.5}$$

Without loss of generality, we can assume that this neighborhood is $\mathfrak{U}$. Since $\exp(\boldsymbol{\vartheta F})(\ell) \in \Sigma$ for every $\ell \in \Sigma$, then (4.5) implies (4.3).

Fix $\ell \in \mathcal{S}$; recalling that $\boldsymbol{\vartheta}(\ell) = \mathbf{0}$, thanks to (4.5) and (4.1), we get for every $\delta\ell \in T_\ell T^*M$ and $i = 1, \dots, m$

$$\begin{aligned} 0 &= \langle dF_{0i}(\phi(\ell)), \partial_\ell \psi(\ell, \boldsymbol{t})|_{\boldsymbol{t} = \boldsymbol{\vartheta}(\ell)} \delta\ell \rangle + \sum_{j=1}^m \langle dF_{0i}(\phi(\ell)), \partial_{t_j} \psi(\ell, \boldsymbol{t})|_{\boldsymbol{t} = \boldsymbol{\vartheta}(\ell)} \rangle \langle d\vartheta_j(\ell), \delta\ell \rangle \\ &= \langle dF_{0i}(\ell), \delta\ell \rangle + \sum_{j=1}^m F_{j0i}(\ell) \langle d\vartheta_j(\ell), \delta\ell \rangle \end{aligned}$$

and then (4.4). □

Remark 4.4. It is easy to see that, as a particular case of (4.4), we get

$$\langle d\vartheta_i(\ell), \boldsymbol{F}_j(\ell) \rangle = -\delta_{ij} \qquad \forall\, \ell \in \mathcal{S}, \quad i, j = 1, \dots, m. \tag{4.6}$$

The Hamiltonian H_0 is defined by means of the map ϕ, as follows.

Definition 4.5. We define the Hamiltonian function $H_0 : \mathfrak{U} \to \mathbb{R}$ as

$$H_0(\ell) = F_0 \circ \phi(\ell)$$

and we set

$$\chi = H_0 - F_0.$$

In order to prove the properties of H_0, we start by computing the first and second derivatives of χ on $\mathcal{S}$.

Proposition 4.6. *For every $\ell \in \mathcal{S}$ and $\delta\ell \in T_\ell T^*M$, it holds $\langle d\chi(\ell), \delta\ell\rangle = 0$ and*

$$D^2\chi(\ell)[\delta\ell]^2 = -\sum_{r,s=1}^{m} (\mathbb{L}_\ell^{-1})_{rs} \langle dF_{0r}(\ell), \delta\ell\rangle \langle dF_{0s}(\ell), \delta\ell\rangle. \tag{4.7}$$

Proof. By definition we get

$$d\chi(\ell) = dF_0(\phi(\ell))\Big(\partial_\ell \psi(\ell, \boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)} + \sum_{j=1}^{m} \partial_{t_j}\psi(\ell, \boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)} d\vartheta_j(\ell)\Big) - dF_0(\ell). \tag{4.8}$$

For $\ell \in \mathcal{S}$, by (4.1) and thanks to the fact that $\boldsymbol{\vartheta}(\ell) = 0$ on $\mathcal{S}$, we obtain

$$d\chi(\ell) = dF_0(\ell)\Big(\mathrm{id} + \sum_{i=1}^{m} \boldsymbol{F}_i(\ell) d\vartheta_i(\ell)\Big) - dF_0(\ell) = 0.$$

Therefore $D^2\chi(\ell)$ is a well defined quadratic form for every $\ell \in \mathcal{S}$. To prove (4.7), we perform the computations in any chart. Fix $\ell \in \mathcal{S}$ and $\delta\ell \in T_\ell T^*M$ and recall (4.1), (4.2); from (4.8) it follows the following expression:

$$\begin{aligned} D^2\chi(\ell)[\delta\ell]^2 = {} & D^2F_0(\ell)[\delta\ell + \sum_{i=1}^{m} \boldsymbol{F}_i(\ell)\langle d\vartheta_i(\ell), \delta\ell\rangle]^2 - D^2F_0(\ell)[\delta\ell]^2 \\ & + dF_0(\ell)\Big(\partial^2_{\ell\,\ell}\psi(\ell, \mathbf{0}) + \sum_{i=1}^{m} \boldsymbol{F}_i(\ell) D^2\vartheta_i(\ell)\Big)[\delta\ell]^2 \\ & + dF_0(\ell)\Big(2\sum_{i=1}^{m} \partial^2_{\ell\,t_i}\psi(\ell, \mathbf{0})[\delta\ell]\,\langle d\vartheta_i(\ell), \delta\ell\rangle \\ & + \sum_{i,j=1}^{m} \partial^2_{t_j\,t_i}\psi(\ell, \mathbf{0})\langle d\vartheta_j(\ell), \delta\ell\rangle\langle d\vartheta_i(\ell), \delta\ell\rangle\Big). \end{aligned}$$

First of all, we notice that the first line in the right hand side of the equation is equal to

$$2\sum_{i=1}^{m} D^2F_0(\ell)[\delta\ell, \boldsymbol{F}_i(\ell)]\langle d\vartheta_i(\ell), \delta\ell\rangle + \sum_{i,j=1}^{m} D^2F_0(\ell)[\boldsymbol{F}_i(\ell), \boldsymbol{F}_j(\ell)]\langle d\vartheta_j(\ell), \delta\ell\rangle\langle d\vartheta_i(\ell), \delta\ell\rangle.$$

Since $\partial^2_{\ell\,\ell}\psi(\ell, \mathbf{0}) = 0$ and $\langle dF_0(\ell), \boldsymbol{F}_i(\ell)\rangle = F_{0i}(\ell) = 0$ on $\mathcal{S}$, we deduce that

$$dF_0(\ell)\Big(\partial^2_{\ell\,\ell}\psi(\ell, \mathbf{0}) + \sum_{i=1}^{m} \boldsymbol{F}_i(\ell) D^2\vartheta_i(\ell)\Big)[\delta\ell]^2 = 0.$$

Finally, from $\partial_{t_i}\psi(\ell,\mathbf{0}) = \boldsymbol{F}_i(\ell)$, we obtain that $\partial^2_{\ell t_i}\psi(\ell,\mathbf{0})[\delta\ell] = \langle D\boldsymbol{F}_i(\ell),\delta\ell\rangle$. Then we have that

$$\begin{aligned} D^2\chi(\ell)[\delta\ell]^2 = &\, 2\sum_{i=1}^m \left(D^2F_0(\ell)[\delta\ell,\boldsymbol{F}_i(\ell)] + \langle dF_0(\ell), D\boldsymbol{F}_i(\ell)[\delta\ell]\rangle\right)\langle d\vartheta_i(\ell),\delta\ell\rangle \\ &+ \sum_{i,j=1}^m \left(D^2F_0(\ell)[\boldsymbol{F}_i(\ell),\boldsymbol{F}_j(\ell)] + dF_0(\ell)\partial^2_{t_jt_i}\psi(\ell,\mathbf{0})\right)\langle d\vartheta_j(\ell),\delta\ell\rangle\langle d\vartheta_i(\ell),\delta\ell\rangle \end{aligned}$$

that can be written intrinsically as

$$\begin{aligned} D^2\chi(\ell)[\delta\ell]^2 &= 2\sum_{i=1}^m L_{\delta\ell}L_{\boldsymbol{F}_i}F_0(\ell)\langle d\vartheta_i(\ell),\delta\ell\rangle + \sum_{i,j=1}^m \partial^2_{t_j\,t_i}(F_0\circ\psi)(\ell,\mathbf{0})\langle d\vartheta_j(\ell),\delta\ell\rangle\langle d\vartheta_i(\ell),\delta\ell\rangle \\ &= -2\sum_{i=1}^m \langle dF_{0i}(\ell),\delta\ell\rangle\langle d\vartheta_i(\ell),\delta\ell\rangle + \sum_{i,j=1}^m F_{ji0}(\ell)\langle d\vartheta_j(\ell),\delta\ell\rangle\langle d\vartheta_i(\ell),\delta\ell\rangle. \end{aligned}$$

Thanks to (4.4), this is equal to

$$D^2\chi(\ell)[\delta\ell]^2 = -\sum_{i,j=1}^m F_{ji0}(\ell)\langle d\vartheta_j(\ell),\delta\ell\rangle\langle d\vartheta_i(\ell),\delta\ell\rangle = -\sum_{r,s=1}^m (\mathbb{L}_\ell^{-1})_{rs}\langle dF_{0r}(\ell),\delta\ell\rangle\langle dF_{0s}(\ell),\delta\ell\rangle. \qquad \square$$

Thanks to its definition, H_0 satisfies the following properties.

Theorem 4.7. *The Hamiltonian H_0 has the following properties.*

(1) *$F_0 = H_0$ and $\boldsymbol{F}_0 = \boldsymbol{H}_0$ on $\mathcal{S}$.*
(2) *The vector field $\boldsymbol{H}_0$ is tangent to Σ.*
(3) *$F_0 \le H_0$ on Σ.*

Proof. Since $\phi|_{\mathcal{S}}$ is the identity, then $H_0 = F_0$ on $\mathcal{S}$. Moreover

$$\begin{aligned} dH_0(\ell) &= dF_0\circ\left(\partial_\ell\psi(\ell,\boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)} + \sum_{i=1}^m \partial_{t_i}\psi(\ell,\boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)}d\vartheta_i(\ell)\right) \\ &= dF_0\circ\left(\mathrm{id} + \sum_{i=1}^m \boldsymbol{F}_i(\ell)d\vartheta_i(\ell)\right) = dF_0(\ell), \end{aligned} \tag{4.9}$$

since $\boldsymbol{\vartheta}(\ell) = \mathbf{0}$ and $F_{0i}(\ell) = 0$ on $\mathcal{S}$. This ends the proof of (1).

To prove (2), fix $\ell\in\Sigma$ and observe that $\partial_{t_i}\psi(\ell,\boldsymbol{t}) \in \mathrm{Lie}(\mathfrak{f})(\psi(\ell,\boldsymbol{t}))$, $i = 1,\dots,m$, so that by Assumption 5 we have

$$\langle dF_0(\phi(\ell)), \partial_{t_i}\psi(\ell,t)|_{t=\boldsymbol{\vartheta}(\ell)}\rangle = 0.$$

Therefore (4.9) leads to

$$dH_0(\ell) = dF_0(\phi(\ell))\circ\partial_\ell\psi(\ell,\boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)}, \qquad \ell\in\Sigma,$$

and then we get easily

$$\boldsymbol{H}_0(\ell) = \left[\partial_\ell\psi(\ell,\boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)}\right]^{-1}\boldsymbol{F}_0(\phi(\ell)).$$

Since $\phi(\ell)\in\mathcal{S}$, $\boldsymbol{F}_0(\phi(\ell)) \in T_{\phi(\ell)}\Sigma$ and $\left[\partial_\ell\psi(\ell,\boldsymbol{t})|_{\boldsymbol{t}=\boldsymbol{\vartheta}(\ell)}\right]^{-1}$ maps $T_{\phi(\ell)}\Sigma$ onto $T_\ell\Sigma$, then (2) is proved.

Statement (3) is a straight consequence of Proposition 4.6, since for all $\ell\in\mathcal{S}$, $\delta\ell\in T_\ell\mathcal{S}$ and $i,j = 1,\dots,m$, one gets $D^2\chi(\ell)[\delta\ell]^2 = D^2\chi(\ell)[\delta\ell,\boldsymbol{F}_i] = 0$ and $D^2\chi(\ell)[\boldsymbol{F}_j,\boldsymbol{F}_i] = -F_{ij0}$. $\square$

We finally define the super-maximized Hamiltonian H_t as follows.

Definition 4.8. We denote with H_t the following time-dependent Hamiltonian:

$$H_t(\ell) = H_0(\ell) + \sum_{i=1}^{m} \widehat{u}_i(t) F_i(\ell), \tag{4.10}$$

and with $\mathcal{H}_t$ the Hamiltonian flow generated by H_t.

Notice that $\boldsymbol{H}_t$ is tangent to Σ and to the reference extremal $\widehat{\lambda}(\cdot)$.

5. The result

In this section we state and prove our main result. It relies on the following Hamiltonian sufficient conditions, that we state and prove here below.

Theorem 5.1 (Geometric sufficient condition)**.** *Let $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ be an admissible triple for the minimum-time problem (1.1)–(1.2) with associate adjoint covector $\widehat{\lambda}$, and let Assumptions 1–5 be satisfied. Suppose that there exist a neighborhood V of $\widehat{q}_0$ and a smooth function $\mathfrak{a} : V \to \mathbb{R}$ with the following properties:*

(i) $d\mathfrak{a}(\widehat{q}_0) = \widehat{\lambda}(0)$
(ii) *the Lagrangian submanifold $\Lambda = \{d\mathfrak{a}(q) : q \in V\}$ is contained in Σ and satisfies*

$$\ker \pi_* \mathcal{H}_{t*} \cap T_{\widehat{\lambda}(0)}\Lambda = \{0\} \qquad \forall\, t \in [0, \widehat{T}]. \tag{5.1}$$

Then $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ is a strict strong-local minimizer for the minimum-time problem (1.2) between $\mathcal{I}_{\widehat{q}_0}$ and $\widehat{q}_f$ (or between $\widehat{q}_0$ and $\mathcal{I}_{\widehat{q}_f}$) and a strong-local minimizer for the minimum-time problem (1.2) between $\mathcal{I}_{\widehat{q}_0}$ and $\mathcal{I}_{\widehat{q}_f}$.

Proof. Consider the map

$$\mathrm{id} \times \pi \circ \mathcal{H}_t : (t, \ell) \in [0, \widehat{T}] \times \Lambda \mapsto (t, \pi \circ \mathcal{H}_t(\ell)) \in [0, \widehat{T}] \times M.$$

Hypothesis (ii) and the compactness of the reference trajectory imply that (possibly restricting V) there exists a neighborhood $\mathcal{O}$ of the graph of $\widehat{\xi}(\cdot)$ in $[0, \widehat{T}] \times \Lambda$ such that $\mathrm{id} \times \pi \circ \mathcal{H}$ is a diffeomorphism between $[0, \widehat{T}] \times \Lambda$ and $\mathcal{O}$.

Let $(\xi, \mathbf{v}, T)$ be an admissible triple for the control system (1.2) such that its graph is contained in $\mathcal{O}$, $\xi(0) \in \mathcal{I}_{\widehat{q}_0}$, $\xi(T) \in \mathcal{I}_{\widehat{q}_f}$, and $T \le \widehat{T}$, and denote its lift on Λ with $\ell(t)$, namely

$$\ell(t) = (\pi \circ \mathcal{H}_t)^{-1}(\xi(t)).$$

Let us choose a curve $\mu_0 : [0, 1] \to \Lambda$ joining $\widehat{\ell}_0$ with $\ell(0)$ satisfying $\pi\mu_0(t) \in \mathcal{I}_{\widehat{q}_0}$, $\forall\, t \in [0, 1]$, and a curve $\mu_f : [T, \widehat{T}] \to \Lambda$ joining $\widehat{\ell}_0$ with $\ell(T)$ satisfying $\pi \circ \mathcal{H}_t(\mu_f(t)) \in \mathcal{I}_{\widehat{q}_f}$, $\forall\, t \in [T, \widehat{T}]$. Let us now define the following paths in $[0, \widehat{T}] \times \Lambda$:

$$\begin{aligned}
\gamma_1 &= (t, \widehat{\ell}_0) && t \in [0, \widehat{T}] \\
\gamma_2 &= (0, \mu_0(t)) && t \in [0, 1] \\
\gamma_3 &= (t, \ell(t)) && t \in [0, T] \\
\gamma_4 &= (t, \mu_f(t)) && t \in [T, \widehat{T}],
\end{aligned}$$

and let $\gamma = (-\gamma_1) \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$.

Consider the following 1-form on $[0, \widehat{T}] \times T^*M$

$$\omega(t, \ell) = \mathcal{H}_t^* \varsigma - H_t \circ \mathcal{H}_t(\ell) \mathrm{d}t, \tag{5.2}$$

where ς is the canonical Liouville form on T^*M. It is easy to prove (see [1], Prop. 17.1) that ω is exact on $[0,\widehat{T}]\times\Lambda$. In particular, since γ is a closed path contained in $[0,\widehat{T}]\times\Lambda$, then $\int_\gamma\omega=0$. Let us now evaluate the single components of this integral. It is easy to see that

$$\int_{\gamma_1}\omega=\int_{\gamma_2}\omega=0.$$

Moreover

$$\begin{aligned}\int_{\gamma_3}\omega&=\int_0^T\langle\mathcal{H}_t(\ell(t)),\dot{\xi}(t)\rangle-H_t\circ\mathcal{H}_t(\ell(t))\,\mathrm{d}t\\&=\sum_{i=1}^m\int_0^T\langle\mathcal{H}_t(\ell(t)),(\mathrm{v}_i(t)-\widehat{u}_i(t))f_i(\xi(t))\rangle\,\mathrm{d}t-\int_0^T\chi\circ\mathcal{H}_t(\ell(t))\,\mathrm{d}t\le 0.\end{aligned}$$

Therefore

$$0\le\int_{\gamma_4}\omega=\int_T^{\widehat{T}}\langle\mathcal{H}_t(\ell(t)),\pi_*\circ\mathcal{H}_{t*}\dot{\mu}_f(t)\rangle\,\mathrm{d}t-\int_T^{\widehat{T}}H_t\left(\mathcal{H}_t(\mu_f(t))\right)\,\mathrm{d}t.$$

The first term is zero, since $\pi_*\circ\mathcal{H}_{t*}\dot{\mu}_f(t)$ is tangent to $\mathcal{I}_{\pi\circ\mathcal{H}_t(\mu_f(t))}$, by construction, and $\mathcal{H}_t(\ell(t))$ is contained in Σ. Then

$$0\le\int_{\widehat{T}}^T H_t\left(\mathcal{H}_t(\mu_f(t))\right)\,\mathrm{d}t=\int_{\widehat{T}}^T 1+\mathcal{O}(t)\,\mathrm{d}t\le(T-\widehat{T})+o(T-\widehat{T}),$$

which is a contradiction, therefore $T\ge\widehat{T}$.

If the final point $\widehat{q}_f$ is fixed, we claim that the reference extremal is strictly optimal. To prove the claim, we consider an admissible triple $(\xi,\mathbf{v},T)$ as above, such that $T=\widehat{T}$ and $\xi(T)=\widehat{q}_f$.

In particular, since in this case $\int_{\gamma_4}\omega=0$, we obtain that

$$0=\int_{\gamma_3}\omega=-\int_0^T\chi\circ\mathcal{H}_t(\ell(t))\,\mathrm{d}t,$$

which implies that $\lambda(t)=\mathcal{H}_t(\ell(t))\in\mathcal{S}$ for every $t\in[0,T]$, then, in particular, that $\dot{\lambda}(t)\in T_{\lambda(t)}\mathcal{S}$, that is $\langle dF_{0j}(\lambda(t)),\dot{\lambda}(t)\rangle=0$ for $j=1,\dots,m$. By computations it is possible to show that

$$\dot{\lambda}(t)=\overrightarrow{\widehat{F}}_t(\lambda(t))+\sum_{i=1}^m(\mathrm{v}_i(t)-\widehat{u}_i(t))\,\mathcal{H}_{t*}(\pi\circ\mathcal{H}_t)_*^{-1}f_i(\xi(t)).$$

Thanks to (5.1), possibly restricting V, we can find a family of smooth functions $\mathfrak{a}^t:V^t\to\mathbb{R}$, where V^t is a neighborhood of $\widehat{\xi}(t)$ and $\mathfrak{a}^0=\mathfrak{a}$, such that $\mathcal{H}_t(\Lambda)=\{d\mathfrak{a}^t(q):q\in V^t\}$ for every $t\in[0,T]$. In particular, for every $t\in[0,T]$ we have that

$$\begin{aligned}\lambda(t)&=d\mathfrak{a}^t(\xi(t))\\\mathcal{H}_{t*}(\pi\circ\mathcal{H}_t)_*^{-1}f_i(\xi(t))&=d\mathfrak{a}^t_*f_i(\xi(t)),\quad i=1,\dots,m.\end{aligned}$$

Then

$$\begin{aligned}\langle dF_{0j}(\lambda(t)),\mathcal{H}_{t*}(\pi\circ\mathcal{H}_t)_*^{-1}f_i(\xi(t))\rangle=\sigma(d\mathfrak{a}^t_*f_i(\xi(t)),\boldsymbol{F}_{0j}(\lambda(t)))&=L_{f_i}L_{f_{0j}}\mathfrak{a}^t(\xi(t))\\&=L_{f_{0j}}L_{f_i}\mathfrak{a}^t(\xi(t))+L_{f_{i0j}}\mathfrak{a}^t(\xi(t))\\&=\langle\lambda(t),f_{i0j}(\xi(t))\rangle,\end{aligned}$$

since $L_{f_i}\mathfrak{a}^t(\xi(t))$ is identically null, being $\mathcal{H}_t(\Lambda) = d\mathfrak{a}^t(V^t)$ contained in Σ. This implies that, for every $j = 1, \dots, m$, it holds

$$0 = \langle dF_{0j}(\lambda(t)), \dot{\lambda}(t)\rangle = F_{00j}(\lambda(t)) + \sum_{i=1}^{m} \mathrm{v}_i(t) F_{i0j}(\lambda(t)),$$

that is, in particular, that $\mathbf{v}$ is a solution of equation (2.16), and therefore $\dot{\lambda}(t) = \boldsymbol{F}_{\mathcal{S}}(\lambda(t))$. Since $\widehat{\lambda}(t)$ is solution of the same equation, and both λ and $\widehat{\lambda}$ pass through $\widehat{q}_f$, then they coincide.

The same argument shows that the reference triple is a strict strong-local minimizer for the minimum-time problem (1.2) between $\widehat{q}_0$ and $\mathcal{I}_{\widehat{q}_f}$. □

Now we state and prove the main result.

Theorem 5.2. *Let $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ be an admissible triple of the minimum-time problem (1.1)–(1.2) with associate adjoint covector $\widehat{\lambda}$, and let $\widehat{\lambda}$ be a normal singular extremal. If Assumptions 1–6 are satisfied, then $\widehat{\xi}(\cdot)$ is a minimum-time trajectory between $\mathcal{I}_{\widehat{q}_0}$ and $\mathcal{I}_{\widehat{q}_f}$, and hence between N_0 and N_f. Moreover, the reference trajectory is strictly optimal among all admissible trajectories between $\widehat{q}_0$ and $\mathcal{I}_{\widehat{q}_f}$ and among all admissible trajectories between $\mathcal{I}_{\widehat{q}_0}$ and $\widehat{q}_f$.*

Proof. The thesis comes straightaway once proved that the coercivity assumption (Assumption 6) allows us to define a smooth function α_ρ that satisfies the hypotheses of Theorem 5.1. In particular, using the adapted coordinates (3.15) introduced in Section 3.2, we define α_ρ as follows:

$$\alpha_\rho(x) = \sum_{i=R+1}^{n} \widehat{p}_i x_i + \frac{\rho}{2} \sum_{i=R+1}^{n} x_i^2.$$

It is easy to see that α_ρ satisfies property (i) and that (3.16) implies that Λ is a Lagrangian submanifold contained in Σ. To prove (5.1) we need to exploit the links between the flow of the Hamiltonian H_t'' defined in equation (3.19) and $\mathcal{H}_{t*}$, as done in [11, 25, 27].

It is known that the pull-back flow $\mathcal{G}_t = \widehat{\mathcal{F}}_t^{-1} \circ \mathcal{H}_t$ is the Hamiltonian flow relative to the Hamiltonian $G_t : T^*M \to \mathbb{R}$ defined by

$$G_t = \left(H_t - \widehat{F}_t\right) \circ \widehat{\mathcal{F}}_t = \chi \circ \widehat{\mathcal{F}}_t$$

(see [23]). Since $DG_t(\widehat{\ell}_0) = 0$, then $G_t'' = \frac{1}{2} D^2 G_t|_{\widehat{\ell}_0}$ is a well defined quadratic form and its associated Hamiltonian flow is $\mathcal{G}_{t*} : T_{\widehat{\ell}_0}(T^*M) \to T_{\widehat{\ell}_0}(T^*M)$.

Let $\beta = -\sum_{i=R+1}^{n} \widehat{p}_i x_i$; then the linear map $\iota : T^*_{\widehat{q}_0}M \times T_{\widehat{q}_0}M \to T_{\widehat{\ell}_0}(T^*M)$, defined as follows

$$\iota(\omega, \delta x) = -\omega + d(-\beta)_* \delta x,$$

establishes an anti-symplectic isomorphism between $T^*_{\widehat{q}_0}M \times T_{\widehat{q}_0}M$ and $T_{\widehat{\ell}_0}(T^*M)$. In particular, it determines an equivalence between the Hamiltonian functions G_t'' and H_t'', *i.e.* the following identities hold:

$$H_t'' = -G_t'' \circ \iota \tag{5.3}$$

$$\overrightarrow{H_t''} = \iota^{-1} \circ \overrightarrow{G_t''} \circ \iota$$

$$\mathcal{H}_t'' = \iota^{-1} \circ \mathcal{G}_{t*} \circ \iota. \tag{5.4}$$

We need to prove only (5.3), since the other two equations are a direct consequence (see [11] and references therein for details).

Consider $\ell \in \mathcal{S}$, $\delta\ell \in T_\ell(T^*M)$, and set $\ell_t = \widehat{\mathcal{F}}_t(\ell)$. Then, thanks to (4.7), we have that

$$\begin{aligned} D^2G_t(\ell)[\delta\ell]^2 &= D^2\chi(\ell_t) \circ \widehat{\mathcal{F}}_{t*} \otimes \widehat{\mathcal{F}}_{t*} \\ &= -\sum_{r,s=1}^{m} (\mathbb{L}_{\ell_t}^{-1})_{rs} \langle dF_{0r}(\ell_t), \widehat{\mathcal{F}}_{t*}\delta\ell\rangle \langle dF_{0s}(\ell_t), \widehat{\mathcal{F}}_{t*}\delta\ell\rangle \\ &= -(\langle dF_{01}(\ell_t), \widehat{\mathcal{F}}_{t*}\delta\ell\rangle, \dots, \langle dF_{0m}(\ell_t), \widehat{\mathcal{F}}_{t*}\delta\ell\rangle)(\mathbb{L}_{\ell_t}^{-1}) \begin{pmatrix} \langle dF_{01}(\ell_t), \widehat{\mathcal{F}}_{t*}\delta\ell\rangle \\ \vdots \\ \langle dF_{0m}(\ell_t), \widehat{\mathcal{F}}_{t*}\delta\ell\rangle \end{pmatrix} \\ &= -2H_t'' \circ \iota^{-1}\delta\ell. \end{aligned}$$

By computations, it is easy to see that the space L'' defined in equation (3.20) satisfies the equality $\iota L'' = \{d\alpha_{\rho*}\delta x : \delta x \in T_{\widehat{q}_0}M\} = L$, therefore equations (3.21) and (5.4) imply

$$\ker \pi_* \mathcal{G}_{t*}|_L = \{0\} \qquad \forall\, t \in [0, \widehat{T}].$$

To end the proof it is sufficient to notice that $\widehat{F}(t)$ is a lifted Hamiltonian, so that $\widehat{\mathcal{F}}_{t*}$ is an isomorphism on the vertical fibers. □

6. Examples

The classical Dubins and dodgem car problems concern the motion of a car on the plane $\mathbb{R}^2$ with constant speed and controlled (bounded) angular velocity (see for instance [1, 13]). In particular, Dubins problem looks for minimum-time trajectories between fixed initial and final positions and orientations, while in the dodgem car problem the final orientation is free.

As shown in [21] (see also [12]), this problem can be reformulated on the manifold $\mathbb{R}^2 \times \mathrm{SO}(2)$, where $\mathrm{SO}(2)$ is the group of positively oriented rotations on $\mathbb{R}^2$. A great advantage of this formulation is that the extension to higher dimensions is straightforward: we denote with (q, R) the elements of $\mathbb{R}^N \times \mathrm{SO}(N)$, where $\mathrm{SO}(N)$ is the group of positively oriented rotations on $\mathbb{R}^N$, and we consider the control system

$$\begin{cases} \dot{q}(t) = R(t)e_1 \\ \dot{R}(t) = \sum_{j=1}^{N-1} u_j(t)R(t)A_j, \qquad |\mathbf{u}| \le 1, \end{cases} \tag{6.1}$$

where e_1 is the first element of the canonical basis of $\mathbb{R}^N$, and, for every $j = 1, \dots, N-1$, A_j is the anti-symmetric matrix defined by

$$(A_j)_{lm} = \begin{cases} -1 & \text{if } l = 1, m = j+1 \\ 1 & \text{if } l = j+1, m = 1 \\ 0 & \text{otherwise.} \end{cases}$$

This system models the motion of a point in the N-dimensional space with constant speed equal to 1, where we control the orientation velocity. For both Dubins and dodgem car problems, the initial condition consists in fixing the initial point q_0 and the initial (unit length) velocity $\mathbf{v}_0$ of the trajectory on $\mathbb{R}^N$. Namely, $N_0 = \{q_0\} \times \mathcal{I}_{\mathbf{v}_0}$, where $\mathcal{I}_{\mathbf{v}} = \{S \in \mathrm{SO}(N) : Se_1 = \mathbf{v}\}$. Dubins problem looks for minimum-time trajectories joining N_0 with $N_f = \{q_f\} \times \mathcal{I}_{\mathbf{v}_f}$, for some fixed $q_f \in \mathbb{R}^N$, $\mathbf{v}_f \in \mathbb{R}^N$ with $|\mathbf{v}_f| = 1$, while dodgem car problem searches minimum-time trajectories from N_0 to $N_f = \{q_f\} \times \mathrm{SO}(N)$, for some fixed $q_f \in \mathbb{R}^N$.

The system (6.1) can be embedded in the matrix group $\mathrm{GL}(N+1)$ (non singular $(N+1)$-dimensional matrices), via the map

$$(q, R) \in \mathbb{R}^N \times \mathrm{SO}(N) \mapsto g = \begin{pmatrix} 1 & 0 \\ q & R \end{pmatrix} \in \mathrm{GL}(N+1).$$

This formulation is suitable also to consider the Dubins problem on other homogeneous spaces different from $\mathbb{R}^N$, that is the N-dimensional sphere S^N and the N-dimensional hyperbolic space $\mathbb{H}^N$, defined as $\mathbb{H}^N = \{\boldsymbol{x} \in \mathbb{R}^{N+1} : -x_0^2 + \sum_{i=1}^N x_i^2 = 1,\ x_0 > 0\}$. Briefly, a pair *(point, orientation)* in $S^N \times \mathrm{SO}(N)$ can be represented in the group $G = \mathrm{SO}(N+1)$ in the following way: the first column of the matrix $g \in \mathrm{SO}(N+1)$ gives the coordinate representation (in $\mathbb{R}^{N+1}$) of the point, the other N columns determine an orthonormal frame in the tangent space to the sphere at the point. Analogously, a pair *(point, orientation)* in $\mathbb{H}^N \times \mathrm{SO}(N)$ can be represented in the group $G = \mathrm{SO}(1,n)$, as above: the first column of the matrix g gives the coordinate representation (in $\mathbb{R}^{N+1}$) of the point, while the other N columns determine an orthonormal frame in the tangent space. More details on these representations can be found in Appendix 6 and in [20, 21], where the authors study the geodesic problem for curves with bounded curvature.

We can then write Dubins problem on $M \in \{\mathbb{R}^n, S^N, \mathbb{H}^N\}$ in the following unified way:

$$\min T \tag{6.2}$$

subject to

$$\begin{cases} \dot{g}(t) = g(t)\begin{pmatrix} 0 & -\varepsilon e_1^{\mathsf{T}} \\ e_1 & 0 \end{pmatrix} + \sum_{i=1}^{N-1} u_i(t) g(t) \begin{pmatrix} 0 & 0 \\ 0 & A_i \end{pmatrix}, \qquad |\mathbf{u}| \le 1 \\ g \in G \\ g(0) \in N_0,\ g(T) \in N_f, \end{cases} \tag{6.3}$$

where G, ε and the manifolds of the constraints depend on the manifold M as shown below:

M	ε	G	N_0	N_f
$\mathbb{R}^N$	$\varepsilon = 0$	$\mathbb{R}^N \rtimes \mathrm{SO}(N)$	$\left\{g : g\mathfrak{e}_1 = \begin{pmatrix} 1 \\ \widehat{q}_0 \end{pmatrix},\ g\mathfrak{e}_2 = \begin{pmatrix} 0 \\ \boldsymbol{v}_0 \end{pmatrix}\right\}$	$\left\{g : g\mathfrak{e}_1 = \begin{pmatrix} 1 \\ \widehat{q}_f \end{pmatrix},\ g\mathfrak{e}_2 = \begin{pmatrix} 0 \\ \boldsymbol{v}_f \end{pmatrix}\right\}$
S^N	$\varepsilon = 1$	$\mathrm{SO}(N+1)$	$\{g : g\mathfrak{e}_1 = \widehat{q}_0,\ g\mathfrak{e}_2 = \boldsymbol{v}_0\}$	$\{g : g\mathfrak{e}_1 = \widehat{q}_f,\ g\mathfrak{e}_2 = \boldsymbol{v}_f\}$
$\mathbb{H}^N$	$\varepsilon = -1$	$\mathrm{SO}(1,N)$	$\{g : g\mathfrak{e}_1 = \widehat{q}_0,\ g\mathfrak{e}_2 = \boldsymbol{v}_0\}$	$\{g : g\mathfrak{e}_1 = \widehat{q}_f,\ g\mathfrak{e}_2 = \boldsymbol{v}_f\}$

where $\mathfrak{e}_i$ denotes the ith element of the canonical basis of $\mathbb{R}^{N+1}$.

The control system (6.3) is a control-affine system of the form (1.2) with $m = N-1$ and $\dim M = N(N+1)/2$, and the corresponding left-invariant vector fields are defined by $f_i(g) = gA_i$, $i = 0, \dots, N-1$. We recall that left invariant vector fields satisfy the following relation: $g[A_i, A_j] = [f_i, f_j](g)$, where $[\cdot,\cdot]$ denotes also the usual matrix commutator. Thanks to this equation, the commutation properties of the matrices A_i extend also to their associated left-invariant vector fields. In particular, the following properties are easily verified:

(i) Lie($\{A_i : i = 1, \dots, N-1$) is 2-step bracket-generating and isomorphic to $\mathfrak{so}(N)$ (Lie algebra of antisymmetric N-dimensional matrices); $\mathfrak{so}(N)$ has dimension $R = N(N-1)/2$;
(ii) the matrices $\{[A_i, A_j] : i, j = 1, \dots, N-1\}$ generate the derived sub-algebra $[\mathfrak{so}(N), \mathfrak{so}(N)]$, which is isomorphic to $\mathfrak{so}(N-1)$ and has dimension $\frac{(N-1)(N-2)}{2}$;
(iii) the matrices $\{A_0, [A_0, A_i], A_i, [A_i, A_j] : i < j = 1, \dots, N-1\}$ are linearly independent and form a basis for the Lie algebra of G;
(iv) for $i, j = 1, \dots, N-1$ the matrix commutators of the kind $[A_i, [A_j, A_0]]$ satisfy the following relations:

$$[A_i, [A_i, A_0]] = -A_0$$
$$[A_i, [A_j, A_0]] = 0 \quad \text{if } i \neq j.$$

(v) the matrices $\{A_0, A_{01}, \dots, A_{0m}\}$ mutually commute.
(vi) A_0 commutes with every element of $\{[A_i, A_j] : i, j = 1, \dots, N-1\}$.

We notice that the submanifolds N_0 and N_f are integral manifolds of the derived sub-algebra $\{[f_i, f_j] : i, j = 1, \dots, m\}$. Indeed, it is easy to verify that the Lie sub-algebra is contained in the tangent spaces of N_0 and N_f; a dimensional computation proves the claim.

Let us now consider singular extremal for the problem (6.2)–(6.3). First of all, we remark that this problem does not admit abnormal singular extremals, thanks to property (iii). Moreover, thanks to property (v) and from equation (2.13), we get that the reference control $\widehat{\mathbf{u}}(\cdot)$ is identically zero, and that the matrix $\mathbb{L}_{\widehat{\lambda}(t)} = -F_0(\widehat{\lambda}(t))\mathbb{I}_n$. In particular, singular trajectories are the integral curves of the drift f_0; with each of these curves we associate the adjoint covector $p(t)$ that satisfies the differential equation $\dot{p}(t) = -p\partial_q f_0$, with initial condition $p(0) \in \{f_i, f_{ij}, f_{0i} : \ i, j = 1, \dots, n-1\}^\perp$ and $\langle p(0), f_0\rangle = 1$ (thanks to (iii), these conditions uniquely define $p(0)$). It is easy to prove that the pair $(p(t), q(t))$ is a normal singular extremal for both Dubins and dodgem car problems, and that it satisfies Assumptions 1–5; in particular $\mathbb{L}_{\widehat{\lambda}(t)} = -\mathbb{I}_n$.

Remark 6.1. We stress that even if in this problem we consider bounded controls, nevertheless we do not need to strengthen the natural optimality conditions. Indeed, the High Order Goh condition reduces to Goh condition; moreover, when considering the second variation, the linear quadratic problem (3.10)–(3.11)–(3.12) coincides with (3.14)–(3.13), since the tangent space to N_0 is in direct sum with the linear span of the controlled vector fields and their sum coincides with the Lie algebra of the controlled vector fields.

We now compute explicitly the second variation. First of all, we compute the space $\mathcal{W}$ of the admissible variations, that is we shall solve the Cauchy problem for $\zeta(t)$, (3.11)–(3.12).

Since the reference controls are null, the reference flow reduces to $\widehat{S}_t = \exp(tf_0)$. The time derivatives of the pull-back vector fields give

$$\dot{g}^i_t(\widehat{q}_0) = \exp(-tf_0)_* \Big[f_0, f_i\Big] \circ \exp(tf_0)(\widehat{q}_0) = \exp(-tf_0)_* (f_{0i}) \circ \exp(tf_0)(\widehat{q}_0)$$
$$\ddot{g}^i_t(\widehat{q}_0) = \exp(-tf_0)_* \Big[f_0, f_{0i}\Big] \circ \exp(tf_0)(\widehat{q}_0) = \exp(-tf_0)_* (f_{00i}) \circ \exp(tf_0)(\widehat{q}_0) = 0,$$

then $\dot{g}^i_t(\widehat{q}_0) = \dot{g}^i_0(\widehat{q}_0) = f_{0i}(\widehat{q}_0)$ for every $t \in [0, \widehat{T}]$ and every $i = 1, \dots, N-1$, and then

$$g^i_t(\widehat{q}_0) = f_i(\widehat{q}_0) + tf_{0i}(\widehat{q}_0).$$

The solution of (3.11)–(3.12) is then

$$\zeta(t) = \sum_{i=1}^{N(N-1)/2} \epsilon_i f_i(\widehat{q}_0) + \sum_{i=1}^{N-1} \int_0^t w_i(s)\, ds\ f_{0i}(\widehat{q}_0).$$

From the boundary condition $\zeta(\widehat{T}) = 0$ and from (v) we get that the admissible variations $(\boldsymbol{\epsilon}, \boldsymbol{w}(\cdot)) \in \mathcal{W}$ satisfy the constraints

$$\begin{cases} \int_0^{\widehat{T}} w_i(t)\ \mathrm{d}t = 0 & i = 1, \dots, N-1 \\ \epsilon_j = 0 & j = 1, \dots, \frac{N(N-1)}{2}, \end{cases} \tag{6.4}$$

then $\zeta(t) = \sum_{i=1}^{N-1} \left(\int_0^t w_i(s)\, ds\right) f_{0i}(\widehat{q}_0)$. Then the quadratic form (3.14) is given by

$$J''[(\mathbf{0}, \boldsymbol{w})]^2 = \frac{1}{2} \sum_{i=1}^{N-1} \int_0^{\widehat{T}} w_i(t)^2\ \mathrm{d}t + \sum_{i,j=1}^{N-1} \int_0^{\widehat{T}} w_i(t) \left(\int_0^t w_j(s)\ \mathrm{d}s\right) L_{f_{0j}} L_{f_{0i}} \beta(\widehat{x}_0)\ \mathrm{d}t.$$

Integrating by parts the second term and thanks to conditions (6.4) we get

$$J''[(\mathbf{0}, \boldsymbol{w})]^2 = \frac{1}{2}\|\boldsymbol{w}\|^2_{L^2} + \frac{1}{2} \sum_{i=1}^{N-1} \left(\int_0^{\widehat{T}} w_i(t)\ \mathrm{d}t\right)^2 = \frac{1}{2}\|\boldsymbol{w}\|^2_{L^2}.$$

Therefore the second variation is coercive.

Remark 6.2. In this paper we considered the second variation associated with the sub-problem with fixed final point, and we proved that its coercivity is a sufficient condition for the optimality also if the final condition is not fixed. In particular, this implies that sufficient optimality conditions for Dubins problem are also sufficient for the optimality of the extremal in dodgem car problem.

We would like to remark that, in the example considered in this section, the extended second variation associated with the original boundary conditions is not coercive. Indeed, the final constraint $g(\widehat{T}) \in N_f$ imposes the constraint $\zeta(\widehat{T}) \in \widehat{S}^{-1}_{\widehat{T}*}(T_{g(\widehat{T})}N_f)$. In particular, since N_f is an integral manifold of the derived sub-algebra $\{[f_i, f_j] : i, j = 1, \ldots, N-1\}$, then $T_{g(\widehat{T})}N_f = \{[f_i, f_j](g(\widehat{T})) : i, j = 1, \ldots, N-1\}$ and, by (vi) and the fact that the reference flow is the flow associated with the drift, it turns out that $\widehat{S}^{-1}_{\widehat{T}*}(T_{g(\widehat{T})}N_f) = \{[f_i, f_j](g(0)) : i, j = 1, \ldots, N-1\}$. It is easy to prove that any non-zero variation of the form $(\boldsymbol{\epsilon}, \boldsymbol{w} \equiv 0)$ with $\epsilon_i = 0$ for $i \leq N-1$ is admissible for the problem with final constraint N_f, but $J''[(\boldsymbol{\epsilon}, \mathbf{0})]^2 = 0$.

Therefore, $\mathcal{W}$ is the maximal subspace of variations where we can require coercivity of the extended second variation.

Appendix A. Necessity of the HOGC

This section is devoted to prove that the HOGC is a necessary optimality condition. To be more precise we prove the following result.

Theorem A.1. *Let $U = \mathbb{R}^m$ and let $(\widehat{\xi}, \widehat{\mathbf{u}}, \widehat{T})$ be an optimal triple for the problem (1.1)–(1.2). Then there exists an adjoint covector $\widehat{\lambda} : [0, \widehat{T}] \to T^*M$, such that*

$$\langle \widehat{\lambda}(t), f(\widehat{\xi}(t)) \rangle = 0 \qquad \forall\, f \in \mathrm{Lie}(\mathfrak{f}),\ t \in [0, \widehat{T}].$$

This theorem is already known when the reference control is smooth (see [8]); here we remove the smoothness hypothesis. The proof follows the outlines of the so called "higher order maximum principles" based on the "good" needle-like variations or *g-variations* (see [5], Def. 2.3)

We use the results contained in [5–7], where a general theory is considered to take into account g-variations which occur only at one point of the reference trajectory. For our purpose, we need only to consider needle-like variations coming from the values of a vector field along the reference trajectory. For this reason we summarize here only some results.

We start by fixing some notations. We denote with $S(t, t_0, q_0, \mathbf{v})$ the solution at time t of the control problem (1.2) associated with the control function $\mathbf{v}$, with initial condition at t_0 equal to q_0. Moreover, we use the notation $\widehat{S}_{t,t_0}(q_0) = \widehat{S}(t, t_0, q_0) = S(t, t_0, q_0, \widehat{\mathbf{u}})$, where $\widehat{\mathbf{u}}$ is the reference control.

In Definition A.2 here below, we put together ([5], Def. 2.1) and the properties required in ([5], Prop. 2.4) in order to single out the g-variations we need.

Definition A.2. A vector field f on M is a right g-variation of order 2 along $[0, \widehat{T}]$ of the reference couple $(\widehat{\xi}, \widehat{\mathbf{u}})$, if there exists $N > 0$ such that for every $\bar{s} \in [0, \widehat{T}]$, there exist two positive numbers c, $\bar{\epsilon}$ and a family of control maps $\{\boldsymbol{\nu}_\epsilon : \epsilon \in [0, \bar{\epsilon}]\} \subset L^1_{loc}([0, \widehat{T}], \mathbb{R}^m)$ with the following properties:

(1) $\boldsymbol{\nu}_\epsilon(t) = \widehat{\mathbf{u}}(t)$ outside the interval $[\bar{s}, \bar{s} + (N\epsilon)^2]$.
(2) There exists a neighborhood V of $\widehat{\xi}(\bar{s})$ such that the map

$$(q, \epsilon) \mapsto \widehat{S}_{\bar{s}, \bar{s}+(N\epsilon)^2} \circ S(\bar{s} + (N\epsilon)^2, \bar{s}, q, \boldsymbol{\nu}_\epsilon)$$

is continuous on $V \times [0, \bar{\epsilon}]$.
(3) The map $\epsilon \mapsto D_q\, \widehat{S}_{\bar{s}, \bar{s}+(N\epsilon)^2} \circ S(\bar{s} + (N\epsilon)^2, \bar{s}, q, \boldsymbol{\nu}_\epsilon)\big|_{q=\widehat{\xi}(\bar{s})}$ is continuous.
(4) $\widehat{S}_{\bar{s}, \bar{s}+(N\epsilon)^2} \circ S(\bar{s} + (N\epsilon)^2, \bar{s}, \widehat{\xi}(\bar{s}), \boldsymbol{\nu}_\epsilon) = \widehat{\xi}(\bar{s}) + \epsilon c f(\widehat{\xi}(\bar{s})) + o(\epsilon)$.

The core of the proof of the main result of this section is the following Lemma.

Lemma A.3. *If $U = \mathbb{R}^m$, then every $f \in \mathrm{Lie}(\mathfrak{f})$ is a right g-variation of order 2 along $[0, \widehat{T}]$ of the reference couple $(\widehat{\xi}, \widehat{\mathbf{u}})$.*

Proof. Let us consider the driftless control system

$$\dot{\zeta} = \sum_{i=1}^{m} u_i f_i \circ \zeta, \tag{A.1}$$

denoting its solutions at time t, relative to the control $\mathbf{u}$, and with initial condition $\zeta(t_0) = \zeta_0$, as $\widetilde{S}(t, t_0, \zeta_0, \mathbf{u})$.

Set $\boldsymbol{a} = \widehat{\xi}(\bar{s})$. We perform the proof in an adapted coordinate frame centered at $\boldsymbol{a}$, analogous to the frame (3.15) described in Section 3. In this frame $\boldsymbol{a} = \mathbf{0}$, $T_{\boldsymbol{a}}\mathcal{I}_{\boldsymbol{a}} = \mathbb{R}^R$ and $\mathcal{I}_{\boldsymbol{a}}$ is a neighborhood of $\mathbf{0}$ in $\mathbb{R}^R$, that we call $\mathcal{I}_{\boldsymbol{a}}$ in what follows.

If $\bar{\boldsymbol{t}} = (\bar{t}_1, \dots, \bar{t}_R)$ is sufficiently small, there exists a choice of R vector fields $\{f_{i_1}, \dots, f_{i_R}\} \in \{f_1, \dots, f_m\}$ such that the map

$$\boldsymbol{t} = (t_1, \dots, t_R) \in \mathbb{R}^R \mapsto \exp(t_R f_{i_R}) \circ \cdots \circ \exp(t_1 f_{i_1})(\boldsymbol{a}) \in \mathcal{I}_{\boldsymbol{a}} \tag{A.2}$$

has maximal rank at $\boldsymbol{t} = \bar{\boldsymbol{t}}$ (see [22], Thm. 1). This implies that there exist a $\delta > 0$ and a small neighborhood of $\boldsymbol{b} = \exp(t_R f_{i_R}) \circ \cdots \circ \exp(t_1 f_{i_1})(\boldsymbol{a})$ in $\mathcal{I}_{\boldsymbol{b}}$ such that the map (A.2) is invertible between the ball of radius δ centered in $\bar{\boldsymbol{t}}$, denoted as $B_\delta(\bar{\boldsymbol{t}})$, and the neighborhood of $\boldsymbol{b}$.

For $\boldsymbol{t} \in \mathbb{R}^R$, let us rewrite the map (A.2) as $\widetilde{S}(1, 0, \boldsymbol{a}, \mathbf{u}^{\boldsymbol{t}})$, for the piecewise-constant control $\mathbf{u}^{\boldsymbol{t}} \in L^\infty([0,1], \mathbb{R}^m)$, defined, for $j = 1, \dots, m$, by

$$u_j^{\boldsymbol{t}} \colon s \in \left[\frac{k-1}{R}, \frac{k}{R}\right] \mapsto \begin{cases} R t_k & \text{if } j = i_k \\ 0 & \text{if } j \neq i_k \end{cases} \qquad k = 1, \dots, R.$$

It is clear that $\|\mathbf{u}^{\boldsymbol{t}}\|_{L^\infty}$ is uniformly bounded for $\boldsymbol{t} \in B_\delta(\bar{\boldsymbol{t}})$. Let $\mathbf{u_0} \in L^\infty([0,1], \mathbb{R}^m)$ be the control map which satisfies

$$\widetilde{S}(1, 0, q, \mathbf{u_0})) = \exp(-\bar{t}_1 f_{i_1}) \circ \cdots \circ \exp(-\bar{t}_R f_{i_R})(q),$$

so that $\widetilde{S}(1, 0, \boldsymbol{b}, \mathbf{u_0})) = \boldsymbol{a}$ and $\|\mathbf{u_0}\|_{L^\infty} = \|\mathbf{u}^{\bar{\boldsymbol{t}}}\|_{L^\infty}$.

For $\boldsymbol{t} \in B_\delta(\bar{\boldsymbol{t}})$, let us consider the control map $\boldsymbol{\nu_t} \in L^\infty([0,2], \mathbb{R}^m)$ defined by

$$\boldsymbol{\nu_t} \colon s \mapsto \begin{cases} \mathbf{u}^{\boldsymbol{t}}(s) & \forall\, s \in [0,1] \\ \mathbf{u_0}(s-1) & \forall\, s \in (1,2]. \end{cases}$$

By definition there exists an $M > 0$ such that $\|\boldsymbol{\nu_t}\|_{L^1([0,2])} \leq M$ for every $\boldsymbol{t} \in B_\delta(\bar{\boldsymbol{t}})$; moreover, possibly restricting δ, the map $\boldsymbol{t} \in B_\delta(\bar{\boldsymbol{t}}) \mapsto \widetilde{S}(2, 0, \boldsymbol{a}, \boldsymbol{\nu_t})$ is well defined and covers a compact neighborhood U of $\boldsymbol{a}$ in $\mathcal{I}_{\boldsymbol{a}}$ contained in the local coordinate chart.

For $\epsilon > 0$, we define the control variation $\boldsymbol{\nu}_{\boldsymbol{t},\epsilon} \in L^1([0,2], \mathbb{R}^m)$ as

$$\boldsymbol{\nu}_{\boldsymbol{t},\epsilon}(s) = \begin{cases} \epsilon^{-1} \boldsymbol{\nu_t}(s\epsilon^{-2}) & \forall\, s \in [0, 2\epsilon^2] \\ 0 & \forall\, s \in (2\epsilon^2, 2], \end{cases} \tag{A.3}$$

and the control function

$$\widetilde{\boldsymbol{\nu}}_{\boldsymbol{t},\epsilon}(s) = \begin{cases} \widehat{\mathbf{u}}(\bar{s} + s) + \boldsymbol{\nu}_{\boldsymbol{t},\epsilon}(s) & \forall\, s \in [0, 2\epsilon^2] \\ \widehat{\mathbf{u}}(\bar{s} + s) & \forall\, s \in [-\bar{s}, 0) \cup (2\epsilon^2, \widehat{T} - \bar{s}]. \end{cases} \tag{A.4}$$

It is easy to see that $\|\boldsymbol{\nu}_{\boldsymbol{t},\epsilon}\|_{L^1} = \epsilon \|\boldsymbol{\nu_t}\|_{L^1} \leq M\epsilon$ and that

$$\widetilde{S}(2\epsilon^2, 0, \boldsymbol{a}, \boldsymbol{\nu}_{\boldsymbol{t},\epsilon}) = \epsilon \widetilde{S}(2, 0, \boldsymbol{a}, \boldsymbol{\nu_t}). \tag{A.5}$$

It is clear that $\widetilde{\boldsymbol{\nu}}_{\boldsymbol{t},\epsilon}$ satisfies property *(1)* of Lemma A.3. To prove the other properties, consider the pull-back system $\eta(s) = \widehat{S}^{-1}_{\bar{s}+s,\bar{s}} \circ S(\bar{s}+s,\bar{s},q,\widetilde{\boldsymbol{\nu}}_{\boldsymbol{t},\epsilon})$, which is solution of the following Cauchy problem

$$\begin{cases} \dot{\eta}(s) = \sum_{i=1}^m \nu^i_{\boldsymbol{t},\epsilon}(s)\Big(\widehat{S}^{-1}_{\bar{s}+s,\bar{s}}\Big)_* f_i \circ \widehat{S}_{\bar{s}+s,\bar{s}}(\eta(s)) \\ \eta(0) = q, \end{cases} \tag{A.6}$$

where $\nu^i_{\boldsymbol{t},\epsilon}$ denotes the i-component of $\boldsymbol{\nu}_{\boldsymbol{t},\epsilon}$. Since $\|\boldsymbol{\nu}_{\boldsymbol{t},\epsilon}\|_{L^1} \leq M\epsilon$, possibly restricting δ, there exist $\bar{\epsilon} \in (0,1)$ and a neighborhood V of $\boldsymbol{a}$ such that $\eta(2\epsilon^2)$ belongs to U, for all $q \in V$, $\boldsymbol{t} \in B_\delta(\bar{\boldsymbol{t}})$ and $\epsilon \in [0,\bar{\epsilon}]$.

Fix $\boldsymbol{t} \in B_\delta(\bar{\boldsymbol{t}})$. It is not difficult to verify that the map $\epsilon \in [0,\bar{\epsilon}] \mapsto \boldsymbol{\nu}_{\boldsymbol{t},\epsilon}$ is strongly continuous in $L^1([0,2],\mathbb{R}^m)$; therefore, (2) and (3) of Lemma A.3 are consequences of the properties of system (A.6), see [24].

Finally, to verify property (4), we consider the system (A.6) with initial condition $\eta(0) = \boldsymbol{a}$ and the system (A.1) with the same initial condition and control map $\mathbf{u} = \boldsymbol{\nu}_{\boldsymbol{t},\epsilon}$. We get

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s}|\eta(s)-\zeta(s)| &\leq \sum_{i=1}^m |\nu^i_{\boldsymbol{t},\epsilon}(s)|\Big(|(\widehat{S}^{-1}_{\bar{s}+s,\bar{s}})_* f_i \circ \widehat{S}(\bar{s}+s,\bar{s},\eta(s)) - f_i(\zeta(s))|\Big) \\ &\leq \sum_{i=1}^m |\nu^i_{\boldsymbol{t},\epsilon}(s)|\Big(|f_i(\eta(s)) - f_i(\zeta(s))| + |(\widehat{S}^{-1}_{\bar{s}+s,\bar{s}})_* f_i \circ \widehat{S}(\bar{s}+s,\bar{s},\eta(s)) - f_i(\eta(s))|\Big). \end{aligned}$$

Possibly restricting $\bar{\epsilon}$ and $\bar{\delta}$, $\eta(s)$ and $\zeta(s)$ belong to the compact neighborhood U, therefore there exists a constant $C > 0$ such that

$$\frac{\mathrm{d}}{\mathrm{d}s}|\eta(s)-\zeta(s)| \leq C\sum_{i=1}^m |\nu^i_{\boldsymbol{t},\epsilon}(s)||\eta(s)-\zeta(s)| + C\epsilon^{-1}s.$$

By Gronwall inequality we obtain

$$|\eta(2\epsilon^2) - \zeta(2\epsilon^2))| \leq \mathrm{e}^{CM\epsilon}\int_0^{2\epsilon^2} C\epsilon^{-1}s\,ds = 4C\mathrm{e}^{CM\epsilon}\epsilon^3.$$

so that

$$\eta(2\epsilon^2) = \zeta(2\epsilon^2) + o(\epsilon).$$

Since $\zeta(2\epsilon^2) = \epsilon\widetilde{S}(2,0,\boldsymbol{a},\boldsymbol{\nu}_{\boldsymbol{t}})$ and $\widetilde{S}(2,0,\boldsymbol{a},\boldsymbol{\nu}_{\boldsymbol{t}})$ covers a neighborhood of $\boldsymbol{a}$ in $\mathbb{R}^R$, Lemma A.3 is proved. □

Following ([5], Def. 2.9), we define for $t \in [0,\widehat{T}]$

$$\mathcal{K}(t) = \cup_{\tau \leq t}\Big\{(\widehat{S}_{t,\tau})_* f(\widehat{\xi}(\tau))\colon f \in \mathrm{Lie}(\mathfrak{f})\Big\} \subset T_{\widehat{\xi}(t)}M.$$

The results in ([7], Thm. 2.10) and ([5], Prop. 2.4) prove that if the convex hull of $\mathcal{K}(t)$ is the whole $T_{\widehat{\xi}(t)}M$, then $\widehat{\xi}(t)$ is in the interior of the reachable set from $\widehat{q}_0$ at time t.

Since a necessary condition for the trajectory $\widehat{\xi}$ to be time optimal is that $\widehat{\xi}(t)$ belongs to boundary of the reachable set from $\widehat{q}_0$ at each time $t \in [0,\widehat{T})$ (see for instance Sect. 3 in [7]), then it is not difficult to see that there must exist a non zero covector $\widehat{\ell}_f \in T^*_{\widehat{q}_f}M$ such that

$$\langle \widehat{\ell}_f, v\rangle = 0 \qquad \forall\, v \in \mathcal{K}(\widehat{T}).$$

Theorem A.1 follows, by definition of $\mathcal{K}(\widehat{T})$.

Appendix B. Orthonormal frame bundles on canonical space forms

In this section we give more details about the lifting of Dubins and dodgem car problem on Lie groups. For details, we refer to [20, 21] and references therein.

Let $M \in \{\mathbb{R}^n, S^n, \mathbb{H}^n\}$. We recall that the hyperbolic space $\mathbb{H}^n$ is defined as $\mathbb{H}^n = \{\boldsymbol{x} \in \mathbb{R}^{n+1} : -x_0^2 + \sum_{i=1}^n x_i^2 = 1,\ x_0 > 0\}$. The manifolds $\mathbb{R}^n$ and S^n inherit a natural Riemannian structure from $\mathbb{R}^n$ and $\mathbb{R}^{n+1}$, respectively. As for $\mathbb{H}^n$, its Riemannian metric is given by the Lorentzian quadratic form $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = -x_0 y_0 + \sum_{i=2}^n x_i y_i$.

The Dubins problem on M can be lifted to a minimum-time problem on the bundle of positive-oriented orthonormal frames on M, denoted with $\mathcal{F}_+(M)$, as we show below.

For $M = \mathbb{R}^n$, let us fix some positively oriented orthonormal frame $\{e_1, \dots, e_n\}$ attached at the point $\boldsymbol{q} = 0$ in $\mathbb{R}^n$. Given a point $\tilde{\boldsymbol{q}} \in \mathbb{R}^n$ and a positively oriented orthonormal frame $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ attached at $\tilde{\boldsymbol{q}}$, we can associate to them a pair $(\boldsymbol{x}, R) \in \mathbb{R}^n \times \mathrm{SO}(n)$, where $\boldsymbol{x}$ denotes the coordinate representation of $\tilde{\boldsymbol{q}}$ with respect to the basis $\{e_1, \dots, e_n\}$, and $\boldsymbol{v}_i = Re_i$ for every $i = 1, \dots, n$. In other words, the bundle of positively oriented orthonormal frames can be identified with the orbit through $(0, \{e_1, \dots, e_n\})$ of the semi-direct product $G = \mathbb{R}^n \rtimes \mathrm{SO}(n)$, that is the group of pairs $(\boldsymbol{x}, R) \in \mathbb{R}^n \times \mathrm{SO}(n)$ equipped with the operation $(\boldsymbol{x}, R) \cdot (\boldsymbol{y}, S) = (\boldsymbol{x} + R\boldsymbol{y}, RS)$. This construction provides a coordinate system on $\mathcal{F}_+(\mathbb{R}^n)$. Moreover, every element $(\boldsymbol{x}, R) \in G$ can be represented by the following matrix $g \in \mathrm{GL}_{n+1}(\mathbb{R})$

$$g = \begin{pmatrix} 1 & 0 \dots 0 \\ x_1 & \\ \vdots & R \\ x_n & \end{pmatrix}.$$

As the manifolds S^n and $\mathbb{H}^n$ are embedded in $\mathbb{R}^{n+1}$, we can repeat the same construction and find some group G such that all the elements of $\mathcal{F}_+(M)$ are given by the orbit of G through some fixed orthonormal frame $\{\mathfrak{e}_1, \dots, \mathfrak{e}_{n+1}\}$ of $\mathbb{R}^{n+1}$ centered at some fixed point $\boldsymbol{x}_0$.

Indeed, every point $\boldsymbol{q} \in S^n$ can be represented with respect to the canonical basis $\{\mathfrak{e}_1, \dots, \mathfrak{e}_{n+1}\}$ by a unit vector $\boldsymbol{x} \in \mathbb{R}^{n+1}$. The tangent space to S^n at $\boldsymbol{q}$ is given by the span of n unit vectors $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \in \mathbb{R}^{n+1}$ orthogonal to $\boldsymbol{x}$. A choice of these unit vectors determines an orthonormal frame on the tangent space. Therefore, the bundle $\mathcal{F}_+(S^n)$ can be regarded as the orbit of $\mathrm{SO}(n+1)$ applied to the standard orthonormal frame $\{\mathfrak{e}_1, \dots, \mathfrak{e}_{n+1}\}$ of $\mathbb{R}^{n+1}$, in the following way: to a frame $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ attached at a point $\boldsymbol{q} \in S^n$ there corresponds the matrix $g \in \mathrm{SO}(n+1)$ such that the coordinates of $\boldsymbol{q}$ are given by $\boldsymbol{x} = g\mathfrak{e}_1$ and $\boldsymbol{v}_i = g\mathfrak{e}_{i+1},\ i = 1, \dots, n$, that is

$$g = \begin{pmatrix} x^1 & v_1^1 & \dots & v_n^1 \\ \vdots & \vdots & & \\ x^{n+1} & v_1^{n+1} & \dots & v_n^{n+1} \end{pmatrix}$$

(here x^j and v_i^j denote respectively the jth component of the vectors $\boldsymbol{x}$ and $\boldsymbol{v}_i$).

For what concerns the hyperboloid $\mathbb{H}^n$, we consider the *Lorentz group* SO(1,n), defined as the group of transformation that preserve the $(n+1)$-dimensional matrix

$$\mathbb{I}(1, n) = \begin{pmatrix} -1 & 0 \dots 0 \\ 0 & \mathbb{I}_n \end{pmatrix},$$

where $\mathbb{I}_n$ is the n-dimensional identity matrix. It can be proved as above that the bundle $\mathcal{F}_+(\mathbb{H}^n)$ can be identified with the connected component $\mathrm{SO}_0(1,n)$ of SO(1,n) that contains the group identity.

For any $M \in \{\mathbb{R}^n, S^n, \mathbb{H}^n\}$, the tangent vector fields in $\mathcal{F}_+(M)$ are identified with the left-invariant vector fields on the isometry group of M, which is respectively $\mathbb{R}^n \rtimes \mathrm{SO}(n)$ for $\mathbb{R}^n$, $\mathrm{SO}(n+1)/\mathrm{SO}(n)$ for S^n and $\mathrm{SO}(1,n)$ for $\mathbb{H}^n$. Therefore, we can prove that the tangent vectors are of the form

$$g \begin{pmatrix} 0 & -\epsilon a_1 \ldots -\epsilon a_n \\ a_1 & \\ \vdots & U \\ a_n & \end{pmatrix}, \tag{B.1}$$

where $\epsilon = 0$ for $M = \mathbb{R}^n$, $\epsilon = 1$ for $M = S^n$ and $\epsilon = -1$ for $M = \mathbb{H}^n$, U is an anti-symmetric matrix of dimension n, and g is the element of $\mathcal{F}_+(M)$ where the vector is attached.

Let γ be a curve in M. When lifting γ to a curve of orthonormal frames, that is to a curve $g(t) \in G$ such that its projection on M coincides with γ, it is possible to choose the lifted curve in such a way that the first element $\boldsymbol{v}_1(t)$ of the frame attached at $\gamma(t)$ is equal to $\dot{\gamma}(t)$. In particular, this sets $a_1 = 1$ and $a_j = 0$ for $j \geq 2$ in equation (B.1). This kind of lifting is called *Darboux frame*.

There is still a freedom of choice of the form of the matrix U in equations (B.1). Systems with the form (6.3) are called *Serret-Frenet curves* (see [20, 21] and references therein).

References

[1] A.A. Agrachev and Yu. L. Sachkov, Control Theory from the Geometric Viewpoint. Springer-Verlag (2004).

[2] A. Agrachev, G. Stefani and P. Zezza, An invariant second variation in optimal control. *Int. J. Control* **71** (1998) 689–715.

[3] M.S. Aronna, J.F. Bonnans, A.V. Dmitruk and P. Lotito, Quadratic order conditions for bang-singular extremals. *AIMS J. Numer. Algebra Control Optim.* **2** (2012) 511–546.

[4] A. Bacciotti and G. Stefani, On the relationship between global and local controllability. *Math. Syst. Theory* **16** (1983) 79–91.

[5] R.M. Bianchini, Good needle-like variations. In vol. 64 of *Proceedings of Symposia in Pure Mathematics* (1999).

[6] R.M. Bianchini, *Variational cones and high-order maximum principles*. Technical report, Dipartimento di Matematica "Ulisse Dini", viale Morgagni 67/a, Firenze (1994).

[7] R.M. Bianchini, Variational Approach to Some Optimization Control Problems. In Geometry in Nonlinear Control and Differential Inclusions. Edited by G. Ferreyra R. Gardner H. Hermes and H. Sussmann (1995).

[8] R.M. Bianchini and G. Stefani, A High Order Maximum Principle, in Analysis and Control of Linear Systems. Edited by R.E. Saeks C.I. Byrnes, C.F. Martin (1988).

[9] A. Bressan and F. Rampazzo, Impulsive control systems with commutative vector fields. *J. Optim. Theory Appl.* **71** (1991) 67–84.

[10] A. Bressan and F Rampazzo, Impulsive control systems without commutativity assumptions. *J. Optim. Theory Appl.* **81** (1994) 435–457.

[11] F.C. Chittaro and G. Stefani, Singular extremals in multi-input time-optimal problem: a sufficient condition. *Control Cyber.* **39** (2010).

[12] F.C. Chittaro and G. Stefani, Minimum-Time Strong Optimality of a Singular Arc: extended Dubins Problem. In *52nd IEEE Conference on Decision and Control* (2013).

[13] B.D. Craven, Control and Optimization. *Chapman & Hall Mathematics Series*. Chapman & Hall, London, New York (1995).

[14] A. V. Dmitruk. Quadratic condition for a weak minimum for singular regimes in optimal control problems. *Soviet Math Dokl.* **18** (1977).

[15] A. V. Dmitruk, Jacobi type conditions for singular extremals. *Control Cybernet.* **37** (2008) 285–306.

[16] R. Gabasov and F.M. Kirillova, High order necessary conditions for optimality. *SIAM J. Control* **10** (1972) 127–168.

[17] B. S. Goh, The second variation for singular Bolza problems. *SIAM J. Control Optim.* **4** (1966) 309–325.

[18] M. Guerra and A. Sarychev, Fréchet generalized trajectories and minimizers for variational problems of low coercivity. *J. Dyn. Contr. Syst.* **21** (2015) 351–377.

[19] M.R. Hestenes, Application of the theory of quadratic forms in Hilbert spaces to the calculus of variations. *Pac. J. Math.* **1** (1951) 525–581.

[20] V. Jurdjevic, Geometric Control Theory. Cambridge University Press (1997).

[21] V. Jurdjevic and F. Monroy-Pérez, Variational Problems on Lie Groups and Their Homogeneous Spaces: Elastic Curves, Tops, and Constrained Geodesic Problems (2002).

[22] A.J. Krener, A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems. *SIAM J. Control Optim.* **12** (1974) 43–52.

[23] J.E. Marsden and T.S. Ratiu, Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems. *Texts Appl. Math.* Springer (1999).

[24] E.J. McShane, Unified Integration. Academic Press (1983).

[25] L. Poggiolini and G. Stefani, Bang-singular-bang extremals: sufficient optimality conditions. *J. Dyn. Control Syst.* **17** (2011) 469–514.

[26] G. Stefani, Minimum-time optimality of a singular arc: second order sufficient conditions. In vol. 1 of *43rd IEEE Conference on Decision and Control* (2004).

[27] G. Stefani, Strong Optimality of Singular Trajectories. Geometric Control and nonsmooth analysis. Edited by F. Ancona A. Bressan P. Cannarsa F. Clarke and P. R. Wolenski (2008).

[28] G. Stefani and P. Zezza. Constrained regular LQ-control problems. *SIAM J. Control Optim.* **35** (1997) 876–900.