SINGULAR EXTREMALS IN L^1-OPTIMAL CONTROL PROBLEMS: SUFFICIENT OPTIMALITY CONDITIONS*

Francesca C. Chittaro[1,**] and Laura Poggiolini[2]

Abstract. In this paper we are concerned with generalised L^1-minimisation problems, *i.e.* Bolza problems involving the absolute value of the control with a control-affine dynamics. We establish sufficient conditions for the strong local optimality of extremals given by the concatenation of bang, singular and inactive (zero) arcs. The sufficiency of such conditions is proved by means of Hamiltonian methods. As a by-product of the result, we provide an explicit invariant formula for the second variation along the singular arc.

1. INTRODUCTION

L^1-minimisation problems, that is, optimal control problems aiming to minimise the L^1 norm of the control, have shown to model quite effectively fuel consumption optimisation problems in engineering [7, 8, 26, 27] and some systems appearing in neurobiology [5]. In a recent paper dealing with the problem of minimising the fuel consumption of an academic vehicle [6], a generalised version of this problem is studied: the cost to be minimised is the *absolute work*, modelled as the integral of the absolute value of the control, weighted by the absolute value of a function dependent on the state (such kind of problems have been called *generalised L^1 optimal control problems* in [10]).

Besides the well known solutions given by concatenations of singular and bang arcs, generalised L^1-optimal control problems are known to present a new category of extremal arcs, in which the control is neither singular nor bang, that is, its value is uniquely determined by Pontryagin Maximum Principle, but is not an extremum point of the control set. In these arcs the control is identically zero, so they are commonly called *zero arcs, zero thrust arcs, inactivated arcs* or *cost arcs*. The property of L^1 minimisation of generating zero arcs is well known and exploited in practical situations, see for instance [18].

The Pontryagin Maximum Principle, suitably generalised for non-smooth frameworks (see *e.g.* [14]), provides a set of necessary conditions for the optimality of admissible trajectories. The aim of this paper is to give a

---

[1] Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France.
[2] DiMaI, Università di Firenze, 50139 Firenze, Italy.

** Corresponding author: francesca-carlotta.chittaro@univ-tln.fr

set of sufficient conditions for the optimality of admissible trajectories that satisfy the Pontryagin Maximum Principle and whose associated extremal contains bang, singular and inactivated arcs.

Before proceeding with further discussion, let us state the problem in detail. Let M be a smooth n-dimensional manifold, $\psi \colon M \to \mathbb{R}$ a smooth function, f_0, f_1 two smooth vector fields on M. Finally let $T > 0$ be fixed and let q_0, q_T be two given points in M. Consider a Bolza optimal control problem of the following form:

$$\text{minimise} \int_0^T |u(t)\psi(\xi(t))| \, \mathrm{d}t \tag{1.1}$$

over all absolutely continuous trajectories that are Carathéodory solutions of the boundary value problem (admissible trajectories)

$$\begin{cases} \dot{\xi}(t) = (f_0 + u(t)f_1) \circ \xi(t), \\ \xi(0) = q_0, \quad \xi(T) = q_T, \\ u \in L^\infty([0,T],[-1,1]). \end{cases} \tag{1.2}$$

For a solution $\widehat{\xi}$ of the Cauchy problem (1.2), according to [3], we adopt the following notion of optimality.

Definition 1.1 (Strong local optimality). The trajectory $\widehat{\xi}$ is a strong local minimiser of problem (1.1)–(1.2) if there exists a neighbourhood $\mathbf{O}$ of its graph in $\mathbb{R} \times M$ such that $\widehat{\xi}$ is a minimiser among the admissible trajectories whose graphs are contained in $\mathbf{O}$, independently of the values of the associated control function. We say that $\widehat{\xi}$ is a strict strong local minimiser if it is the only minimising trajectory whose graph is in $\mathbf{O}$.

Pontryagin Maximum Principle states that, if $\widehat{\xi}$ is locally optimal, then it must be an *extremal trajectory*, *i.e.* the projection on the manifold M of the solution of a suitable Hamiltonian system defined on the cotangent bundle T^*M, see *e.g.* [1, 14]. If we limit ourselves to consider normal extremals only, then such Hamiltonian system is determined by the values of the two Hamiltonian functions $\Phi^\pm(\ell) = \langle \ell, f_1(\pi\ell) \rangle \pm |\psi(\pi\ell)|$, $\ell \in T^*M$, whose role is analogous to the one played by the switching functions in smooth control-affine optimal control problems. In particular, when neither of the two functions vanishes, the maximised Hamiltonian is realised by one and only one admissible control value, which, according to the sign of these two functions, is either ± 1 or 0. In the first case, in analogy with the control-affine case, we say that we have a *regular bang control*, while in the latter case we say that we have a *zero* or *inactivated control*; on the other hand, if either $\Phi^+(\ell)$ or $\Phi^-(\ell)$ vanishes along a nontrivial arc of a Pontryagin extremal, then only the sign of the control is prescribed, and we say that we have a *singular control*.

In [6] the authors consider a specific parameter-dependent problem fitting in the class of problem (1.1)–(1.2). They show that, according to the values of the parameters, the optimal extremal trajectories are given by the concatenations of bang-zero-bang arcs or of bang-singular-zero-bang arcs. Inspired by this result, we look for sufficient optimality conditions for such extremals: in [9, 10], we focus on extremals made by a concatenation of bang-zero-bang arcs; here we consider the case of a concatenation of bang-singular-zero-bang arcs and provide an adequate set of sufficient conditions.

Our approach, successfully applied also for other classes of problems (see for instance [2, 4, 12, 20, 23, 25, 30]) relies on Hamiltonian methods. The main steps in these methods are the following:

- if the flow generated by the maximised Hamiltonian is well defined, and the maximised Hamiltonian is at least C^2, find a Lagrangian submanifold Λ_1 of the cotangent bundle that projects injectively onto the base manifold, and such that at each time $t \in [0,T]$ its image under the flow generated by the maximised Hamiltonian projects one-to-one onto a neighbourhood of $\widehat{\xi}(t)$ in the base manifold;
- thanks to the local invertibility of this flow, lift all admissible trajectories, with graph belonging to the neighbourhood of the graph of the reference one, to Λ_1;
- estimate the cost associated with every trajectory, by means of a line integral in Λ_1.

As long as this construction is possible – that is, as long as the image of the manifold via the flow generated by the maximised Hamiltonian is invertible – it is possible to show that the reference trajectory is indeed a local minimiser.

In the series of papers [4, 11, 20, 23, 28], the existence of a suitable manifold Λ_1 is shown to be related to the coerciveness of the second variation associated with some sub-problem of the original one. Indeed, the second variation is written as an accessory linear-quadratic control problem on the tangent space; by the classical theory of linear-quadratic systems (see *e.g.* [29], Thm. 2.6), if the second variation is coercive, then the image, under the linearised flow, of the space of transversality conditions of the accessory problem projects injectively onto the base manifold. A good candidate for the manifold Λ_1 is thus given by the image of such manifold of transversality conditions, under a suitably defined symplectic (or anti-symplectic) isomorphism.

We stress that one of the main strengths of this approach relies on the fact that all trajectories whose graph is close to the reference one can be lifted to Λ_1, regardless of the value of the associated control, thus yielding optimality in the strong topology.

In the case of bang-bang extremals, the maximised Hamiltonian is not C^2. Nevertheless, the Hamiltonian methods described above can be applied with minor adjustments, see *e.g.* [4, 19–22]. When the extremal is singular or is the concatenation of bang and singular arcs, then a more sophisticated construction is required. In particular, it is no longer possible to use the flow associated with the maximised Hamiltonian of the control system (see Sect. 5.1 for details). However, as observed for the first time in [28], Hamiltonian methods work even if the maximised Hamiltonian is replaced by a suitable *over-maximised Hamiltonian*. In the present paper, this construction is made possible thanks to a set of *regularity assumptions* (Asm. 3.1–3.9), holding along the reference trajectory.

The problem under study presents another tough issue: indeed, we are dealing with a Bolza problem containing a singular arc. As it happens, for instance, for singular extremals of the minimum-time problem or of a Mayer problem, the second variation is degenerate, thus not coercive. This problem can be surmounted by means of a Goh transformation, which provides a non-degenerate second variation, defined on a larger Hilbert space. Moreover, since we are dealing with a Bolza problem, the construction of such second variation is particularly elaborate and, up to the authors' knowledge, this is the first time it is computed, at least in the invariant form we are using.

The paper is organised as follows: in Section 2 we state the notations that we are going to use throughout the paper; in Section 3 we state the first part of our assumptions. In Section 4 we define the extended second variation. In Section 5, by using the regularity assumptions, we construct the over-maximised flow and in Section 6 we prove that the projected over-maximised flow emanating from an appropriate Lagrangian manifold is locally invertible. Finally, in Section 7 we state and prove the main result of the paper. The result is illustrated in Section 8 with an example for which explicit analytical computations can be done. For the sake of readability, some technical computations are postponed to the Appendices.

2. NOTATIONS

We denote with TM and with T^*M the tangent bundle and the cotangent bundle to M, respectively. π denotes the canonical projection of T^*M on M; the elements of T^*M are denoted with ℓ.

In the following, small letters f, g, k denote vector fields on the manifold M, and the corresponding capital letters are used to denote the corresponding Hamiltonian lift, *i.e.* $F(\ell) = \langle \ell, f(\pi\ell) \rangle$. Given a vector field f on M, the Lie derivative at a point $q \in M$ of a smooth function $\varphi \colon M \to \mathbb{R}$ with respect to f is denoted with $L_f\varphi(q) = \langle d\varphi(q), f(q) \rangle$, and $L_f^2\varphi(q) = L_f(L_f\varphi)(q)$. The Lie bracket of two vector fields f, g is denoted as commonly with $[f, g]$. In particular, for Lie brackets of indexed vector fields as f_0, f_1, we adopt the following notations:

$$f_{ij} = [f_i, f_j], \qquad f_{ijk} = [f_i, f_{jk}].$$

Analogously, $F_{ij}(\ell) = \langle \ell, f_{ij}(\pi\ell) \rangle$ and $F_{ijk}(\ell) = \langle \ell, f_{ijk}(\pi\ell) \rangle$.

The symbol ς denotes the Poincaré-Cartan invariant on T^*M, defined as $\varsigma_\ell = \ell \circ \pi_*$ $\forall \ell \in T^*M$. The symbol $\boldsymbol{\sigma}_\ell = \mathrm{d}\varsigma_\ell$ denotes the canonical symplectic form on T^*M. With each Hamiltonian function F we associate the Hamiltonian vector field $\vec{F}$ on T^*M defined by

$$\langle \mathrm{d}F(\ell) , \cdot \rangle = \boldsymbol{\sigma}_\ell(\cdot, \vec{F}(\ell)).$$

In this paper, a special role is played by the switching time $\widehat{\tau}_1$ between the first bang arc and the singular one. Thus, we shall always consider flows starting from time $\widehat{\tau}_1$, evolving backwards in time up to time $t = 0$ or forward in time up to time $t = T$. Capital cursive letters are used for the Hamiltonian flows associated with some Hamiltonian vector fields: for instance, $\mathcal{F}_t$ denotes the flow, from time $\widehat{\tau}_1$ to time t, associated with the Hamiltonian vector field $\vec{F}$.

Throughout the paper, the superscript $\widehat{\cdot}$ is used for objects related to the reference trajectory $\widehat{\xi}$. In particular, $\widehat{u}(t)$ denotes the control associated with $\widehat{\xi}$, $\widehat{f}_t$ the vector field $f_0 + \widehat{u}(t)f_1$ and $\widehat{S}_t$ its flow, *i.e.* $\widehat{S}_t(q)$ is the solution at time t of the Cauchy problem

$$\begin{cases} \dot{\xi}(t) = \widehat{f}_t \circ \xi(t), \\ \xi(\widehat{\tau}_1) = q. \end{cases}$$

Analogously, $\widehat{F}_t(\ell) = \langle \ell , \widehat{f}_t(\pi\ell) \rangle$ is the Hamiltonian associated with $\widehat{f}_t$, and $\widehat{\mathcal{F}}_t$ denotes its associated Hamiltonian flow. Finally, we define the function $\widehat{\psi}_t \colon M \to \mathbb{R}$ as $\widehat{\psi}_t = \psi \circ \widehat{S}_t$.

3. Regularity assumptions

We consider an admissible trajectory $\widehat{\xi}$ of the control system (1.2) whose associated control $\widehat{u}$ has the following structure:

$$\widehat{u}(t) = \begin{cases} 1 & t \in (0, \widehat{\tau}_1), \\ \widehat{u}_S(t) \in (0,1) & t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ 0 & t \in (\widehat{\tau}_2, \widehat{\tau}_3), \\ -1 & t \in (\widehat{\tau}_3, T). \end{cases} \tag{3.1}$$

where $0 < \widehat{\tau}_1 < \widehat{\tau}_2 < \widehat{\tau}_3 < T$. $\widehat{\xi}$ is called the *reference trajectory*. The times $\widehat{\tau}_i$, $i = 1, 2, 3$ are called *(reference) switching times* and, analogously, the points $\widehat{q}_i = \widehat{\xi}(\widehat{\tau}_i)$ are called *(reference) switching points*.

This section is devoted to the statement of the necessary conditions for optimality and the discussion of the regularity assumptions along the reference trajectory.

Assumption 3.1. Along the reference trajectory, the cost ψ does not vanish for $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$ and $t = \widehat{\tau}_3$. Without loss of generality, we assume that $\psi \circ \widehat{\xi}|_{[\widehat{\tau}_1, \widehat{\tau}_2]} > 0$.

Assumption 3.2. For every $t \in [0, T]$ such that $\psi(\widehat{\xi}(t)) = 0$, it holds $L_{\widehat{\dot{\xi}}(t)} \psi(\widehat{\xi}(t)) \neq 0$.

Assumptions 3.1–3.2 deal with the behaviour of the function ψ; in particular, a direct consequence of Assumption 3.2 is that the zeroes of the cost function ψ along $\widehat{\xi}$ are isolated, thus finite. Since the zeroes of ψ are the only non-smoothness points of the running cost, we can apply a classical version of PMP. Indeed, due to non-smoothness, more general versions of PMP (see, *e.g.*, [14], Thm. 22.2, [16]) would be required. Nevertheless, thanks to Assumption 3.2, we can rearrange the optimal control problem into a hybrid control problem, as defined in ([14], Sect. 22.5), setting the surface $S = \{(t, x, y) \colon \psi(x) = 0, y = x\}$ as switching surface. The analogue of the classical PMP for hybrid optimal control problems is the Hybrid Maximum Principle (see [31],

Thm. 17.4.1 and [14], Thm. 22.26) which, under Assumption 3.2, reduces to the standard smooth version of PMP, see [1]. For this reason, here below we refer to the classical notions of Pontryagin extremal and PMP.

Definition 3.3 (Pontryagin extremals)**.** For every $u \in [-1, 1]$ and for $p_0 \in \{0, 1\}$, we consider the pre-Hamiltonian function associated with the optimal control problem (1.1)–(1.2)

$$h(\ell, u) := F_0(\ell) + uF_1(\ell) - p_0 |u\,\psi(\pi\ell)|,$$

and we define the maximised Hamiltonian as

$$H_{\max}(\ell) = \max_{u \in [-1, 1]} h(\ell, u).$$

Let $\lambda \colon [0, T] \to T^*M$ be an absolutely continuous curve such that $\psi(\pi\lambda(t))$ vanishes only for a finite number of times $0 < t_1 < \ldots < t_k < T$. λ is called a Pontryagin extremal of problem (1.1)–(1.2), if there exist a constant $p_0 \in \{0, 1\}$ and an admissible control $u(t)$ (called *extremal control*) such that

$$
\begin{aligned}
(\lambda(t), p_0) &\neq 0, & \forall\, t \in [0, T], & \\
\dot\lambda(t) &= \vec{h}(\lambda(t), u(t)), & \text{a.e. } t \in [0, T], & \quad (3.2a) \\
h(\lambda(t), u(t)) &= H_{\max}(\lambda(t)), & \text{a.e. } t \in [0, T], & \quad (3.2b) \\
\pi\lambda(0) &= q_0, \quad \pi\lambda(T) = q_T. &
\end{aligned}
$$

If $p_0 = 1$, then λ is called a *normal Pontryagin extremal*, if $p_0 = 0$ we say that λ is an *abnormal Pontryagin extremal*.

As discussed in the Introduction, in the case of normal extremals, the Hamiltonian functions

$$\Phi^{\pm}(\ell) := F_1(\ell) \pm |\psi(\pi\ell)|$$

play the same role as the switching functions do in control-affine problems. Indeed, if both Φ^- and Φ^+ are non-zero along an extremal, then the extremal control is uniquely determined by equation (3.2b): in particular, it is zero if Φ^- and Φ^+ have different signs, it is $+1$ if both are positive, and it is -1 if both are negative. If only one between Φ^- and Φ^+ is zero, then PMP prescribes only the sign of the extremal control. The last case, where both Φ^- and Φ^+ are zero, is highly degenerate and in this case equation (3.2b) gives no information about the value of the extremal control. For these reasons, Φ^- and Φ^+ are called *switching functions*. Subarcs of a Pontryagin extremal are thus classified according to the signs of the switching functions.

Definition 3.4. Let $\lambda \colon [0, T] \to T^*M$ be a normal Pontryagin extremal for problem (1.1)-(1.2), and let $I \subset [0, T]$ be an open interval.
  If $\Phi^-(\lambda(t))\Phi^+(\lambda(t)) > 0$ for every $t \in I$, then $\lambda|_I$ is called a *regular bang arc*.
  If $\Phi^-(\lambda(t))\Phi^+(\lambda(t)) < 0$ for every $t \in I$, then $\lambda|_I$ is called an *inactivated* or *zero arc*.
  If $\Phi^-(\lambda(t))\Phi^+(\lambda(t)) = 0$ for every $t \in I$, then $\lambda|_I$ is called a *singular arc*.
  In particular, if one between $\Phi^-(\lambda(t))$ and $\Phi^+(\lambda(t))$ is different from zero for every $t \in I$, then $\lambda|_I$ is a *non-degenerate singular arc*. Else we call it *degenerate*.

Assumption 3.5. There exists a normal Pontryagin extremal $\widehat{\lambda}$ associated with the reference control $\widehat{u}$ such that $\pi\widehat{\lambda}(t) = \widehat{\xi}(t)$ for every $t \in [0, T]$. We assume that $\widehat{\lambda}|_{[0,\widehat{\tau}_1)}$ and $\widehat{\lambda}|_{(\widehat{\tau}_3, T]}$ are regular bang arcs, that $\widehat{\lambda}|_{(\widehat{\tau}_1, \widehat{\tau}_2)}$ is a non-degenerate[1] singular arc, and that $\widehat{\lambda}|_{(\widehat{\tau}_2, \widehat{\tau}_3)}$ is an inactivated arc.

---

[1] We point out that degenerate singular arcs occur if and only if for some $t \in (\widehat{\tau}_1, \widehat{\tau}_2)$ $F_1(\lambda(t)) = \psi(\pi\lambda(t)) = 0$; for the extremal $\widehat{\lambda}$ this situation is precluded by Assumption 3.1.

Notation. *We set $\widehat{\ell}_i = \widehat{\lambda}(\widehat{\tau}_i)$, for $i = 0, 1, 2, 3$, and $\widehat{\ell}_T = \widehat{\lambda}(T)$.*
   Moreover, we define the following constants:

$$a_i = sign\big(\psi(\widehat{q}_i)\big), \ \ i = 1, 2, 3.$$

Thanks to Assumption 3.1, it follows that $a_1 = a_2 = 1$.

Remark 3.6. We recall that, if the reference extremal is optimal, then it must satisfy PMP and the switching functions must satisfy the mild version of the inequalities appearing in Definition 3.4; the only additional requirements in Assumption 3.5 are the regularity of the arcs and the fact that the extremal is normal.

Remark 3.7. Assumption 3.1 ensures that, in a neighbourhood of $\widehat{\lambda}([0, T])$, the switching surfaces $\{\Phi^+ = 0\}$ and $\{\Phi^- = 0\}$ do not intersect each other. Together with the fact that the zeroes of ψ along $\widehat{\xi}$ are finite, this fact guarantees that the Hamiltonian vector field associated with $\widehat{F}_t$ is well defined along the reference trajectory, but for at most a finite number of times, that is, the switching times and the zeroes of $\psi(\pi\widehat{\lambda}(t))$.

Assumption 3.5 and equation (3.1) imply the following conditions on the sign of the switching functions along the reference extremal:

$$\Phi^-\big(\widehat{\lambda}(t)\big) > 0 \qquad\qquad t \in [0, \widehat{\tau}_1), \qquad\qquad (3.3)$$

$$\Phi^-\big(\widehat{\lambda}(t)\big) = 0 \qquad\qquad t \in [\widehat{\tau}_1, \widehat{\tau}_2], \qquad\qquad (3.4)$$

$$\Phi^-\big(\widehat{\lambda}(t)\big) < 0 < \Phi^+\big(\widehat{\lambda}(t)\big) \qquad t \in (\widehat{\tau}_2, \widehat{\tau}_3), \qquad\qquad (3.5)$$

$$\Phi^+\big(\widehat{\lambda}(t)\big) < 0 \qquad\qquad t \in (\widehat{\tau}_3, T]. \qquad\qquad (3.6)$$

Equations (3.3)–(3.6) yield a set of *higher order necessary conditions*. Indeed, combining equations (3.3) and (3.4), we obtain that $\frac{\mathrm{d}}{\mathrm{d}t}\Phi^-\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_1^-} \leq 0$, while (3.4) gives $\frac{\mathrm{d}}{\mathrm{d}t}\Phi^-\big(\widehat{\lambda}(t)\big) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi^-\big(\widehat{\lambda}(t)\big) = 0$ for every $t \in (\widehat{\tau}_1, \widehat{\tau}_2)$. Explicit computations show that $\frac{\mathrm{d}}{\mathrm{d}t}\Phi^-\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_1^\pm} = \{F_0, \Phi^-\}(\widehat{\ell}_1)$. By continuity this implies that $\frac{\mathrm{d}}{\mathrm{d}t}\Phi^-\big(\widehat{\lambda}(t)\big)\Big|_{t=\widehat{\tau}_1} = 0$, so that we must have $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi^-\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_1^-} \geq 0$.

Analogously, from equations (3.4) and (3.5), we obtain that $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi^-\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_2^+} \leq 0$.

At time $\widehat{\tau}_3$, $\Phi^+\big(\widehat{\lambda}(t)\big)$ is differentiable and changes sign from positive to negative, thus $\frac{\mathrm{d}}{\mathrm{d}t}\Phi^+\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_3} \leq 0$. The regularity assumptions at the switching points consist in a strengthening of the above inequalities.

Assumption 3.8 (Regularity at the switching points.)**.**

$$(F_{001} + F_{101})\,(\widehat{\ell}_1) + a_1 L_{f_{01}}\psi(\widehat{q}_1) - a_1 L_{f_0+f_1}L_{f_0}\psi(\widehat{q}_1) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi^-\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_1^-} > 0$$

$$F_{001}(\widehat{\ell}_2) - a_2 L_{f_0}^2 \psi(\widehat{q}_2) = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi^-\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_2^+} < 0$$

$$\mathfrak{r}_3 = F_{01}(\widehat{\ell}_3) + a_3 L_{f_0}\psi(\widehat{q}_3) = \frac{\mathrm{d}}{\mathrm{d}t}\Phi^+\big(\widehat{\lambda}(t)\big)|_{t=\widehat{\tau}_3} < 0.$$

A well-known second order necessary optimality condition concerning singular arcs is given by the Generalised Legendre condition (see for instance [1], Thm. 20.16), which in our context reduces to

$$F_{101}\big(\widehat{\lambda}(t)\big) + a_1\big(L_{f_{01}}\psi\big(\widehat{\xi}(t)\big) - L_{f_1}L_{f_0}\psi\big(\widehat{\xi}(t)\big)\big) \geq 0.$$

We assume that the inequality here above holds in the strict form.

Assumption 3.9 (Strong generalised Legendre Condition (SGLC)). For all $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$

$$F_{101}(\widehat{\lambda}(t)) + a_1 \big(L_{f_{01}} \psi \big(\widehat{\xi}(t) \big) - L_{f_1} L_{f_0} \psi \big(\widehat{\xi}(t) \big) \big) > 0. \tag{3.7}$$

For the purpose of future computations, we introduce the following notation:

$$\mathbb{L}(\ell) = F_{101}(\ell) + a_1 \big(L_{f_{01}} \psi \big(\pi \ell \big) - L_{f_1} L_{f_0} \psi \big(\pi \ell \big) \big) \qquad \ell \in T^*M,$$

so that equation (3.7) reads $\mathbb{L}(\widehat{\lambda}(t)) > 0, \ t \in [\widehat{\tau}_1, \widehat{\tau}_2]$.

Assumption 3.9 yields some geometric properties of two subsets of T^*M which are crucial for our construction:

$$\Sigma^- = \{ \ell \in T^*M \colon \Phi^-(\ell) = 0 \},$$
$$S^- = \{ \ell \in \Sigma^- \colon F_{01}(\ell) - L_{f_0} \psi(\pi \ell) = 0 \}.$$

Indeed, thanks to Assumption 3.9, it is easy to see that, in a neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$, Σ^- is a codimension one embedded submanifold of T^*M. Moreover, Assumption 3.9 implies that $\vec{\Phi}^-$ is not tangent to S^-, so that S^- is a codimension one embedded submanifold of Σ^-. We finally notice that, for every $\ell \in S^-$, the tangent space to Σ^- at ℓ splits in the following direct sum

$$T_\ell \Sigma^- = T_\ell S^- \oplus \mathbb{R} \vec{\Phi}^-(\ell). \tag{3.8}$$

We end the section with a technical assumption concerning again the zeroes of the running cost along the reference trajectory. Indeed, as already stressed, the flow generated by the maximised Hamiltonian, appearing in (3.2a)–(3.2b), depends on how many times the function ψ changes sign along the reference trajectory, as each of these points bears a non-smoothness point. This issue has been accurately treated in [10], and the same computations carried out there could extend with no modifications to the current problem. Thus, due to the complexity introduced by the presence of a singular arc, and in order to simplify the presentation, we make the following assumption.

Assumption 3.10. Along the first bang-arc, the singular and the last bang arc, the function ψ is positive.

Remark 3.11. We stress that this assumption does not cause any loss of generality; the result (Thm. 7.1) holds true also if we drop it, provided that the other assumptions are satisfied and that the second variation and the maximised flow are suitably computed, according to the rules given in [10].

4. THE SECOND VARIATION

Following the approach initiated in [2], we write the second variation as an accessory problem, that is, an LQ optimal control problem defined on the tangent space to M at $\widehat{q}_1$.[2] The admissible control functions of the accessory problem are the admissible control variations of the original optimal control problem, that is, all functions $\delta u \in L^\infty([0, T], \mathbb{R})$ such that $\widehat{u} + \delta u$ is still an admissible control for problem (1.2); the set of admissible control variations is then completed as a suitable subspace of the Hilbert space $L^2([0, T], \mathbb{R})$ (see [2], Rem. 6). However, facing the problem from the most general point of view, that is, considering all possible admissible variations, is not only cumbersome, but possibly pointless: in many general cases, indeed, the space of admissible variations is "too big", so that the second variation cannot be coercive on it (see for instance [24] for bang-singular concatenations and [12] for an example in the case of totally singular extremals). On the other hand, in many cases ([4, 12, 23]) it has been proved that it is possible to properly reduce the set

---

[2] In principle, the basepoint could be any point of the reference trajectory. The choice of $\widehat{q}_1$ considerably simplifies the expression of the second variation, as it permits to neglect variations along the first bang arc.

of admissible variations, and still obtain sufficient conditions for optimality in terms of the coerciveness of the second variation. The goal is then to find the "smallest" space of admissible variations such that the coerciveness of the second variation on it still implies that the Hamiltonian flow is invertible.

In particular, in [4] it has been shown that, for bang-bang extremals, it is sufficient to consider only the variations of the switching times. As it will be proved in the paper, it turns out that, for problem (1.1)–(1.2), an appropriate set of admissible variations is constituted by the variation of the third switching time and the variation of the control function along the singular arc.

Clearly, this reduction considerably simplifies the expression of the second variation. The rest of this section is devoted to its construction.

It is well known that computing higher order derivatives on manifolds is a delicate task, as they are not invariant under change of coordinates. To overcome this problem and obtain an intrinsic expression of the second variation, we transform the original problem into a Mayer one and we pull-back the linearisation of the system along the reference trajectory to the tangent space $T_{\widehat{q}_1}M$. Setting $\boldsymbol{\xi} = (\xi^0, \xi) \in \mathbb{R} \times M$, problem (1.1)–(1.2) reads

$$\text{minimise } \xi^0(T) - \xi^0(0)$$

among all solutions of the control system

$$\begin{cases} \dot{\xi}^0(t) = |u(t)\psi(\xi(t))|, \\ \dot{\xi}(t) = (f_0 + uf_1) \circ \xi(t), \\ \boldsymbol{\xi}(0) = (0, \widehat{q}_0), \quad \boldsymbol{\xi}(T) \in \mathbb{R} \times \{\widehat{q}_T\}, \\ u \in [-1, 1]. \end{cases} \tag{4.1}$$

It is immediate to see that the covector $\boldsymbol{\lambda} = (\lambda^0, \widehat{\lambda}) \in \mathbb{R} \times T^*M$, with $\lambda^0(t) \equiv -1$ satisfies normal PMP.

The reference flow from time $\widehat{\tau}_1$ associated with the system (4.1) is denoted as $\widehat{\boldsymbol{S}}_{\boldsymbol{t}}$, and is given by

$$\widehat{\boldsymbol{S}}_{\boldsymbol{t}}(c_1, q) = \begin{pmatrix} \widehat{S}_t^0(c_1, q) \\ \widehat{S}_t(q) \end{pmatrix} = \begin{pmatrix} c_1 + \int_{\widehat{\tau}_1}^t |\widehat{u}(s)\psi(\widehat{S}_s(q))|\mathrm{d}s \\ \widehat{S}_t(q) \end{pmatrix} = \begin{pmatrix} c_1 + \int_{\widehat{\tau}_1}^t |\widehat{u}(s)\widehat{\psi}_s(q))|\mathrm{d}s \\ \widehat{S}_t(q) \end{pmatrix}.$$

Remark 4.1. Notice that $\widehat{S}_t^0$ does depend on c_1, while its differential does not. In what follows, with some abuse of notation, we write $\mathrm{d}\widehat{S}_t^0(q)$ for the differential of $q \mapsto \widehat{S}_t^0(c_1, \cdot)$ at a point q.

Consider some $\tau_3 \in (\widehat{\tau}_2, T)$ and a measurable control function $v \colon [\widehat{\tau}_1, \widehat{\tau}_2] \to (0, 1)$, and let $\boldsymbol{\xi}$ be the solution of (4.1), starting from the point $\boldsymbol{\xi}(0) = (0, \widehat{q}_0)$ associated with the control

$$u(t) = \begin{cases} 1 & t \in [0, \widehat{\tau}_1), \\ v(t) & t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ 0 & t \in (\widehat{\tau}_2, \tau_3), \\ -1 & t \in (\tau_3, T]. \end{cases}$$

We consider the piecewise-affine reparametrization of time $\varphi \colon [0, T] \to [0, T]$ defined by

$$\dot{\varphi}(t) = \begin{cases} 1 & t \in [0, \widehat{\tau}_2), \\ \frac{\tau_3 - \widehat{\tau}_2}{\widehat{\tau}_3 - \widehat{\tau}_2} & t \in (\widehat{\tau}_2, \widehat{\tau}_3), \\ \frac{T - \tau_3}{T - \widehat{\tau}_3} & t \in (\widehat{\tau}_3, T], \end{cases} \qquad \varphi(0) = 0,$$

and we set $\boldsymbol{\eta}_t := \begin{pmatrix} \eta_t^0 \\ \eta_t \end{pmatrix} = \widehat{\boldsymbol{S}}_t^{-1}(\boldsymbol{\xi}(\varphi(t)))$. Let $\gamma_0, \gamma_T \colon M \to \mathbb{R}$ be two smooth functions such that

$$\mathrm{d}\gamma_0(\widehat{q}_0) = \widehat{\ell}_0, \qquad \mathrm{d}\gamma_T(\widehat{q}_T) = -\widehat{\ell}_T,$$

and let $\widehat{\gamma}_0 := \gamma_0 \circ \widehat{S}_0$, $\widehat{\gamma}_T := \gamma_T \circ \widehat{S}_T$. Then, the cost can be written in terms of the pull-back trajectory $\boldsymbol{\eta}_t$ as

$$J(u) = \widehat{S}_T^0(\boldsymbol{\eta}_T) - \widehat{S}_0^0(\boldsymbol{\eta}_0) + \widehat{\gamma}_0(\eta_0) + \widehat{\gamma}_T(\eta_T). \tag{4.2}$$

Thanks to PMP, it is possible to show that the first variation of J evaluated at $\widehat{u}$ is null (see Appendix A for more details).

In order to compute the second variation of J, we introduce the pullbacks to time $\widehat{\tau}_1$ of the vector fields governing the dynamics of $\boldsymbol{\eta}_t$:

$$g_t^1 := \widehat{S}_{t*}^{-1} f_1 \circ \widehat{S}_t, \quad k_3 := \widehat{S}_{\widehat{\tau}_3*}^{-1} f_0 \circ \widehat{S}_{\widehat{\tau}_3}, \quad k_4 := \widehat{S}_{\widehat{\tau}_3*}^{-1}(f_0 - f_1) \circ \widehat{S}_{\widehat{\tau}_3}, \quad k = k_4 - k_3 := -\widehat{S}_{\widehat{\tau}_3*}^{-1} f_1 \circ \widehat{S}_{\widehat{\tau}_3}.$$

Setting $\varepsilon := -(\tau_3 - \widehat{\tau}_3)$, the second variation of the optimal control problem is given by

$$\begin{aligned} J''[\delta v, \varepsilon]^2 := &\int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) L_{\delta\eta(s)} \big(\widehat{\psi}_s + L_{g_s^1} \big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_s^0 \big) \big)(\widehat{q}_1) \mathrm{d}s \\ &- \frac{\varepsilon^2}{2} L_k^2 \big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \big)(\widehat{q}_1) + \frac{\varepsilon^2}{2} L_{[k_4,k_3]} \big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \big)(\widehat{q}_1) \\ &- \frac{\varepsilon^2}{2} L_{k_4} \widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1), \end{aligned} \tag{4.3}$$

where $\delta v \in L^\infty([\widehat{\tau}_1, \widehat{\tau}_2])$ and $\delta\eta_t$ is the linearisation of η_t at $u = \widehat{u}$, and satisfies the control system

$$\dot{\delta\eta}_t = \begin{cases} 0 & t \in [0, \widehat{\tau}_1), \\ \delta v(t) g_t^1(\widehat{q}_1) & t \in (\widehat{\tau}_1, \widehat{\tau}_2), \\ -\frac{\varepsilon}{\widehat{\tau}_3 - \widehat{\tau}_2} k_3(\widehat{q}_1) & t \in (\widehat{\tau}_2, \widehat{\tau}_3), \\ \frac{\varepsilon}{T - \widehat{\tau}_3} k_4(\widehat{q}_1) & t \in (\widehat{\tau}_3, T], \end{cases} \qquad \begin{aligned} \delta\eta_0 &= 0, \\ \delta\eta_T &= 0. \end{aligned} \tag{4.4}$$

The second variation (4.3) is degenerate, as the quadratic term in δv (Legendre term) is missing; to overcome this issue, we perform a Goh transformation, that is, we integrate the control variation and we add an additional variation ε_0, in the same spirit of [23]:

$$w(t) := \int_t^{\widehat{\tau}_2} \delta v(s) \mathrm{d}s, \quad t \in [\widehat{\tau}_1, \widehat{\tau}_2], \qquad \varepsilon_0 := w(\widehat{\tau}_1).$$

We thus obtain the *extended admissible variations* as the triples $\delta e := (\varepsilon_0, \varepsilon, w) \in \mathbb{R} \times \mathbb{R} \times L^2([\widehat{\tau}_1, \widehat{\tau}_2])$ such that the system

$$\begin{cases} \dot{\zeta}(t) = w(t)\dot{g}_t^1(\widehat{q}_1), \\ \zeta(\widehat{\tau}_1) = \varepsilon_0 f_1(\widehat{q}_1), \\ \zeta(\widehat{\tau}_2) = -\varepsilon k(\widehat{q}_1) \end{cases} \tag{4.5}$$

admits a solution, and the extended second variation as the quadratic form

$$J_e''[\delta e]^2 = \frac{\varepsilon_0^2}{2} L_{f_1}\big(\psi + L_{f_1}\big(\widehat{\gamma}_T + \widehat{S}_T^0\big)\big)(\widehat{q}_1) + \frac{1}{2}\int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \big(w^2(s)R(s) + 2w(s)L_{\zeta(s)}A(s,\widehat{q}_1)\big)\,\mathrm{d}s$$
$$- \frac{\varepsilon^2}{2} L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \frac{\varepsilon^2}{2} L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1) + \frac{\varepsilon^2}{2} L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1),$$

$$(4.6)$$

where we put

$$A(s,q) := \big(L_{g_s^1}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_s^0\big) + L_{g_s^0}\widehat{\psi}_s\big)(q),$$
$$R(s) := L_{[g_s^1,g_s^1]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_s^0\big)(\widehat{q}_1) + L_{g_s^1}\widehat{\psi}_s(\widehat{q}_1) - L_{g_s^1}L_{g_s^0}\widehat{\psi}_s(\widehat{q}_1).$$

Remark 4.2. We remark that $R(s) = \mathbb{L}(\widehat{\lambda}(s))$, which is positive by Assumption 3.9.

We can now state our final assumption.

Assumption 4.3. The quadratic form J_e'' (4.6) is coercive on the space of admissible variations

$$\mathcal{W} := \big\{\delta e = (\varepsilon_0, \varepsilon, w) \in \mathbb{R} \times \mathbb{R} \times L^2([\widehat{\tau}_1, \widehat{\tau}_2]) \text{ such that system (4.5) admits a solution}\big\}.$$

4.1. Consequences of coerciveness of J_e''

Lemma 4.4. *Assume that Assumption 4.3 holds true. Then* $f_1(\widehat{q}_1) \neq 0$.

Proof. Assume by contradiction that $f_1(\widehat{q}_1) = 0$. Then $\delta e := \big(\varepsilon_0 = 1, \varepsilon = 0, w \equiv 0\big)$ is a non-trivial admissible variation and $J_e''[\delta e]^2 = L_{f_1}\big(\psi + L_{f_1}\big(\widehat{\gamma}_T + \widehat{S}_T^0\big)\big)(\widehat{q}_1) = 0$, since $f_1(\widehat{q}_1) = 0$. Thus we have a contradiction. $\square$

Lemma 4.5. *Under Assumption 4.3, there exist a neighbourhood $U_{\widehat{q}_1}$ of $\widehat{q}_1$ in M and a smooth function $\alpha \colon U_{\widehat{q}_1} \to \mathbb{R}$ such that*

$$J_e''[\delta e]^2 = \frac{\varepsilon_0^2}{2} L_{f_1}^2\big(\alpha + \widehat{\gamma}_T + \widehat{S}_T^0\big)(\widehat{q}_1) + \frac{1}{2}\int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \big(w^2(s)R(s) + 2w(s)L_{\zeta(s)}A(s,\widehat{q}_1)\big)\,\mathrm{d}s$$
$$- \frac{\varepsilon^2}{2} L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \frac{\varepsilon^2}{2} L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1) + \frac{\varepsilon^2}{2} L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1).$$

In particular, it holds

$$L_{f_1}\alpha\big(q\big) = \psi(q) \quad \forall q \in U_{\widehat{q}_1}, \qquad \mathrm{d}\alpha(\widehat{q}_1) = \widehat{\ell}_1.$$

Proof. Thanks to Lemma 4.4, we can choose local coordinates $(y_1, y_2, \ldots, y_n)$ around $\widehat{q}_1$ such that $f_1 \equiv \dfrac{\partial}{\partial y_1}$ in a neighbourhood $U_{\widehat{q}_1}$ of $\widehat{q}_1$.

Let $\widehat{\ell}_1 = \big(\widehat{p}_1, \widehat{p}_2, \ldots, \widehat{p}_n\big)$. Thus $\widehat{p}_1 = F_1(\widehat{\ell}_1) = \psi(\widehat{q}_1)$. Possibly shrinking $U_{\widehat{q}_1}$, let $\alpha \colon U_{\widehat{q}_1} \to \mathbb{R}$ be the solution of the Cauchy problem

$$\begin{cases} \dfrac{\partial \alpha}{\partial y_1}(y_1, y_2, \ldots, y_n) = \psi(y_1, y_2, \ldots, y_n), \\ \alpha(0, y_2, \ldots, y_n) = \sum_{j=2}^{n} \widehat{p}_j y_j. \end{cases}$$

Then, by construction, $\mathrm{d}\alpha(\widehat{q}_1) = \big(\psi(\widehat{q}_1), \widehat{p}_2, \ldots, \widehat{p}_n\big) = \widehat{\ell}_1$. The term $L_{f_1}\psi(\widehat{q}_1)$ in J''_e can thus be replaced by $L^2_{f_1}\alpha(\widehat{q}_1)$, and this proves the claim. $\qquad\square$

We claim that the term $L^2_{f_1}\big(\alpha + \widehat{\gamma}_T + \widehat{S}^0_T\big)(\widehat{q}_1)$ in the expression for J''_e can be replaced by $\mathrm{D}^2\big(\alpha + \widehat{\gamma}_T + \widehat{S}^0_T\big)(\widehat{q}_1)[f_1(\widehat{q}_1)]^2$. Indeed, by definition of γ_T, it holds $\mathrm{d}\gamma_T(\widehat{q}_T) = -\widehat{\ell}_T$; since, by (A.1), we have that $\widehat{\ell}_T = \big(\widehat{\ell}_1 + \mathrm{d}\widehat{S}^0_T\big)\widehat{S}^{-1}_{T*}$, we obtain that $\mathrm{d}\alpha(\widehat{q}_1) = \widehat{\ell}_1 = -\big(\mathrm{d}\widehat{\gamma}_T + \mathrm{d}\widehat{S}^0_T\big)(\widehat{q}_1)$, and we are done. We can thus write J''_e on $\mathcal{W}$ as

$$
\begin{aligned}
J''_e[\delta e]^2 = {} & \frac{1}{2}\mathrm{D}^2\big(\alpha + \widehat{\gamma}_T + \widehat{S}^0_T\big)(\widehat{q}_1)[\varepsilon_0 f_1(\widehat{q}_1)]^2 + \frac{1}{2}\int_{\widehat{\tau}_1}^{\widehat{\tau}_2}\big(w^2(s)R(s) + 2w(s)L_{\zeta(s)}A(s,\widehat{q}_1)\big)\,\mathrm{d}s \\
& - \frac{\varepsilon^2}{2}L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \frac{\varepsilon^2}{2}L^2_k\big(\widehat{\gamma}_T + \widehat{S}^0_T - \widehat{S}^0_{\widehat{\tau}_2}\big)(\widehat{q}_1) + \frac{\varepsilon^2}{2}L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}^0_T - \widehat{S}^0_{\widehat{\tau}_2}\big)(\widehat{q}_1).
\end{aligned}
\tag{4.7}
$$

We now extend the space of admissible variations, in the following way: we remove the constraint on $\zeta(\widehat{\tau}_1)$ and we consider the control system

$$
\begin{cases}
\dot{\zeta}(s) = w(s)\dot{g}^1_s(\widehat{q}_1), \\
\zeta(\widehat{\tau}_1) = \delta z \in T_{\widehat{q}_1}M, \\
\zeta(\widehat{\tau}_2) = -\varepsilon k(\widehat{q}_1).
\end{cases}
\tag{4.8}
$$

The associated space of variations is thus defined as

$$
\mathcal{W} := \big\{\delta e = (\delta z, \varepsilon, w) \in T_{\widehat{q}_1}M \times \mathbb{R} \times L^2([\widehat{\tau}_1, \widehat{\tau}_2])\ \text{such that system (4.8) admits a solution}\big\}.
$$

Applying ([15], Thm. 11.6), we can easily verify that (4.7) defines a Legendre form on $\mathcal{W}$. For $s \in \mathbb{R}^+$, we define $\theta(y_1, y_2, \ldots, y_n) := \frac{s}{2}\sum_{j=2}^n y_j^2$ where $(y_1, y_2, \ldots y_n)$ are the coordinates defined in the proof of Lemma 4.5. Applying ([15], Thm. 13.2), we obtain that, under Assumption 4.3 and if s is large enough, then the quadratic form

$$
\boldsymbol{J}''[\delta e]^2 := J''_e[\delta e]^2 + \frac{1}{2}\mathrm{D}^2\theta[\delta z]^2
$$

is coercive on $\mathcal{W}$.

We can write $\boldsymbol{J}''$ explicitly as

$$
\begin{aligned}
\boldsymbol{J}''[\delta e]^2 = {} & \frac{1}{2}\mathrm{D}^2\big(\alpha + \theta + \widehat{\gamma}_T + \widehat{S}^0_T\big)(\widehat{q}_1)[\delta z]^2 + \frac{1}{2}\int_{\widehat{\tau}_1}^{\widehat{\tau}_2}\big(w^2(s)R(s) + 2w(s)L_{\zeta(s)}A(s,\widehat{q}_1)\big)\,\mathrm{d}s \\
& - \frac{\varepsilon^2}{2}L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \frac{\varepsilon^2}{2}L^2_k\big(\widehat{\gamma}_T + \widehat{S}^0_T - \widehat{S}^0_{\widehat{\tau}_2}\big)(\widehat{q}_1) + \frac{\varepsilon^2}{2}L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}^0_T - \widehat{S}^0_{\widehat{\tau}_2}\big)(\widehat{q}_1).
\end{aligned}
\tag{4.9}
$$

5. Construction of the over-maximised flow

As briefly mentioned in the Introduction, Hamiltonian methods to prove the optimality of an extremal need two ingredients: a Lagrangian submanifold Λ_1 of the cotangent bundle, containing a point of the reference extremal, and a over-maximised Hamiltonian flow (that is, the flow associated with a Hamiltonian function which is greater than or equal to $H_{\max}$, and which coincides with it along the reference extremal, at least up to the first order); the optimality of the reference extremal is proved if, at each time $t \in [0, T]$, the image of Λ_1 under the flow of the over-maximised Hamiltonian projects one-to-one onto a neighbourhood of $\widehat{\xi}(t)$ in the base manifold (see Sect. 7).

This section is devoted to the construction of the over-maximised Hamiltonian and of its associated flow, that we refer to as *over-maximised flow*. It is defined patching together some piecewise smooth Hamiltonian flows, each one defined on a suitable neighbourhood of each arc. We stress that, to achieve this task, to coerciveness of J_e'' is not needed; though, we extensively use Assumptions 3.1–3.10, which are assumed to hold true throughout the whole section.

5.1. The over-maximised Hamiltonian near the singular arc

By definition, the singular arc evolves on the hypersurface $\Sigma^- = \{\Phi^- = 0\}$ and $\Phi^+\big(\widehat{\lambda}(t)\big) > 0$. Thus, in a sufficiently small neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$, the maximised Hamiltonian is given by

$$H_{\max} = \begin{cases} F_0 + \Phi^- & \text{if } \Phi^- \geq 0, \\ F_0 & \text{if } \Phi^- \leq 0. \end{cases}$$

Therefore, $H_{\max}$ is continuous, but its associated Hamiltonian vector field is not well defined on Σ^-: for every smooth Hamiltonian $v(t, \ell)$, every Hamiltonian of the form $F_0 + v\Phi^-$ coincides with $H_{\max}$ on Σ^-.

On the other hand, no Hamiltonian of the form $F_0 + v\Phi^-$ can be used in place of the maximised Hamiltonian: indeed, by Assumption 3.9, for every $t \in (\widehat{\tau}_1, \widehat{\tau}_2)$ and every $\mathcal{U}$ neighbourhood of $\widehat{\lambda}(t)$ in Σ^-, there exists some $\ell \in \mathcal{U}$ such that $F_{01}(\ell) - L_{f_0}\psi(\pi\ell) < 0$. For any choice of $v > 0$, the flow of $\vec{F}_0 + v\vec{\Phi}^-$ sends ℓ in the region $\{\Phi^- < 0\}$, where $F_0 + v\Phi^-$ is no more the maximised Hamiltonian. Indeed, whatever the choice of v, the Hamiltonian vector field $\vec{F}_0 + v\vec{\Phi}^-$ is tangent to Σ only on $S^- = \{\ell \in \Sigma^- : F_{01}(\ell) - L_{f_0}\psi(\pi\ell) = 0\}$. We refer to [23] for a detailed description of the phenomenon.

However, as proposed in [28], the flow associated with $H_{\max}$ may be replaced by the flow of a suitable over-maximised Hamiltonian tangent to Σ^-, at least for $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$. The construction of such a flow in a neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$ in Σ^- relies on the definition of a C^1 over-maximised Hamiltonian that agrees with the maximised one at least up to the first order along the reference extremal, and whose Hamiltonian vector field is tangent to Σ^- (see [12, 23, 28, 30] for similar constructions).

In order to do so, following the steps of [11, 12], we substitute F_0 with a suitable Hamiltonian H_0 which is constant along the integral lines of Φ^-.

Lemma 5.1. *There exist a neighbourhood $\mathcal{U}$ of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$ in T^*M and a smooth function $\vartheta\colon \mathcal{U} \to \mathbb{R}$ such that*

$$\big(F_{01} - L_{f_0}\psi \circ \pi\big) \circ \exp\big(\vartheta(\ell)\vec{\Phi}^-\big)(\ell) = 0 \qquad \forall \ell \in \mathcal{U},$$

and, for any $\ell \in \mathcal{U} \cap S^-$,

$$\langle \mathrm{d}\vartheta(\ell), (\cdot) \rangle = -\frac{1}{\mathbb{L}(\ell)} \big(\langle \mathrm{d}F_{01}(\ell), (\cdot) \rangle - L_{\pi_*(\cdot)}L_{f_0}\psi\big(\pi\ell\big)\big). \tag{5.1}$$

Proof. By Assumption 3.9, we can apply the implicit function theorem to the function $(s, \ell) \mapsto \big(F_{01} - L_{f_0}\psi \circ \pi\big) \circ \exp\big(s\vec{\Phi}^-\big)(\ell)$ at the point $(0, \widehat{\lambda}(t))$, and obtain the result. $\square$

Thanks to Lemma 5.1, we can define the following Hamiltonian $H_0\colon \mathcal{U} \to \mathbb{R}$

$$H_0(\ell) = F_0 \circ \exp(\vartheta(\ell)\vec{\Phi}^-)(\ell).$$

Proposition 5.2. *The Hamiltonian H_0 satisfies the following properties.*

1. For every $\ell \in \Sigma^- \cap \mathcal{U}$, $\vec{H}_0(\ell)$ is tangent to Σ^- and is given by

$$\vec{H}_0(\ell) = \exp(-s\vec{\Phi}^-)_* \vec{F}_0 \circ \exp(s\vec{\Phi}^-)(\ell)|_{s=\vartheta(\ell)}. \tag{5.2}$$

    In particular, it coincides with $\vec{F}_0(\ell)$ if $\ell \in S^- \cap \mathcal{U}$.

2. *Possibly shrinking $\mathcal{U}$, $H_0(\ell) \geq F_0(\ell)$ for any $\ell \in \Sigma^- \cap \mathcal{U}$. Equality holds if and only if $\ell \in S^-$.*

3. *For every smooth function $v(t,\ell) \colon \mathbb{R} \times T^*M \to \mathbb{R}$, the Hamiltonian vector field associated with $H_0(\ell) + v(t,\ell)\Phi^-(\ell)$ is tangent to Σ^-.*

Proof. Claim 1 easily follows from the fact that $\langle \mathrm{d}F_0, \vec{\Phi}^- \rangle = 0$ on S^-. Indeed, by construction, H_0 is constant along the integral lines of $\vec{\Phi}^-$, so that $\langle \mathrm{d}\Phi^-, \vec{H}_0 \rangle = \langle \mathrm{d}H_0, \vec{\Phi}^- \rangle = 0$, for any $\ell \in \mathcal{U}$, *i.e.* $\vec{H}_0$ is tangent to Σ^-. Equation (5.2) can be verified with simple computations.

Since $\vartheta = 0$ on S^-, by definition and equation (5.2) we have that $H_0 = F_0$ and $\vec{H}_0 = \vec{F}_0$ on S^-. Thus the differential of $H_0 - F_0$ is identically zero on S^-, and, for any $\ell \in S^- \cap \mathcal{U}$, we can compute its second derivative:

$$\mathrm{D}^2(H_0 - F_0)(\ell)[\delta\ell]^2 = 2\langle \mathrm{d}\left(L_{\vec{\Phi}^-}F_0\right)(\ell), \delta\ell \rangle \langle \mathrm{d}\vartheta(\ell), \delta\ell \rangle + L_{\vec{\Phi}^-}^2 F_0(\ell)\langle \mathrm{d}\vartheta(\ell), \delta\ell \rangle^2$$
$$= -L_{\vec{\Phi}^-}^2 F_0(\ell)\langle \mathrm{d}\vartheta(\ell), \delta\ell \rangle^2.$$

Noticing that $L_{\vec{\Phi}^-}^2 F_0(\ell) = -\mathbb{L}(\ell)$ and thanks to Assumption 3.9, we see that the expression here above is non negative and it vanishes only if $\langle \mathrm{d}\vartheta(\ell), \delta\ell \rangle = 0$; equation (3.8) and the fact that $\langle \mathrm{d}\vartheta(\ell), \vec{\Phi}^-(\ell) \rangle = -1$ for $\ell \in \Sigma^- \cap \mathcal{U}$ prove claim 2.

Since $\vec{H}_0$ is tangent to Σ^-, then $\vec{H}_0 + u\vec{\Phi}^-$ is tangent to Σ^- too, for every $u \in \mathbb{R}$. Let v be a smooth function on $\mathbb{R} \times T^*M$. Then the Hamiltonian field associated with $H_0 + v(t,\ell)\Phi^-$ is given by

$$\vec{H}_0(\ell) + v(t,\ell)\vec{\Phi}^- + \Phi^-(\ell)\vec{v}(t,\ell),$$

and, by definition of Σ^-, this completes the proof. $\qquad\square$

5.2. The over-maximised flow

    Thanks to the regularity of the bang and zero arcs, for any $t \in [0, \hat{\tau}_1) \cup (\hat{\tau}_2, \hat{\tau}_3) \cup (\hat{\tau}_3, T]$, it is possible to find a neighbourhood of $\hat{\lambda}(t)$ where $H_{\max}$ and its associated vector field are unambiguously defined. As observed in the previous section, this is no longer true for $t \in [\hat{\tau}_1, \hat{\tau}_2]$. On the other hand, thanks to Proposition 5.2, we know that any over-maximised Hamiltonian of the form $H_0 + v\Phi^-$ may replace the maximised Hamiltonian in a neighbourhood of the singular arc in Σ^-.

    Here below we show how to concatenate the flows of $H_{\max}$ and of the over-maximised Hamiltonian, in order to obtain a flow defined for all $t \in [0, T]$.

The first bang arc.

    We first construct the over-maximised flow for $t \in [0, \hat{\tau}_1]$.

    Proposition 5.2 guarantees that $H_0 \geq F_0$ in a neighbourhood of $\hat{\lambda}([\hat{\tau}_1, \hat{\tau}_2])$ contained in Σ^- only. In other words, if we want to use H_0 to construct the over-maximised Hamiltonian, we have to be sure that, for $t \in [\hat{\tau}_1, \hat{\tau}_2]$, the image of the sub-manifold Λ_1 under the over-maximised flow is in Σ^-. For this reason, it is convenient to start from $\hat{\tau}_1$ and construct the flow integrating backward in time.

    Fix some $\epsilon > 0$ and consider a sufficiently small tubular neighbourhood $\mathbf{U}$ of $\hat{\lambda}((\hat{\tau}_1 - \epsilon, \hat{\tau}_1 + \epsilon))$ in T^*M. The manifold Σ^- separates $\mathbf{U}$ in two regions, one in which $\Phi^- > 0$ (and $H_{\max} = F_0 + \Phi^-$), the other one in which $\Phi^- < 0 < \Phi^+$ (and $H_{\max} = F_0$), since $\psi(q) \neq 0$ in a neighbourhood of $\hat{\xi}(\hat{\tau}_1)$; in particular, the first bang arc is contained in the first region. On the other hand, by Assumption 3.9, the manifold S^- separates $\mathbf{U} \cap \Sigma^-$ into two regions, in which $F_{01} - L_{f_0}\psi \circ \pi$ has different sign. Consider now a small neighbourhood of $\hat{\ell}_1$ in Σ^-. The trajectories obtained by integrating backward in time the flow generated by $\vec{F}_0 + \vec{\Phi}^-$, starting at $t = \hat{\tau}_1$ from a point ℓ satisfying $F_{01}(\ell) - L_{f_0}\psi(\pi\ell) < 0$, immediately leave Σ^- and enter in the region $\{\Phi^- > 0\}$; in particular, they evolve with the maximised flow and stay close to the reference extremal (if ℓ is sufficiently close to $\hat{\ell}_1$). The same happens for trajectories starting from a point ℓ such that $F_{01}(\ell) - L_{f_0}\psi(\pi\ell) = 0$, thanks to Assumption 3.8.

On the contrary, the integral curves of $\vec{F}_0 + \vec{\Phi}^-$ with an initial condition ℓ satisfying $F_{01}(\ell) - L_{f_0}\psi(\pi\ell) > 0$ immediately enter into the region $\{\Phi^- < 0\}$, so that they are not integral curves of $\vec{H}_{\max}$ and may soon leave $\mathbf{u}$. To fix this issue, for initial conditions belonging to the region where $F_{01}(\ell) - L_{f_0}\psi(\pi\ell) > 0$, we substitute the flow of $\vec{F}_0 + \vec{\Phi}^-$ with the one of $\vec{H}_0 + \vec{\Phi}^-$, until the trajectories reach S^-. This construction is explained in Proposition 5.5 here below, whose proof relies on the following lemma.

Lemma 5.3. *There exist a neighbourhood $\mathcal{O}_1$ of $\widehat{\ell}_1$ in Σ^- and a smooth function $t_1 \colon \mathcal{O}_1 \to \mathbb{R}$ satisfying $t_1(\widehat{\ell}_1) = \widehat{\tau}_1$ such that*

$$(F_{01} - (L_{f_0}\psi) \circ \pi) \circ \exp\big((t_1(\ell) - \widehat{\tau}_1)(\vec{H}_0 + \vec{\Phi}^-)\big)(\ell) = 0 \quad \forall \ell \in \mathcal{O}_1.$$

Moreover, $t_1(\ell) \gtreqless \widehat{\tau}_1$ if and only if $F_{01}(\ell) - L_{f_0}\psi(\pi\ell) \lesseqgtr 0$.

Proof. The existence of the function t_1 is a straightforward application of the implicit function theorem to the function $\varphi(t, \ell) = (F_{01} - (L_{f_0}\psi) \circ \pi) \circ \exp((t - \widehat{\tau}_1)(\vec{H}_0 + \vec{\Phi}^-))(\ell)$ at $(\widehat{\tau}_1, \widehat{\ell}_1)$, which is possible since $\frac{\partial}{\partial t}\varphi(t, \ell)|_{(\widehat{\tau}_1, \widehat{\ell}_1)} > 0$, by Assumption 3.8.

The sign of $t_1(\ell) - \widehat{\tau}_1$ is determined by the fact that $t_1(\ell) = \widehat{\tau}_1$ for every $\ell \in S^- \cap \mathcal{O}_1$, and again by Assumption 3.8. $\qquad\square$

We define the piecewise smooth function $\tau_1 \colon \mathcal{O}_1 \to \mathbb{R}$ as

$$\tau_1(\ell) = \min\{t_1(\ell), \widehat{\tau}_1\},$$

and the flow $\mathcal{H}_1 \colon [0, \widehat{\tau}_1] \times \mathcal{O}_1 \to T^*M$ as

$$\mathcal{H}_1(t, \ell) = \begin{cases} \exp\big((t - \widehat{\tau}_1)(\vec{H}_0 + \vec{\Phi}^-)\big)(\ell) & t \in [\tau_1(\ell), \widehat{\tau}_1], \\ \exp\big((t - \tau_1(\ell))(\vec{F}_0 + \vec{\Phi}^-)\big) \circ \exp\big((\tau_1(\ell) - \widehat{\tau}_1)(\vec{H}_0 + \vec{\Phi}^-)\big)(\ell) & t \in [0, \tau_1(\ell)). \end{cases}$$

Remark 5.4. Clearly, if $t_1(\bar{\ell}) \geq \widehat{\tau}_1$ for some $\bar{\ell}$, then $\mathcal{H}_1(t, \bar{\ell})$ is the flow of $\vec{F}_0 + \vec{\Phi}^-$ for every $t \in [0, \widehat{\tau}_1]$.

Proposition 5.5. *The flow $\mathcal{H}_1$ defined above is C^1. Moreover*

$$\Phi^-(\mathcal{H}_1(t, \ell)) = 0 \quad \forall t \in [\tau_1(\ell), \widehat{\tau}_1], \qquad \Phi^-(\mathcal{H}_1(t, \ell)) > 0 \quad \forall t \in [0, \tau_1(\ell)). \tag{5.3}$$

In particular

$$\big(\pi\mathcal{H}_1(t, \cdot)\big)_*\big|_{\widehat{\ell}_1} = \widehat{S}_{t*}\pi_* \quad \forall t \leq \widehat{\tau}_1. \tag{5.4}$$

Proof. At every point (t, ℓ) such that $t \neq \tau_1(\ell)$, the flow is well defined and smooth. Therefore, to prove its regularity on the whole $[0, \widehat{\tau}_1] \times \mathcal{O}_1$, it suffices to verify the continuity of its derivatives at points of the form $(t, \ell) = (\tau_1(\bar{\ell}), \bar{\ell})$. In particular, we can distinguish two cases, that is, $t_1(\bar{\ell}) > \widehat{\tau}_1$ and $t_1(\bar{\ell}) \leq \widehat{\tau}_1$. In the former case, the flow coincides with $\exp((t - \widehat{\tau}_1)(\vec{F}_0 + \vec{\Phi}^-))(\ell)$ for every $t \in [0, \widehat{\tau}_1]$ and for every ℓ in a neighbourhood of $\bar{\ell}$ in Σ^-, thus it is C^1.

If instead $\tau_1(\bar{\ell}) \leq \widehat{\tau}_1$, then the flow starting from points ℓ in a neighbourhood of $\bar{\ell}$ has different expressions according to the sign of $t - \tau_1(\ell)$. However, by straightforward computations, it is easy to prove that they coincide as $(t, \ell) \to (\tau_1(\bar{\ell}), \bar{\ell})$, for every $\bar{\ell} \in \mathcal{O}$ and that the first order partial derivatives are continuous.

Let us now prove equation (5.3). First of all, we recall that $\frac{\partial}{\partial t}\Phi^-(\mathcal{H}_1(t, \ell))|_{t=\tau_1(\ell)} = (F_{01} - (L_{f_0}\psi) \circ \pi) \circ \mathcal{H}_1(t, \ell)|_{t=\tau_1(\ell)}$.

If $t_1(\ell) > \widehat{\tau}_1$, then $\tau_1(\ell) = \widehat{\tau}_1$ and, by Lemma 5.3, $F_{01} - (L_{f_0}\psi) \circ \pi < 0$, so that, if ℓ is close enough to $\widehat{\ell}_1$, equation (5.3) follows immediately from a first order Taylor expansion, with respect to the first variable, at $t = \widehat{\tau}_1$.

Let us now consider the case in which $t_1(\ell) \leq \widehat{\tau}_1$. By construction, $\vec{H}_0 + \vec{\Phi}^-$ is tangent to Σ^-, so that $\Phi^-(\mathcal{H}_1(t,\ell)) = 0$ for $t \in [\tau_1(\ell), \widehat{\tau}_1]$. At $t = \tau_1(\ell)$, $\mathcal{H}_1(t,\ell)$ is in S^-, that is, $(F_{01} - (L_{f_0}\psi) \circ \pi)) \circ \mathcal{H}_1(t,\ell)|_{t=\tau_1(\ell)} = 0$, so that we must look at the second order Taylor expansion of $t \mapsto \Phi^-(\mathcal{H}_1(t,\ell))$ at $\widehat{\tau}_1$. By Assumption 3.8, the second order derivative of this map at $(t,\ell) = (\widehat{\tau}_1, \widehat{\ell}_1)$ is strictly positive, so that, by continuity, it is strictly positive also at $(\tau_1(\ell), \ell)$, for ℓ close enough to $\widehat{\ell}_1$.

We can conclude that there exists a $\epsilon > 0$ and a neighbourhood $\mathcal{O}_1$ of $\widehat{\ell}_1$ in Σ^- such that $\Phi^-(\mathcal{H}_1(t,\ell)) > 0$ for $t \in (\tau_1(\ell) - \epsilon, \tau_1(\ell))$, for every $\ell \in \mathcal{O}_1$. Possibly shrinking $\mathcal{O}_1$, we can conclude that the inequality is satisfied for every $t \in [0, \tau_1(\ell))$. $\qquad\square$

The singular arc. We recall that for any $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, the reference extremal $\widehat{\lambda}$ takes values in S^-. Moreover, thanks to Assumption 3.9 and since $\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Phi^-(\widehat{\lambda}(t)) = 0$ for any $t \in (\widehat{\tau}_1, \widehat{\tau}_2)$, the reference control along the singular arc can be computed in a *feedback Hamiltonian form.* More precisely

$$\widehat{u}(t) = -\frac{F_{001}(\ell) - L_{f_0}^2\psi(\pi\ell)}{\mathbb{L}(\ell)}\Big|_{\ell=\widehat{\lambda}(t)} \qquad \forall t \in (\widehat{\tau}_1, \widehat{\tau}_2).$$

In a neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$ in S^-, we thus define

$$\nu(\ell) = -\frac{F_{001}(\ell) - L_{f_0}^2\psi(\pi\ell)}{\mathbb{L}(\ell)}.$$

We extend ν to a neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$ in Σ^- by setting it constant along the integral lines of $\vec{\Phi}^-$, and then to a full-measure neighbourhood of the range of the singular arc by setting it constant along the integral lines of the Hamiltonian field associated with $F_{01} - L_{f_0}\psi$[3].

We set

$$K(\ell) = H_0(\ell) + \nu(\ell)\Phi^-(\ell),$$

and define the over-maximised flow on the interval $[\widehat{\tau}_1, \widehat{\tau}_2]$ as the flow of $\vec{K}$:

$$\mathcal{K}(t,\ell) = \exp((t - \widehat{\tau}_1)\vec{K})(\ell).$$

Proposition 5.6. *The manifolds Σ^- and S^- are invariant under the action of the flow of K. Moreover, $\vec{\Phi}^-$ is invariant with respect to the flow of K on Σ^-, that is, for ℓ belonging to a small neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \tau_2])$ in Σ^-, it holds*

$$\mathcal{K}(t,\ell)_*\vec{\Phi}^-(\ell) = \vec{\Phi}^- \circ \mathcal{K}(t,\ell) \qquad \forall t \in [\widehat{\tau}_1, \widehat{\tau}_2].$$

This result is proved (in a more general version) in Proposition B.1.

The inactivated arc. The construction of the over-maximised flow on a right hand side neighbourhood of $\widehat{\tau}_2$ presents the same issues as its construction on $[0, \widehat{\tau}_1]$, thus we overcome these difficulties likewise.

---

[3] Indeed, thanks to Assumption 3.9, $T_\ell(T^*M) = T_\ell\Sigma^- \oplus \overrightarrow{F_{01} - L_{f_0}\psi \circ \pi}(\ell)$ for every $\ell \in \Sigma^-$ in a neighbourhood of $\widehat{\lambda}([\widehat{\tau}_1, \widehat{\tau}_2])$.

Lemma 5.7. *Possibly shrinking $\mathcal{O}_1$ and setting $\mathcal{O}_2 = \mathcal{K}(\widehat{\tau}_2, \mathcal{O}_1)$, there exists a smooth function $t_2 \colon \mathcal{O}_2 \to \mathbb{R}$ satisfying $t_2(\widehat{\ell}_2) = \widehat{\tau}_2$ such that*

$$(F_{01} - (L_{f_0}\psi) \circ \pi) \circ \exp((t_2(\ell_2) - \widehat{\tau}_2)\vec{H}_0)(\ell_2) = 0 \quad \forall \ell_2 \in \mathcal{O}_2.$$

Moreover, $t_2(\ell_2) \gtreqless \widehat{\tau}_2$ if and only if $F_{01}(\ell_2) - L_{f_0}\psi(\pi\ell_2) \gtreqless 0$.

As above, we define the piecewise smooth function

$$\tau_2(\ell_2) = \max\{t_2(\ell_2), \widehat{\tau}_2\},$$

and the flow $\mathcal{H}_3$ for $t \geq \widehat{\tau}_2$ as

$$\mathcal{H}_3(t, \ell) = \begin{cases} \exp((t - \widehat{\tau}_2)\vec{H}_0)(\ell_2) & t \in [\widehat{\tau}_2, \tau_2(\ell_2)], \\ \exp((t - \tau_2(\ell_2))\vec{F}_0) \circ \exp((\tau_2(\ell_2) - \widehat{\tau}_2)\vec{H}_0)(\ell_2) & t \in [\tau_2(\ell_2), \tau_2(\ell_2) + \delta(\ell_2)) \end{cases} \tag{5.5}$$

where $\ell_2 = \mathcal{K}(\widehat{\tau}_2, \ell)$ and $\delta(\cdot)$ is a positive function that will be specified here below. The flow $\mathcal{H}_3$ enjoys the same properties of $\mathcal{H}_1$, as stated in the following proposition.

Proposition 5.8. *The flow $\mathcal{H}_3$ defined above is C^1 and*

$$\Phi^-(\mathcal{H}_3(t, \ell)) = 0 \quad \forall t \in [\widehat{\tau}_2(\ell), \tau_2(\ell_2)], \qquad \Phi^-(\mathcal{H}_3(t, \ell)) < 0 \quad \forall t \in (\tau_2(\ell_2), \tau_2(\ell_2) + \delta(\ell_2)).$$

Moreover, for every $t \in [\widehat{\tau}_2, \widehat{\tau}_2 + \delta(\widehat{\ell}_2))$ it holds

$$\big(\pi\mathcal{H}_3(t, \cdot)\big)_*|_{\widehat{\ell}_1} = \exp\big((t - \widehat{\tau}_2)f_0\big)_* \big(\pi\mathcal{K}_{\widehat{\tau}_2}\big)_*|_{\widehat{\ell}_1}. \tag{5.6}$$

We remark that, thanks to Assumption 3.1 and by continuity, $\Phi^+(\mathcal{H}_3(t, \ell)) > 0$ on $[\widehat{\tau}_2, \widehat{\tau}_2 + \delta(\ell_2))$.

The last bang arc. For $t \geq \tau_2(\ell_2)$, F_0 is the maximised Hamiltonian until its integral curves hit the switching surface $\{\Phi^+ = 0\}$. To detect the hitting time, we solve the implicit equation

$$\Phi^+ \circ \exp\big((t - \tau_2(\ell_2))\vec{F}_0\big) \circ \exp\big((\tau_2(\ell_2) - \widehat{\tau}_2)\vec{H}_0\big)(\ell_2) = 0, \tag{5.7}$$

where we recall that $\ell_2 = \mathcal{K}(\widehat{\tau}_2, \ell) \in \mathcal{O}_2$. The derivative with respect to t of the left hand side of (5.7) equals $\mathfrak{r}_3$ for $(t, \ell_2) = (\widehat{\tau}_3, \widehat{\ell}_2)$; thanks to Assumption 3.8 and the implicit function theorem, we obtain that equation (5.7) is satisfied if and only if $(t, \ell_2) = (\tau_3(\ell_2), \ell_2)$, where $\tau_3 \colon \mathcal{O}_2 \to \mathbb{R}$ is a smooth function satisfying $\tau_3(\widehat{\ell}_2) = \widehat{\tau}_3$. In addition, for every $\delta\ell \in T_{\widehat{\ell}_2}(T^*M)$, it holds

$$\langle \mathrm{d}\tau_3(\widehat{\ell}_2), \delta\ell \rangle = -\frac{\boldsymbol{\sigma}_{\widehat{\ell}_3}\big(\exp\big((\widehat{\tau}_3 - \widehat{\tau}_2)\vec{F}_0\big)_*\delta\ell, \vec{\Phi}^+\big)}{\mathfrak{r}_3}. \tag{5.8}$$

We choose $\delta(\cdot) = \tau_3(\cdot) - \tau_2(\cdot)$ in equation (5.5), and we consider the following flow for $t \in [\widehat{\tau}_2, T]$:

$$\mathcal{H}_3(t, \ell) = \begin{cases} \exp\big((t - \widehat{\tau}_2)\vec{H}_0\big)(\ell_2) & t \in [\widehat{\tau}_2, \tau_2(\ell_2)] \\ \exp\big((t - \tau_2(\ell))\vec{F}_0\big) \circ \mathcal{H}_3(\tau_2(\ell_2), \ell_2) & t \in [\tau_2(\ell_2), \tau_3(\ell_2)] \\ \exp\big((t - \tau_3(\ell_2))(\vec{F}_0 - \vec{\Phi}^+)\big) \circ \mathcal{H}_3(\tau_3(\ell_2), \ell_2) & t \geq \tau_3(\ell_2). \end{cases}$$

Thanks to the regularity assumptions, $\mathcal{H}_3$ is an over-maximised flow for every $t \in [\widehat{\tau}_2, T]$.

Finally, the over-maximised flow $\mathcal{H}\colon [0,T] \times \mathcal{O}_1 \to T^*M$ is defined as

$$\mathcal{H}(t,\ell) = \begin{cases} \mathcal{H}_1(t,\ell) & t \in [0,\widehat{\tau}_1], \\ \mathcal{K}(t,\ell) & t \in [\widehat{\tau}_1,\widehat{\tau}_2], \\ \mathcal{H}_3(t,\ell) & t \in [\widehat{\tau}_2,T]. \end{cases}$$

Here below, we will also use the notations

$$\mathcal{H}_t = \mathcal{H}(t,\cdot), \qquad \mathcal{K}_t = \mathcal{K}(t,\cdot).$$

We remark that, for every $t \in [0,T]$, $\mathcal{H}_t$ is the Hamiltonian flow associated with the Hamiltonian

$$H_t(\ell) = \begin{cases} F_0(\ell) + \Phi^-(\ell) & t \in [0,\tau_1(\ell)], \\ H_0(\ell) + \Phi^-(\ell) & t \in [\tau_1(\ell),\widehat{\tau}_1], \\ K(\ell) & t \in [\widehat{\tau}_1,\widehat{\tau}_2], \\ H_0(\ell) & t \in [\widehat{\tau}_2,\tau_2(\ell)], \\ F_0(\ell) & t \in [\tau_2(\ell),\tau_3(\ell)], \\ F_0(\ell) - \Phi^+(\ell) & t \in [\tau_3(\ell),T]. \end{cases}$$

6. INVERTIBILITY

In order to define a manifold Λ_1 such that $\pi\mathcal{H}_t\colon \Lambda_1 \to M$ is locally one-to-one for every t, we shall also exploit the coerciveness of the extended second variation, *i.e.* Assumption 4.3. So, from now, we assume that all the Assumptions 3.1–4.3 are satisfied. We define Λ_1 by means of the functions α and θ appearing in (4.9): namely, we consider the Lagrangian submanifold

$$\Lambda_1 = \{\mathrm{d}(\alpha+\theta)(q)\colon q \in U_{\widehat{q}_1}\}.$$

Remark 6.1. It is immediate to see that $\Lambda_1 \subset \Sigma^-$ and that $\vec{\Phi}^-(\ell) \in T_\ell\Lambda_1$ for every $\ell \in \Lambda_1$.

This section is devoted to the proof of the following result.

Proposition 6.2. *For every $t \in [0,T]$, $t \neq \widehat{\tau}_3$, the flow $\pi\mathcal{H}_t\colon \Lambda_1 \to M$ is a local diffeormorphism from a neighbourhood of $\widehat{\ell}_1$ onto a neighbourhood of $\widehat{\xi}(t)$. $\pi\mathcal{H}_{\widehat{\tau}_3}\colon \Lambda_1 \to M$ is a locally invertible Lipschitz continuous map with Lipschitz continuous inverse.*

The proof is done in several steps: we consider separately the sub-intervals $[0,\widehat{\tau}_1]$, $[\widehat{\tau}_1,\widehat{\tau}_2]$ (illustrated in Fig. 1) and $[\widehat{\tau}_2,T]$ (illustrated in Fig. 2).

Invertibility for $t \in [0,\widehat{\tau}_1]$. The invertibility for $t \in [0,\widehat{\tau}_1]$ is a direct consequence of equation (5.4).

Invertibility for $t \in [\widehat{\tau}_1,\widehat{\tau}_2]$. In order to prove the claim, we introduce the auxiliary Hamiltonian

$$\widehat{H}_t = H_0 + \widehat{u}(t)\Phi^-, \qquad t \in [\widehat{\tau}_1,\widehat{\tau}_2].$$

This new Hamiltonian shares some important features with K. In particular, $\widehat{H}_t$ is an over-maximised Hamiltonian on Σ^- too, and its Hamiltonian vector field is tangent to Σ^-; we denote with $\widehat{\mathcal{H}}_t$ its flow from time $\widehat{\tau}_1$ to time t. The invertibility of $\pi\widehat{\mathcal{H}}_t|_{\Lambda_1}$ is related to the invertibility of $\pi\mathcal{K}_t|_{\Lambda_1}$, as the following result shows.

The proof uses the same arguments of ([23], Lem. 9); we sketch it in the appendix.

Lemma 6.3. *For every $t \in [\widehat{\tau}_1,\widehat{\tau}_2]$, the followings hold*

1. $\widehat{\mathcal{H}}_{t*}(T_{\widehat{\ell}_1}\Lambda_1) = \mathcal{K}_{t*}(T_{\widehat{\ell}_1}\Lambda_1)$.
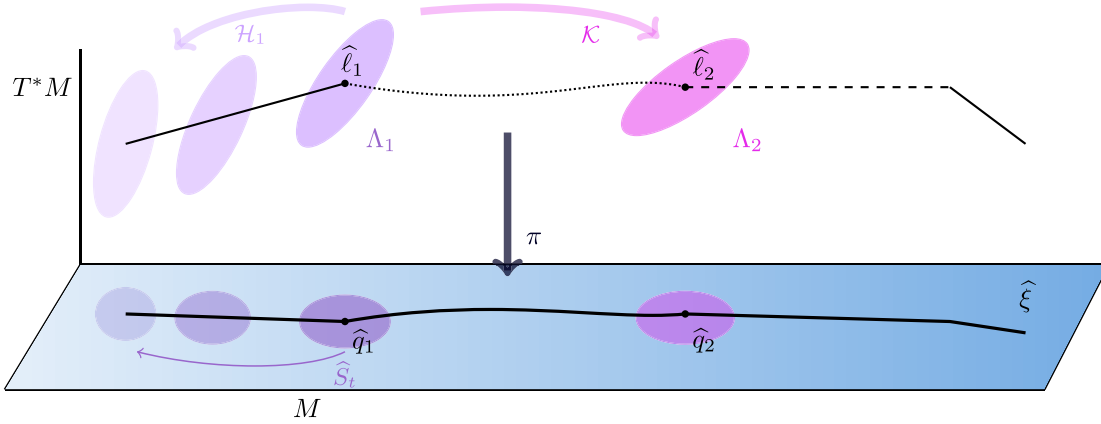
FIGURE 1. Invertibility for $t \in [0, \widehat{\tau}_3)$: $\mathcal{H}_t(\Lambda_1)$ projects diffeormorphically onto a neighbourhood of $\widehat{\xi}(t)$ for every $t \in [0, \widehat{\tau}_3)$. Solid lines denote the bang arcs, dashed lines inactivated arcs, and dotted lines singular arcs.

2. If $\ker\left(\pi\widehat{\mathcal{H}}_t\right)_*|_{T_{\widehat{\ell}_1}\Lambda_1} = 0$, then $\ker\left(\pi\mathcal{K}_t\right)_*|_{T_{\widehat{\ell}_1}\Lambda_1} = 0$.

Taking advantage of Lemma 6.3, the invertibility of $\pi\widehat{\mathcal{H}}_t|_{\Lambda_1}$ implies the one of $\pi\widehat{\mathcal{K}}_t|_{\Lambda_1}$. On the other hand, it turns out that $\widehat{\mathcal{H}}_t$ is directly linked to the Hamiltonian flow associated with the second variation (see details here below); therefore, it is much easier to prove the invertibility of $\pi\widehat{\mathcal{H}}_t|_{\Lambda_1}$ as a consequence of the coerciveness of the second variation. Indeed, consider the subspace $\boldsymbol{\mathcal{V}} \subset \boldsymbol{\mathcal{W}}$ defined by

$$\boldsymbol{\mathcal{V}} := \left\{\delta e \in \boldsymbol{\mathcal{W}} \colon \varepsilon = 0\right\}.$$

Lemma 6.4. *Assume that $\boldsymbol{J}''|_{\boldsymbol{\mathcal{V}}}$ is coercive. Then* $\ker\left(\pi\widehat{\mathcal{H}}_t\right)_*|_{T_{\widehat{\ell}_1}\Lambda_1} = 0$ *for any* $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$.

Proof. We consider the LQ optimal control problem on $T_{\widehat{q}_1}M$ given by

$$\min_{\delta e \in \boldsymbol{\mathcal{V}}} \boldsymbol{J}''[\delta e]^2. \tag{6.1}$$

The maximised Hamiltonian associated by PMP with this LQ problem is

$$H_t''(\delta p, \delta z) = \frac{1}{2R(t)}\left(\langle \delta p, \dot{g}_t^1(\widehat{q}_1)\rangle - L_{\delta z}A(t, \widehat{q}_1)\right)^2 \qquad (\delta p, \delta z) \in T_{\widehat{q}_1}^*M \times T_{\widehat{q}_1}M.$$

Since $\delta e \in \boldsymbol{\mathcal{V}}$, δz is free so that PMP applied to problem (6.1) gives the following transversality conditions at the initial point

$$(\delta p, \delta z) \in L_{\widehat{\tau}_1}'' := \{(\delta p, \delta z) \colon \delta z \in T_{\widehat{q}_1}M, \delta p = -\mathrm{D}^2\big(\theta + \alpha + \widehat{\gamma}_T + \widehat{S}_T^0\big)[\delta z, \cdot]\}.$$

Denote with $\mathcal{H}_t''$ the Hamiltonian flow of H_t''. In order to compare $\mathcal{H}_t''$ with $\widehat{\mathcal{H}}_t$, we define the anti-symplectic isomorphism $\iota \colon T_{\widehat{q}_1}^*M \times T_{\widehat{q}_1}M \to T_{\widehat{\ell}_1}(T^*M)$ as

$$\iota(\delta p, \delta x) = -\delta p + \mathrm{d}\big(-\widehat{S}_T^0 - \widehat{\gamma}_T\big)_*\delta x.$$
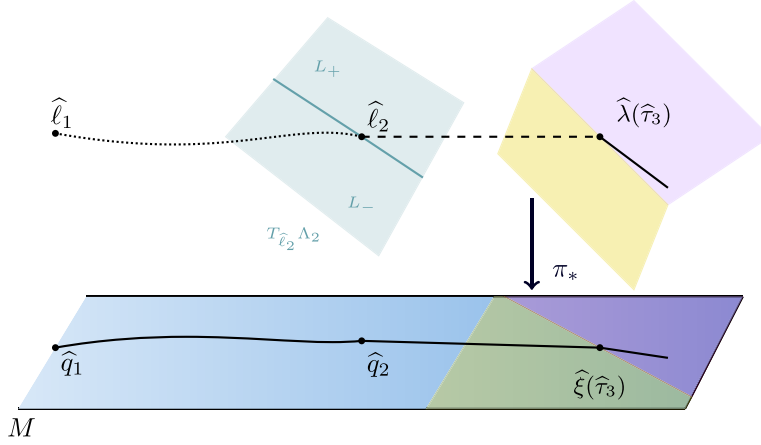
FIGURE 2. Invertibility at $t = \widehat{\tau}_3$. L_+ (respectively, L_-) denotes the half space of all $\delta\ell_2 \in T_{\widehat{\ell}_2}\Lambda_2$ such that $\langle \mathrm{d}\widehat{\tau}_3(\widehat{\ell}_2), \delta\ell_2\rangle \geq 0$ (respectively, ≤ 0). In purple, $\mathcal{H}_{\widehat{\tau}_3*}L_+$; in yellow, $\mathcal{H}_{\widehat{\tau}_3*}L_-$. These semi-planes project without intersections on $T_{\widehat{\xi}(\widehat{\tau}_3)}M$.

By definition, $\boldsymbol{\sigma} \circ \iota \otimes \iota = -\widehat{\sigma}$, where $\widehat{\sigma}$ denotes the standard symplectic structure on $T^*_{\widehat{q}_1}M \times T_{\widehat{q}_1}M$. It is immediate to verify that

$$\iota L''_{\widehat{\tau}_1} = T_{\widehat{\ell}_1}\Lambda_1.$$

Moreover, by analogous computations to those in [11, 23], it is easy to prove that

$$\mathcal{H}''_t = \iota^{-1}\widehat{\mathcal{F}}^{-1}_{t*}\widehat{\mathcal{H}}_{t*}\iota \quad t \in [\widehat{\tau}_1, \widehat{\tau}_2], \tag{6.2}$$

which implies that

$$\left(\pi\mathcal{H}''_t\right)^{-1} = \iota^{-1}\left(\pi\widehat{\mathcal{H}}_t\right)^{-1}_*\widehat{S}_{t*}. \tag{6.3}$$

On the other hand, ([29], Thm. 2.6) states that $\boldsymbol{J''}|_{\boldsymbol{\nu}}$ is coercive if and only if $\pi\mathcal{H}''_t \colon L''_{\widehat{\tau}_1} \to T_{\widehat{q}_1}M$ is one to one for every $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, so that the coerciveness of $\boldsymbol{J''}|_{\boldsymbol{\nu}}$ implies that $\left(\pi\widehat{\mathcal{H}}_t\right)_*|_{T_{\widehat{\ell}_1}\Lambda_1}$ is invertible for every $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$. $\qquad\qquad\square$

Coupling Lemma 6.3 with Lemma 6.4, we obtain that, if $\boldsymbol{J''}|_{\boldsymbol{\nu}}$ is coercive, then $(\pi\mathcal{K}_t)|_{\Lambda_1}$ is invertible for $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$.

Remark 6.5. Set $\Lambda_2 := \mathcal{K}_{\widehat{\tau}_2}(\Lambda_1)$. Since Λ_2 is a Lagrangian submanifold of T^*M which projects one to one onto a neighbourhood of M, there exist a neighbourhood $U_{\widehat{q}_2}$ of $\widehat{q}_2$ and a smooth function $\alpha_2 \colon U_{\widehat{q}_2} \to \mathbb{R}$ such that

$$\mathrm{d}\alpha_2(\widehat{q}_2) = \widehat{\ell}_2, \qquad \Lambda_2 := \mathcal{K}_{\widehat{\tau}_2}(\Lambda_1) = \{\mathrm{d}\alpha_2(q) \colon q \in U_{\widehat{q}_2}\}.$$

Invertibility for $t \in [\widehat{\tau}_2, T]$. Thanks to equation (5.6), we obtain that, if $\boldsymbol{J''}|_{\boldsymbol{\nu}}$ is coercive, then $\pi\mathcal{H}_t|_{\Lambda_1}$ is locally invertible for every $t < \widehat{\tau}_3$.

Let $\delta\ell \in T_{\widehat{\ell}_1}\Lambda_1$, and set $\delta\ell_2 = \mathcal{K}_{\widehat{\tau}_2*}\delta\ell$. The first order approximation of $\pi\mathcal{H}_{\widehat{\tau}_3}$ at $\widehat{\ell}_1$, applied to $\delta\ell$, is given by

$$\begin{cases} \exp\left((\widehat{\tau}_3 - \widehat{\tau}_2)f_0\right)_* \pi_*\delta\ell_2 & \text{if } \langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \geq 0, \\ \langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle f_1(\widehat{q}_3) + \exp\left((\widehat{\tau}_3 - \widehat{\tau}_2)f_0\right)_* \pi_*\delta\ell_2 & \text{if } \langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \leq 0, \end{cases}$$

which, up to a pullback, can be written as

$$\begin{cases} \pi_*\delta\ell_2 & \text{if } \langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \geq 0, \\ \pi_*\delta\ell_2 - \langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \widetilde{k}(\widehat{q}_2) & \text{if } \langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \leq 0, \end{cases}$$

where $\widetilde{k} := \widehat{S}_{\widehat{\tau}_2*}k$ is the pullback of $-f_1$ from time $\widehat{\tau}_3$ to time $\widehat{\tau}_2$. Notice that, by Assumption 3.1, $f_1(\widehat{q}_3) \neq 0$, so that $\widetilde{k}(\widehat{q}_2) \neq 0$.

By Clarke's inverse function theorem [13], $\pi\mathcal{H}_{\widehat{\tau}_3}$ is invertible if for every $a \in [0,1]$ and for every $\delta\ell_2 \in T_{\widehat{\ell}_2}\Lambda_2$ it holds

$$\pi_*\delta\ell_2 - a\langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \widetilde{k}(\widehat{q}_2) \neq 0. \tag{6.4}$$

By contradiction, assume there exist some $a \in [0,1]$ and $\delta\ell_2 \in T_{\widehat{\ell}_2}\Lambda_2$, $\delta\ell_2 \neq 0$, such that the left hand side of (6.4) is zero. This implies that $\pi_*\delta\ell_2 = \rho\widetilde{k}(\widehat{q}_2)$, for some $\rho \neq 0$, so that

$$\rho\widetilde{k}(\widehat{q}_2) - a\langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \delta\ell_2 \rangle \widetilde{k}(\widehat{q}_2) = 0,$$

which yields

$$1 - a\langle \mathrm{d}\tau_3(\widehat{\ell}_2), \, \mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2) \rangle = 0,$$

since $\delta\ell_2 = \rho\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2)$. Equivalently, using (5.8),

$$\mathfrak{r}_3 + a\sigma_{\widehat{\ell}_3}\left(\exp((\widehat{\tau}_3 - \widehat{\tau}_2)\vec{F}_0)_*\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2), \vec{\Phi}^+ \right) = 0. \tag{6.5}$$

By Assumption 3.8, the left hand side of (6.5) is negative for $a = 0$. If we show that it is negative also for $a = 1$, then, by linearity, we get a contradiction and we are done. The last part of this section is devoted to prove that this is a consequence of the coerciveness of the second variation on $\mathcal{W}$.

We denote with the symbol $J^{\flat}$ the bilinear form associated with J''. We recall that J'' is coercive on $\mathcal{W}$ if and only if it is coercive both on $\mathcal{V}$ and on $\mathcal{W} \cap \mathcal{V}^{\perp}$, where $\mathcal{V}^{\perp}$ denotes the orthogonal complement to $\mathcal{V}$ with respect to $J^{\flat}$. In order to compute $\mathcal{V}^{\perp}$, we introduce, for every $\delta e = (\delta z, \varepsilon, w) \in \mathcal{W}$, the trajectory $p \colon [\widehat{\tau}_1, \widehat{\tau}_2] \to T_{\widehat{q}_1}^* M$ solution of the Cauchy problem

$$\begin{cases} \dot{p}(t) = -w(t)L.A(t, \widehat{q}_1), \\ p(\widehat{\tau}_1) = -\mathrm{D}^2(\alpha + \theta + \widehat{\gamma}_T + \widehat{S}_T^0)(\widehat{q}_1)[\delta z, \cdot]. \end{cases}$$

Let $\widetilde{\delta e} = (\widetilde{\delta z}, \widetilde{\varepsilon}, \widetilde{w}) \in \mathcal{W}$ be another admissible variation; then $\boldsymbol{J}^{b}[\delta e, \widetilde{\delta e}]$ can be written as

$$
\begin{aligned}
\boldsymbol{J}^{b}[\delta e, \widetilde{\delta e}] &= \frac{1}{2}\mathrm{D}^2\big(\theta + \alpha + \widehat{\gamma}_T + \widehat{S}_T^0\big)(\widehat{q}_1)[\delta z, \widetilde{\delta z}] \\
&\quad + \frac{1}{2}\int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \big(w^2(t)\widetilde{w}(t)R(t) + w(t)L_{\widetilde{\zeta}(t)}A(t,\widehat{q}_1) + \widetilde{w}(t)L_{\zeta(t)}A(t,\widehat{q}_1)\big)\mathrm{d}t \\
&\quad + \frac{\varepsilon\widetilde{\varepsilon}}{2}\Big(L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big) - L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1)\Big) \\
&= \frac{1}{2}\langle\mathrm{D}^2\big(\theta + \alpha + \widehat{\gamma}_T + \widehat{S}_T^0\big)(\widehat{q}_1)[\delta z, \cdot] + p(\widehat{\tau}_1), \widetilde{\delta z}\rangle - \frac{1}{2}\langle p(\widehat{\tau}_2), \widetilde{\zeta}(\widehat{\tau}_2)\rangle \\
&\quad + \frac{1}{2}\int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \widetilde{w}(t)\Big(w(t)R(t) + L_{\zeta(t)}A(t,\widehat{q}_1) + \langle p(t), \dot{g}_t^1\rangle\Big)\mathrm{d}t \\
&\quad + \frac{\varepsilon\widetilde{\varepsilon}}{2}\Big(L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big) - L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1)\Big).
\end{aligned}
$$

From the expression here above, we see that $\delta e \in \mathcal{V}^{\perp}$ if and only if

$$
w(t)R(t) + L_{\zeta(t)}A(t,\widehat{q}_1) + \langle p(t), \dot{g}_t^1\rangle = 0 \qquad \forall t \in [\widehat{\tau}_1, \widehat{\tau}_2], \tag{6.6}
$$

so that, for any $\delta e \in \mathcal{V}^{\perp}$, we get that

$$
\boldsymbol{J}''[\delta e]^2 = -\frac{1}{2}\langle p(\widehat{\tau}_2), \zeta(\widehat{\tau}_2)\rangle + \frac{\varepsilon^2}{2}\Big(L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big) - L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1)\Big).
$$

Remark 6.6. We stress that the solution $(p(t), \zeta(t))$ of the Cauchy problem

$$
\begin{cases}
\dot{\zeta}(t) = w(t)\dot{g}_t^1(\widehat{q}_1), \\
\dot{p}(t) = -w(t)L_{(\cdot)}A(t,\widehat{q}_1), \\
(p(\widehat{\tau}_1), \zeta(\widehat{\tau}_1)) \in L_{\widehat{\tau}_1}'',
\end{cases}
$$

associated with the control $w(\cdot)$ that satisfies equation (6.6) is also the solution of the Hamiltonian system associated with H_t'' with the same initial condition, that is

$$
(p(t), \zeta(t)) = \mathcal{H}_t''(p(\widehat{\tau}_1), \zeta(\widehat{\tau}_1)).
$$

Remark that $\mathcal{V}^{\perp}$ is a 1-dimensional linear space: indeed, for every ε, $\zeta(\tau_2) = -\varepsilon k(\widehat{q}_1)$, so that δz is uniquely determined as the backward solution of the linear system with control (6.6). By homogeneity, we can choose $\delta e \in \mathcal{V}^{\perp}$ with $\varepsilon = -1$. Then $\boldsymbol{J}''|_{\mathcal{V}^{\perp}}$ is coercive if and only if

$$
\begin{aligned}
0 &> \langle p(\widehat{\tau}_2), \zeta(\widehat{\tau}_2)\rangle - L_{[k_4,k_3]}\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1) + L_{k_4}\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) + L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1) \\
&= \langle p(\widehat{\tau}_2), k(\widehat{q}_1)\rangle + \mathfrak{r}_3 + L_k|\widehat{\psi}_{\widehat{\tau}_3}|(\widehat{q}_1) + L_k^2\big(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0\big)(\widehat{q}_1),
\end{aligned} \tag{6.7}
$$

where the expression (6.7) is obtained applying (A.1) and the definition of $\widehat{\gamma}_T$.

We now compute $\langle p(\widehat{\tau}_2), k(\widehat{q}_1)\rangle$ in terms of Hamiltonian flows. Consider the pair $(\delta p, \delta z) \in L_{\widehat{\tau}_1}''$ such that $\pi\mathcal{H}_{\widehat{\tau}_2}''(\delta p, \delta z) = k(\widehat{q}_1)$ (thanks to the invertibility of $\pi\mathcal{H}_{\widehat{\tau}_2}''$, it exists and it is unique), so that, by (6.3), we get

$(\delta p, \delta z) = \iota^{-1}(\pi\widehat{\mathcal{H}}_{\widehat{\tau}_2})_*^{-1}\widehat{S}_{\widehat{\tau}_2 *}k(\widehat{q}_1)$. Thanks to (6.2) we obtain

$$
\begin{aligned}
\langle p(\widehat{\tau}_2)\,,\,k(\widehat{q}_1)\rangle &= \widehat{\sigma}\big(\mathcal{H}''_{\widehat{\tau}_2}(\pi\mathcal{H}''_{\widehat{\tau}_2})^{-1}k(\widehat{q}_1),(0,k(\widehat{q}_1))\big) \\
&= -\boldsymbol{\sigma}_{\widehat{\ell}_1}\Big(\widehat{\mathcal{F}}_{\widehat{\tau}_2 *}^{-1}\widehat{\mathcal{H}}_{\widehat{\tau}_2 *}\iota(\pi\mathcal{H}''_{\widehat{\tau}_2})^{-1}k(\widehat{q}_1),\mathrm{d}\big(-\widehat{\gamma}_T-\widehat{S}^0_T\big)_*k(\widehat{q}_1)\Big) \\
&= -\boldsymbol{\sigma}_{\widehat{\ell}_1}\Big(\widehat{\mathcal{F}}_{\widehat{\tau}_2 *}^{-1}\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2),\mathrm{d}\big(-\widehat{\gamma}_T-\widehat{S}^0_T\big)_*k(\widehat{q}_1)\Big),
\end{aligned}
\tag{6.8}
$$

where α_2 is the function defined in Remark 6.5. Indeed, thanks to (6.3), we obtain that $\iota(\pi\mathcal{H}''_{\widehat{\tau}_2})^{-1}k(\widehat{q}_1) = (\pi\widehat{\mathcal{H}}_{\widehat{\tau}_2})_*^{-1}\widetilde{k}(\widehat{q}_2)$, so that $\iota(\pi\mathcal{H}''_{\widehat{\tau}_2})^{-1}k(\widehat{q}_1) \in T_{\widehat{\ell}_1}\Lambda_1$ and

$$
\widehat{\mathcal{H}}_{\widehat{\tau}_2 *}\iota(\pi\mathcal{H}''_{\widehat{\tau}_2})^{-1}k(\widehat{q}_1) \in \widehat{\mathcal{H}}_{\widehat{\tau}_2 *}(T_{\widehat{\ell}_1}\Lambda_1) = \mathcal{K}_{\widehat{\tau}_2 *}(T_{\widehat{\ell}_1}\Lambda_1) = T_{\widehat{\ell}_2}\Lambda_2.
$$

By cumbersome but standard computations, $\widehat{\mathcal{F}}_{\widehat{\tau}_2 *}\mathrm{d}\big(-\widehat{\gamma}_T-\widehat{S}^0_T\big)_*k(\widehat{q}_1) = \mathrm{d}\big((-\widehat{\gamma}_T-\widehat{S}^0_T+\widehat{S}^0_{\widehat{\tau}_2})\circ\widehat{S}^{-1}_{\tau_2}\big)_*\widetilde{k}(\widehat{q}_2)$. Equation (6.8) thus gives

$$
\begin{aligned}
\langle p(\widehat{\tau}_2)\,,\,k(\widehat{q}_1)\rangle &= -\boldsymbol{\sigma}_{\widehat{\ell}_2}\big(\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2),\mathrm{d}\big((-\widehat{\gamma}_T-\widehat{S}^0_T+\widehat{S}^0_{\widehat{\tau}_2})\circ\widehat{S}^{-1}_{\tau_2}\big)_*\widetilde{k}(\widehat{q}_2)\big) \\
&= -\mathrm{D}^2\big(\alpha_2+(\widehat{\gamma}_T+\widehat{S}^0_T-\widehat{S}^0_{\widehat{\tau}_2})\circ\widehat{S}^{-1}_{\tau_2}\big)[\widetilde{k}(\widehat{q}_2)]^2.
\end{aligned}
$$

Finally, computing the value of $\widehat{\ell}_2$ by means of (A.1), we get

$$
\begin{aligned}
&\langle p(\widehat{\tau}_2),k(\widehat{q}_1)\rangle + L_k^2(\widehat{\gamma}_T+\widehat{S}^0_T-\widehat{S}^0_{\widehat{\tau}_2})(\widehat{q}_1) \\
&= -\mathrm{D}^2\big(\alpha_2+(\widehat{\gamma}_T+\widehat{S}^0_T-\widehat{S}^0_{\widehat{\tau}_2})\circ\widehat{S}^{-1}_{\widehat{\tau}_2}\big)[\widetilde{k}(\widehat{q}_2)]^2 + L_{\widetilde{k}}^2\big((\widehat{\gamma}_T+\widehat{S}^0_T-\widehat{S}^0_{\widehat{\tau}_2})\circ\widehat{S}^{-1}_{\widehat{\tau}_2}\big)(\widehat{q}_2) \\
&= -\mathrm{D}^2\alpha_2[\widetilde{k}(\widehat{q}_2)]^2 - \widehat{\ell}_2\mathrm{D}\widetilde{k}(\widehat{q}_2)\widetilde{k}(\widehat{q}_2) = -L_{\widetilde{k}}^2\alpha_2(\widehat{q}_2).
\end{aligned}
$$

Plugging this equality in equation (6.7), we obtain that

$$
\mathfrak{r}_3 + L_k|\widehat{\psi}_{\widehat{\tau}_3}|(\widehat{q}_1) - L_{\widetilde{k}}^2\alpha_2(\widehat{q}_2) < 0.
$$

It now suffices to notice that

$$
\begin{aligned}
\sigma_{\widehat{\ell}_3}\big(\exp((\widehat{\tau}_3-\widehat{\tau}_2)\vec{F}_0)_*\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2),\vec{\Phi}^+\big) &= \sigma_{\widehat{\ell}_2}\Big(\mathrm{d}\alpha_{2*}\widetilde{k}(\widehat{q}_2),\exp(-(\widehat{\tau}_3-\widehat{\tau}_2)\vec{F}_0)_*\vec{\Phi}^+\Big) \\
&= -L_{\widetilde{k}}^2\alpha_2(\widehat{q}_2) + L_k\widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1),
\end{aligned}
$$

to obtain that equation (6.5) holds true for $a = 1$, and we are done.

7. Main results

We now state the main result of the paper. This section is devoted to its proof.

Theorem 7.1. *Let $\widehat{\xi}$ be an admissible trajectory for system* (1.2) *and satisfying Assumptions 3.1–4.3. Then $\widehat{\xi}$ is a strict strong local minimiser of problem* (1.1)–(1.2).

The proof of the strong local optimality of the reference trajectory follows the same lines of ([10], Thm. 4.1) (see also [30]), thus we are just recalling the main arguments. We shall instead provide all details for the proof of the strictness part.

Proof. We define on $\mathbb{R} \times T^*M$ the one-form $\omega = \mathcal{H}_t^* \varsigma - H_t \circ \mathcal{H}_t \mathrm{d}t$. Applying ([30], Lem. 3.3) we can prove that ω is exact on $[0,T] \times \Lambda_1$.

Let $\mathbf{O} \subset \mathbb{R} \times M$ be a neighbourhood of the graph of $\widehat{\xi}$ such that $\pi\mathcal{H} \colon [0,T] \times \Lambda_1 \to \mathbf{O}$ is invertible, with piecewise C^1 inverse. Consider an admissible trajectory $\xi \colon [0,T] \to M$ of (1.2) whose graph is contained in $\mathbf{O}$, and call $v(t)$ its associated control; set

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

We define the closed path $\gamma \colon [0,2T] \to [0,T] \times \Lambda_1$ as

$$\gamma(t) = \begin{cases} (t, \ell(t)) & t \in [0,T], \\ (2T - t, \widehat{\ell}_1) & t \in [T, 2T]. \end{cases}$$

Integrating ω along γ, and recalling that H_t is an over-maximised Hamiltonian, we obtain that

$$\int_0^T |\widehat{u}(t)\psi(\widehat{\xi}(t))|\mathrm{d}t \leq \int_0^T |v(t)\psi(\xi(t))|\mathrm{d}t, \tag{7.1}$$

that is, $\widehat{\xi}$ is a strong local minimiser.

Assume now that $\widehat{\xi}$ is not a strict minimiser, that is, there exists an admissible trajectory ξ with graph in $\mathbf{O}$ for which equality holds in equation (7.1); this is equivalent to

$$\langle \lambda(t), (f_0 + v(t)f_1)(\xi(t)) \rangle - |v(t)\psi(\xi(t))| = H_t(\lambda(t)) \qquad \text{a.e. } t \in [0,T], \tag{7.2}$$

that is, both $\langle \lambda(t), (f_0 + v(t)f_1)(\xi(t)) \rangle - |v(t)\psi(\xi(t))|$ and $H_t(\lambda(t))$ coincide with $H_{\max}(\lambda(t))$. Since $\xi(0) = \widehat{\xi}(0)$, and by regularity of the first bang arc, then, for t small enough, $h(\lambda(t), u)$ attains its maximum only for $u = \widehat{u}(t) = 1$. This implies that $v(t) = \widehat{u}(t)$ and $\xi(t) = \widehat{\xi}(t)$ as long as $\Phi^-(\lambda(t)) > 0$, that is, for $t \in [0, \widehat{\tau}_1)$.

Analogously, since $\xi(T) = \widehat{\xi}(T)$, we can apply, backward in time, an analogous argument, and prove that $u(t) \equiv \widehat{u}(t)$ for $t \in (\widehat{\tau}_2, T]$ (see [10] for more details).

For $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, equation (7.2) implies that $\lambda(t) \in S^-$. Thus, since $\vec{K}$ is tangent to S^- (see Prop. 5.6), then $\ell(t) \in S^-$ too. We claim that there exists a function $a \colon [\widehat{\tau}_1, \widehat{\tau}_2] \to \mathbb{R}$ such that

$$\dot{\ell}(t) = -a(t)\vec{\Phi}^-(\widehat{\ell}_1). \tag{7.3}$$

If so, since $\dot{\ell}(t)$ is tangent to S^- and $\vec{\Phi}^-$ is transverse to S^-, we obtain that $a(t) \equiv 0$ and $\dot{\ell}(t) \equiv 0$ for $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, that is $\xi(t) = \pi\mathcal{K}_t(\widehat{\ell}_1) = \widehat{\xi}(t)$, which completes the proof.

In order to prove (7.3), we first observe that

$$\dot{\xi}(t) = \pi_*\dot{\lambda}(t) = \pi_*\big(\vec{K}(\lambda(t)) + \mathcal{K}_{t*}\dot{\ell}(t)\big). \tag{7.4}$$

On the other hand, it is not difficult to see that, for every $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, the function

$$\Sigma^- \cap T^*_{\xi(t)}M \ni \ell \mapsto K(\ell) - \langle \ell, f_0(\xi(t)) + v(t)f_1(\xi(t)) \rangle + |v(t)\psi(\xi(t))|$$

has a minimum at $\ell = \lambda(t)$; thus, its derivative along variations in $\Sigma^- \cap T^*_{\xi(t)}M$ (that is, δp with $\langle \delta p, f_1(\xi(t)) \rangle = 0$) must be zero. This means that the derivative with respect to the vertical coordinates (*i.e.*, the directions

contained in $T^*_{\xi(t)}M$) must be parallel to $f_1(\xi(t))$, which implies that

$$\pi_* \vec{K}(\lambda(t)) = f_0(\xi(t)) + v(t)f_1(\xi(t)) + a(t)f_1(\xi(t)) = \dot{\xi}(t) + a(t)f_1(\xi(t)), \tag{7.5}$$

for some real function $a(\cdot)$. Combining (7.4) with (7.5), we obtain that

$$(\pi \mathcal{K}_t)_* \dot{\ell}(t) + a(t)f_1(\xi(t)) = 0.$$

Since $(\pi \mathcal{K}_t)^{-1}_* f_1(\xi(t)) = \vec{\Phi}^-(\widehat{\ell}_1)$, applying $(\pi \mathcal{K}_t)^{-1}_*$ to the equality here above we get the claim. $\qquad\square$

8. Example

In this section, we apply our result to the optimal control problem in $\mathbb{R}^2$

$$\min_{|u(\cdot)| \leq 1} \int_0^T |u(t)x_2(t)|\, \mathrm{d}t$$

subject to the control system

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u - \rho x_2, \\ x_1(0) = 0, \ x_2(0) = 0, \\ x_1(T) = X > 0, \ x_2(T) = 0, \end{cases}$$

where T, ρ and X are given positive constants. This problem, studied in [6], models the fuel consumption minimisation problem for an academic electric vehicle moving in one horizontal direction with friction. The authors prove that, if the final time T is larger than

$$T_{\lim} = \frac{1}{\rho} \log \left((1 + \sqrt{2})e^{\rho^2 X} - 1 + \sqrt{(1 + \sqrt{2})e^{\rho^2 X} - 1)^2 - 1} \right),$$

then the optimal control has the bang-singular-inactivated-bang structure described in equation (3.1). The corresponding trajectories satisfy PMP in normal form, with adjoint covector $\boldsymbol{p}(t) = (p_1(t), p_2(t))$, with $p_1(t) \equiv p_1^0$ for every $t \in [0, T]$. In particular, the following relations hold

$$\widehat{\tau}_1 = \frac{1}{\rho} \log \left(\frac{2}{2 - p_1^0} \right), \quad \widehat{\tau}_3 - \widehat{\tau}_2 = \frac{1}{\rho} \log \left(1 + \sqrt{2} \right), \quad T - \widehat{\tau}_3 = \frac{1}{\rho} \log \left(\frac{p_1^0 + 2(1 + \sqrt{2})}{2(1 + \sqrt{2})} \right), \tag{8.1}$$

and

$$\widehat{u}_S(t) \equiv \sqrt{\rho p_2(0)} = \frac{p_1^0}{2} = \rho p_2(\widehat{\tau}_1) \qquad \forall t \in [\widehat{\tau}_1, \widehat{\tau}_2]. \tag{8.2}$$

Coupling equations (8.1) and (8.2), we deduce that p_1^0 must be positive and smaller than 2. Finally, the authors prove that along these extremals $x_2(t)$ is positive for every $t \in (0, T)$ and

$$x_2(t) \equiv p_2(t) = \frac{p_1^0}{2\rho} \qquad \forall t \in [\widehat{\tau}_1, \widehat{\tau}_2].$$

Using the notations of our paper, the drift f_0, the controlled vector field f_1 and the cost ψ at a point $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ are respectively given by

$$f_0(\boldsymbol{x}) = \begin{pmatrix} x_2 \\ -\rho x_2 \end{pmatrix}, \qquad f_1(\boldsymbol{x}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \psi(\boldsymbol{x}) = x_2. \tag{8.3}$$

Thus, the associated Hamiltonian functions have the following expressions

$$F_0(\boldsymbol{p}, \boldsymbol{x}) = p_1 x_2 - \rho p_2 x_2, \qquad F_1(\boldsymbol{p}, \boldsymbol{x}) = p_2, \qquad \Phi^{\pm}(\boldsymbol{p}, \boldsymbol{x}) = p_2 \pm |x_2|.$$

Here below, we prove that Assumptions 3.1–4.3 are met.

8.1. Regularity assumptions

Since $x_2(t)$ never vanishes in $(0, T)$, then Assumptions 3.1–3.2 and 3.10 are trivially satisfied. To verify that also Assumption 3.5 holds true, we are left to prove that the bang and inactivated arcs are regular. In order to complete this task, we compute the iterated Lie brackets of the vector fields f_0 and f_1:

$$f_{01}(\boldsymbol{x}) \equiv \begin{pmatrix} -1 \\ \rho \end{pmatrix}, \qquad f_{101}(\boldsymbol{x}) \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \mathrm{ad}_{f_0}^k f_{01}(\boldsymbol{x}) = \rho^k f_{01}(\boldsymbol{x}) \equiv \begin{pmatrix} -\rho^k \\ \rho^{k+1} \end{pmatrix}. \tag{8.4}$$

Regularity of the first bang arc. We have to prove that $\Phi^-(\boldsymbol{p}(t), \boldsymbol{x}(t)) > 0$ for $t \in [0, \widehat{\tau}_1)$. The claim follows directly from the computations

$$\Phi^-(\boldsymbol{p}(t), \boldsymbol{x}(t)) = \frac{e^{-\rho t}}{4\rho} \left(e^{\rho t}(2 - p_1^0) - 2 \right)^2,$$

and equation (8.1).

Regularity along the inactivated arc. We must verify that $p_2(t) - x_2(t) < 0 < p_2(t) + x_2(t)$ for $t \in (\widehat{\tau}_2, \widehat{\tau}_3)$. By computations

$$p_2(t) \pm x_2(t) = \frac{p_1^0}{2\rho} \left(-e^{\rho(t-\widehat{\tau}_2)} + 2 \pm e^{-\rho(t-\widehat{\tau}_2)} \right).$$

The claim follows from equation (8.1).

Regularity along the last bang arc. A straightforward computation gives $p_2(t) + x_2(t) \leq 0$ for $t \in [\widehat{\tau}_3, T]$, where the equality holds if and only if $t = \widehat{\tau}_3$.

Regularity at the switching points (Asm. 3.8). At the first switching time $t = \widehat{\tau}_1$, Assumption 3.8 reads

$$0 < \left(F_{001} + F_{101} + L_{f_{01}}\psi - L_{f_0+f_1}L_{f_0}\psi \right)(\widehat{\ell}_1) = -\rho p_1^0 + \rho^2 p_2(\widehat{\tau}_1) + \rho + \rho(1 - \rho x_2(\widehat{\tau}_1)) = \rho(2 - p_1^0),$$

which is verified thanks to equation (8.1).
At $t = \widehat{\tau}_2$, the regularity condition is

$$0 > \left(F_{001} - L_{f_0}^2 \psi \right)(\widehat{\ell}_2) = -\rho p_1^0 + \rho^2 p_2(\widehat{\tau}_2) - \rho^2 x_2(\widehat{\tau}_2) = -\rho p_1^0,$$

which is trivially satisfied.
At $t = \widehat{\tau}_3$, the regularity condition reads

$$0 > \left(F_{01} - L_{f_0}\psi \right)(\widehat{\ell}_3) = -p_1^0 + \rho p_2(\widehat{\tau}_3) + \rho x_2(\widehat{\tau}_3) = -p_1^0,$$

which is verified.

Assumption 3.9 (SGLC). This Assumption is trivially satisfied, since

$$\mathbb{L}(\boldsymbol{p}, \boldsymbol{x}) = F_{101}(\boldsymbol{p}, \boldsymbol{x}) + L_{f_{01}}\psi(\boldsymbol{x}) - L_{f_1}L_{f_0}\psi(\boldsymbol{x})$$
$$= \rho - L_{f_1}(-\rho x_2) = 2\rho > 0 \qquad \forall (\boldsymbol{p}, \boldsymbol{x}) \in T^*\mathbb{R}^2.$$

8.2. Second variation

First of all, we compute the pull-back vector fields, by means of formula (C.1) and of equations (8.3)–(8.4). We obtain the following expressions:

$$g_t^1 \equiv f_1 + \frac{e^{\rho(t-\widehat{\tau}_1)} - 1}{\rho}f_{01}, \qquad \dot{g}_t^1 \equiv e^{\rho(t-\widehat{\tau}_1)}f_{01}, \qquad t \in [\widehat{\tau}_1, \widehat{\tau}_2],$$

$$k_3 = f_0 - \widehat{u}_S\frac{e^{\rho(\widehat{\tau}_2-\widehat{\tau}_1)} - 1}{\rho}f_{01}, \qquad k = -f_1 - \frac{e^{\rho(\widehat{\tau}_3-\widehat{\tau}_1)} - 1}{\rho}f_{01} = \begin{pmatrix} \frac{e^{\rho(\widehat{\tau}_3-\widehat{\tau}_1)}-1}{\rho} \\ -e^{\rho(\widehat{\tau}_3-\widehat{\tau}_1)} \end{pmatrix}.$$

Admissible variations The pullback system (4.5) assumes the form

$$\begin{cases} \zeta_1(t) = -\int_{\widehat{\tau}_1}^t w(s)e^{\rho(s-\widehat{\tau}_1)}\mathrm{d}s, \\ \zeta_2(t) = \varepsilon_0 + \rho\int_{\widehat{\tau}_1}^t w(s)e^{\rho(s-\widehat{\tau}_1)}\mathrm{d}s, \end{cases} \tag{8.5}$$

so that the space of extended admissible variation $\mathcal{W}$ can be identified with the following subspace of $\mathbb{R} \times L^2([\widehat{\tau}_1, \widehat{\tau}_2], \mathbb{R})$:

$$\mathcal{W} = \left\{ (\varepsilon, w): \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} w(s)e^{\rho(s-\widehat{\tau}_1)}\mathrm{d}s = \varepsilon\frac{e^{\rho(\widehat{\tau}_3-\widehat{\tau}_1)} - 1}{\rho} \right\}.$$

It is easy to see that $\widehat{S}_t^0$ is linear with respect to the state; since both $\dot{g}_t^1$ and k are constant with respect to the basepoint, we choose $\widehat{\gamma}_T$ as a linear function of the state, so that all second derivatives of the term $\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0$ are zero. In particular, with this choice we obtain $\mathrm{D}^2(\alpha + \widehat{\gamma}_T + \widehat{S}_T^0)[\varepsilon f_1]^2 = \varepsilon^2$.

Keeping all these facts into account, the second variation reads

$$J_e''[(\varepsilon, w)]^2 = \varepsilon^2(1 + \sqrt{2}\widehat{u}_S) + \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \rho w^2(t) - \rho w(t)\zeta_2(t)e^{-\rho(t-\widehat{\tau}_1)}\mathrm{d}t. \tag{8.6}$$

8.2.1. Coerciveness of the second variation

We recall that, given any subspace $\mathcal{V} \subset \mathcal{W}$, J_e'' is coercive on $\mathcal{W}$ if and only if it is coercive both $\mathcal{V}$ and $\mathcal{V}^\perp$, $\mathcal{V}^\perp$ denoting the orthogonal complement of $\mathcal{V}$ in $\mathcal{W}$ with respect to the bilinear form associated with J_e''. To prove the coerciveness of (8.6), we choose

$$\mathcal{V} = \left\{ (0, w): \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} w(s)e^{\rho(s-\widehat{\tau}_1)}\mathrm{d}s = 0 \right\}.$$

Coerciveness on $\mathcal{V}$. In order to prove the claim, we apply the characterization of coerciveness given in ([2],

Lem. 5), that is, we study the LQ optimal control problem

$$\min \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \rho w^2(t) - \rho w(t) z_2(t) e^{-\rho(t-\widehat{\tau}_1)} \mathrm{d}t$$

subject to

$$\begin{cases} \dot{\boldsymbol{z}}(t) = w(t) e^{\rho(t-\widehat{\tau}_1)} f_{01}, \\ \boldsymbol{z}(\widehat{\tau}_1) = \boldsymbol{z}(\widehat{\tau}_2) = 0, \end{cases} \qquad w \in L^2([\widehat{\tau}_1, \widehat{\tau}_2]): \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} w(t) e^{\rho(t-\widehat{\tau}_1)} \mathrm{d}t = 0.$$

The maximised Hamiltonian G'' associated with this problem is given by

$$G''(\boldsymbol{p}, \boldsymbol{z}, t) = \frac{1}{4\rho} \big((\rho p_2 - p_1) e^{\rho(t-\widehat{\tau}_1)} + \rho z_2 e^{-\rho(t-\widehat{\tau}_1)} \big)^2. \tag{8.7}$$

The solutions of the Hamiltonian system associated with (8.7) are given by

$$\begin{cases} z_1(t) = \rho(a + b(t-\widehat{\tau}_1)) e^{\rho(t-\widehat{\tau}_1)}, \\ z_2(t) = (a + b(t-\widehat{\tau}_1)) e^{\rho(t-\widehat{\tau}_1)}, \\ p_1(t) \equiv p_1(\widehat{\tau}_1), \\ p_2(t) = \frac{p_1(\widehat{\tau}_1)}{\rho} + \big(\frac{2b+a}{\rho} + b(t-\widehat{\tau}_1) \big) e^{-\rho(t-\widehat{\tau}_1)}, \end{cases}$$

for some real constants a, b. The boundary conditions $\boldsymbol{z}(\widehat{\tau}_1) = \boldsymbol{z}(\widehat{\tau}_2) = 0$ are satisfied only for $a = b = 0$, that is $\boldsymbol{z}(t) \equiv 0$. Thanks to ([2], Lem. 5), this implies that J_e'' is coercive on $\mathcal{V}$.

Coerciveness on $\mathcal{V}^\perp$. First of all, we characterize $\mathcal{V}^\perp$. The bilinear form associated with J_e'' is given by the formula

$$J^b[\delta e, \widetilde{\delta e}] = \varepsilon \widetilde{\varepsilon} (1 + \sqrt{2} \widehat{u}_S) + \frac{1}{2} \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} 2\rho w(t) \widetilde{w}(t) - \rho \big(w(t) \widetilde{\zeta}_2(t) + \widetilde{w}(t) \zeta_2(t) \big) e^{-\rho(t-\widehat{\tau}_1)} \mathrm{d}t, \tag{8.8}$$

with $\delta e, \widetilde{\delta e} \in \mathcal{W}$. Introducing $p_2(\cdot)$ as the solution of the Cauchy problem

$$\begin{cases} \dot{p}_2(t) = w(t) e^{\rho(t-\widehat{\tau}_1)}, \\ p_2(\widehat{\tau}_1) = 0, \end{cases}$$

equation (8.8) becomes

$$J^b[\delta e, \widetilde{\delta e}] = \varepsilon \widetilde{\varepsilon} (1 + \sqrt{2} \widehat{u}_S) - \rho \widetilde{\zeta}_2(\widehat{\tau}_2) p_2(\widehat{\tau}_2) + \frac{1}{2} \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \rho \widetilde{w}(t) \Big(2w(t) - \zeta_2(t) e^{-\rho(t-\widehat{\tau}_1)} + p_2(t) e^{-\rho(t-\widehat{\tau}_1)} \Big) \mathrm{d}t. \tag{8.9}$$

It is immediate to see that an admissible variation δe belongs to $\mathcal{V}^\perp$ if and only if $e^{-\rho(t-\widehat{\tau}_1)} \Big(2w(t) - \zeta_2(t) e^{-\rho(t-\widehat{\tau}_1)} + p_2(t) e^{-\rho(t-\widehat{\tau}_1)} \Big)$ does not depend on t (indeed, the orthogonal complement to zero-mean functions in $L^2([\widehat{\tau}_1, \widehat{\tau}_2])$ is the space of constant functions); we thus set

$$C_\varepsilon = e^{-\rho(t-\widehat{\tau}_1)} \Big(2w(t) - \zeta_2(t) e^{-\rho(t-\widehat{\tau}_1)} + p_2(t) e^{-\rho(t-\widehat{\tau}_1)} \Big), \tag{8.10}$$

so that

$$J''[\delta e]^2|_{\mathcal{V}^\perp} = \varepsilon^2(1 + \sqrt{2}\widehat{u}_S) - \varepsilon\rho p_2(\widehat{\tau}_2)e^{(\widehat{\tau}_3-\widehat{\tau}_1)} + \frac{C_\varepsilon}{2}\varepsilon(e^{(\widehat{\tau}_3-\widehat{\tau}_1)} - 1). \qquad (8.11)$$

From equation (8.10), combined with (8.5), we can deduce that $w(\cdot)$ is smooth. Differentiating (8.10) with respect to t, we obtain

$$\dot{w}(t)e^{-\rho(t-\widehat{\tau}_1)} - \rho w(t)e^{-\rho(t-\widehat{\tau}_1)} + \rho\zeta_2(t)e^{-2\rho(t-\widehat{\tau}_1)} = 0. \qquad (8.12)$$

Multiplying (8.12) by $e^{2\rho(t-\widehat{\tau}_1)}$ and differentiating again, we obtain $\ddot{w}(t) = 0$, i.e. $w(\cdot)$ is an affine function. Plugging into the previous equations we obtain

$$C_\varepsilon = \varepsilon\frac{2\sqrt{2} + \rho(\widehat{\tau}_2 - \widehat{\tau}_1)}{\rho(\widehat{\tau}_2 - \widehat{\tau}_1)}$$

and

$$p_2(\widehat{\tau}_2) = \frac{\varepsilon}{\rho^2(\widehat{\tau}_2 - \widehat{\tau}_1)}\Big(-e^{-\rho(\widehat{\tau}_2-\widehat{\tau}_1)}\Big(2\sqrt{2} + (1 + \sqrt{2})\rho(\widehat{\tau}_2 - \widehat{\tau}_1)\Big) + 2\sqrt{2} + \rho(\widehat{\tau}_2 - \widehat{\tau}_1)\Big).$$

Substituting these expressions into (8.11), we obtain

$$J''[\delta e]^2|_{\mathcal{V}^\perp} = \varepsilon^2\left(2 + \sqrt{2} + \frac{2}{\rho(\widehat{\tau}_2 - \widehat{\tau}_1)} + \sqrt{2}\widehat{u}_S\right),$$

which is positive whenever $\varepsilon \neq 0$.

8.2.2. The over-maximised Hamiltonian

Although the expressions of the Hamiltonians H_0 and K are not necessary to deduce the optimality of the reference extremal, we provide their explicit formulas, to give an insight on their construction.

Since, along the extremal, $x_2 > 0$, we do not use the absolute value in the expressions of Φ^-. First of all, we see that $\Sigma^- = \{(\boldsymbol{p}, \boldsymbol{x}): p_2 - x_2 = 0\}$ and $S^- = \{(\boldsymbol{p}, \boldsymbol{x}): p_2 = x_2 \text{ and } p_1 = \rho(p_2 + x_2)\}$. Straightforward computations yield

$$\vartheta(\boldsymbol{p}, \boldsymbol{x}) = \frac{p_1}{2\rho} - \frac{x_2 + p_2}{2}, \qquad H_0(\boldsymbol{p}, \boldsymbol{x}) = \frac{1}{4\rho}\big(p_1 - \rho(p_2 - x_2)\big)^2.$$

In particular, we obtain that

$$H_0(\boldsymbol{p}, \boldsymbol{x}) - F_0(\boldsymbol{p}, \boldsymbol{x}) = \frac{1}{4\rho}\Big((p_1 - 2\rho x_2)^2 + 4\rho^2 x_2(p_2 - x_2)\Big).$$

This formula shows that $H_0 \geq F_0$ only for $(\boldsymbol{p}, \boldsymbol{x}) \in \Sigma^-$, with equality for $(\boldsymbol{p}, \boldsymbol{x}) \in S^-$. Finally $\nu(\boldsymbol{p}, \boldsymbol{x}) = \dfrac{p_1}{2}$ for any $(\boldsymbol{p}, \boldsymbol{x}) \in T^*\mathbb{R}^2$, so that

$$K(\boldsymbol{p}, \boldsymbol{x}) = \frac{1}{4\rho}\big(p_1 - \rho(p_2 - x_2)\big)^2 + \frac{p_1}{2}(p_2 - x_2).$$

9. Conclusions

In this paper, we develop the analysis of sufficient optimality conditions for generalised L^1-optimal control problems started in [10]. In particular, we consider the case in which the reference extremal contains a singular arc. As already observed in the Introduction, this fact brings significant technical difficulties, in particular the necessity of computing the second variation of a singular arc for a Bolza problem, which, to our knowledge, has not been done before.

We believe that [10] and the current paper, altogether, provide a solid basis for the study of sufficient optimality conditions for problems of the form (1.1)–(1.2): indeed, they give an insight of how to figure out more complex cases (as the concatenation of several bang, singular and inactivated extremals).

Two only issues are left over: the possibility of bang-bang concatenations (here prevented by Asm. 3.1), that, in our opinion, can be treated providing minor changes to the existing results, taking advantage of the techniques developed, for instance, in [25]; the case of degenerate singular arcs, which, even if it is a non-generic case, is nevertheless theoretically challenging. The authors are planning to consider this last situation.

Appendix A. Sketch of the computation of the second variation

We first recall that the reference extremal $\widehat{\lambda}$, associated with the control (3.1), satisfies the following equation

$$\widehat{\lambda}(t) = \left(\widehat{\ell}_1 + \int_{\widehat{\tau}_1}^t |u(s)| \ d|\widehat{\psi}_s|(\widehat{q}_1)\mathrm{d}s\right) \circ \widehat{S}_{t*}^{-1} = \left(\widehat{\ell}_1 + \mathrm{d}\widehat{S}_t^0(\widehat{q}_1)\right) \circ \widehat{S}_{t*}^{-1} \qquad \forall t \in [0, T]. \tag{A.1}$$

Moreover the cost realised by the trajectory $\boldsymbol{\xi}$ associated with the control u can be written as $J(u) = \widehat{S}_T^\circ(\boldsymbol{\eta}_T) - \widehat{S}_0^\circ(\boldsymbol{\eta}_0) + \widehat{\gamma}_T(\eta_T) + \widehat{\gamma}_0(\eta_0)$, thanks to equation (4.2). Further, we notice that the variations $\delta u = u - \widehat{u}$ can be encoded in three terms, that is, the variation δv of the control along the singular arc, and the variations $\varepsilon_3 = (\tau_3 - \widehat{\tau}_2) - (\widehat{\tau}_3 - \widehat{\tau}_2)$ and $\varepsilon_4 = (T - \tau_3) - (T - \widehat{\tau}_3)$ of the length of the third and fourth arc, related by the constraint $\varepsilon_3 + \varepsilon_4 = 0$. In particular, there is no variation of the terms $-\widehat{S}_0^\circ(\boldsymbol{\eta}_0) + \widehat{\gamma}_0(\eta_0)$. The second variation $J''[\delta u]$ is thus given by

$$\frac{\partial^2 J}{\partial u^2} = \frac{\partial^2 \eta_T^0}{\partial u^2} + \left(\mathrm{D}^2 \widehat{S}_T^0(\widehat{q}_0) + \mathrm{D}^2 \widehat{\gamma}_T(\widehat{q}_0)\right) \left[\frac{\partial \eta_T}{\partial u}\right]^2 + \langle \mathrm{d}\widehat{S}_T^0(\widehat{q}_0) + \mathrm{d}\widehat{\gamma}_T(\widehat{q}_0), \frac{\partial^2 \eta_T}{\partial u^2}\rangle$$

where, for the differential of $\widehat{S}_t^0$, we have used the notation established in Remark 4.1 and where the derivatives with respect to u have to be intended as the derivatives with respect to the variables $(v, \varepsilon_3, \varepsilon_4)$, which can be computed using the equations

$$\eta_T^0 = \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \left((v(s) - \widehat{u}(s))\widehat{\psi}_s(\eta_s) - L_{\dot{\eta}_s}\widehat{S}_s^0(\eta_s)\right)\mathrm{d}s - \int_{\widehat{\tau}_2}^{\widehat{\tau}_3} L_{\dot{\eta}_s}\widehat{S}_s^0(\eta_s)\mathrm{d}s + \int_{\widehat{\tau}_3}^T \left(\frac{\varepsilon_4}{T - \widehat{\tau}_3}\widehat{\psi}_s(\eta_s) - L_{\dot{\eta}_s}\widehat{S}_s^0(\eta_s)\right)\mathrm{d}s,$$

$$\eta_T = \widehat{q}_1 + \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} (v(s) - \widehat{u}(s))g_s^1(\eta_s)\mathrm{d}s + \int_{\widehat{\tau}_2}^{\widehat{\tau}_3} \frac{\varepsilon_3}{\widehat{\tau}_3 - \widehat{\tau}_2}k_3(\eta_s)\mathrm{d}s + \int_{\widehat{\tau}_3}^T \frac{\varepsilon_4}{T - \widehat{\tau}_3}k_4(\eta_s)\mathrm{d}s.$$

We first show that the first order approximation of J, evaluated at $(v, \varepsilon_3, \varepsilon_4) = (\widehat{u}, 0, 0)$ is null, so that the second variation is intrinsically well defined. We point out that, thanks to (A.1), $\mathrm{d}\widehat{S}_T^0(\widehat{q}_1) + \mathrm{d}\widehat{\gamma}_T(\widehat{q}_1) = -\widehat{\ell}_1$, so

that

$$\left\langle \frac{\partial J}{\partial v}, \delta v(\cdot) \right\rangle = \frac{\partial \eta_T^0}{\partial v} \delta v(\cdot) - \langle \widehat{\ell}_1 \, , \, \frac{\partial \eta_T}{\partial v} \rangle \delta v(\cdot) = \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) \left(\widehat{\psi}_s(\widehat{q}_1) - L_{g_s^1} \widehat{S}_s^0(\widehat{q}_1) \right) \mathrm{d}s + \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) \langle \widehat{\ell}_1 \, , \, g_s^1(\widehat{q}_1) \rangle \mathrm{d}s$$

$$= \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) \left(\psi(\widehat{\xi}(s)) - L_{g_s^1} \widehat{S}_s^0(\widehat{q}_1) - F_1(\widehat{\lambda}(s)) + L_{g_s^1} \widehat{S}_s^0(\widehat{q}_1) \right) \mathrm{d}s = 0.$$

Since only variations where $\varepsilon_4 = -\varepsilon_3$ are admissible, it suffices to show that $\frac{\partial J}{\partial \varepsilon_3} = \frac{\partial J}{\partial \varepsilon_4}$. Notice that

$$L_{k_4} \big(\widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \big)(\widehat{q}_1) = L_{k_4} \big(\widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_3}^0 \big)(\widehat{q}_1) = \int_{\widehat{\tau}_3}^T L_{k_4} \widehat{\psi}_s(\widehat{q}_1) \mathrm{d}s = \widehat{\psi}_T(\widehat{q}_1) - \widehat{\psi}_3(\widehat{q}_1), \tag{A.2}$$

since $\widehat{S}_{\widehat{\tau}_2}^0 = \widehat{S}_{\widehat{\tau}_3}^0$ and, for every $s \in [\widehat{\tau}_3, T]$, $\frac{\mathrm{d}}{\mathrm{d}s} \widehat{\psi}_s = L_{k_4} \widehat{\psi}_s$. Finally, we get

$$\frac{\partial J}{\partial \varepsilon_3} = \frac{\partial \eta_T^0}{\partial \varepsilon_3} + \langle \widehat{\ell}_1 \, , \, \frac{\partial \eta_T}{\partial \varepsilon_3} \rangle = -L_{k_3} \widehat{S}_{\widehat{\tau}_2}^0(\widehat{q}_1) - \langle \widehat{\ell}_1 \, , \, k_3(\widehat{q}_1) \rangle = -L_{k_3} \widehat{S}_{\widehat{\tau}_3}^0(\widehat{q}_1) - \langle \widehat{\ell}_1 \, , \, k_3(\widehat{q}_1) \rangle$$

$$\frac{\partial J}{\partial \varepsilon_4} = \frac{\partial \eta_T^0}{\partial \varepsilon_4} + \langle \widehat{\ell}_1 \, , \, \frac{\partial \eta_T}{\partial \varepsilon_4} \rangle = -L_{k_4} \widehat{S}_T^0(\widehat{q}_1) + \widehat{\psi}_T(\widehat{q}_1) - \langle \widehat{\ell}_1 \, , \, k_4(\widehat{q}_1) \rangle = -L_{k_4} \widehat{S}_{\widehat{\tau}_3}^0(\widehat{q}_1) + \widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \langle \widehat{\ell}_1 \, , \, k_4(\widehat{q}_1) \rangle$$

so that

$$\frac{\partial J}{\partial \varepsilon_3} - \frac{\partial J}{\partial \varepsilon_4} = L_{k_4 - k_3} \widehat{S}_{\widehat{\tau}_3}^0(\widehat{q}_1) - \widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \langle \widehat{\ell}_1 \, , \, (k_4 - k_3)(\widehat{q}_1) \rangle$$

$$= L_{k_4 - k_3} \widehat{S}_{\widehat{\tau}_3}^0(\widehat{q}_1) - \widehat{\psi}_{\widehat{\tau}_3}(\widehat{q}_1) - \langle -\mathrm{D}\widehat{S}_{\widehat{\tau}_3}^0 \widehat{S}_{\widehat{\tau}_3*}^{-1} + \widehat{\ell}_3 \, , \, f_1(\widehat{q}_3) \rangle = -\Phi^+(\widehat{\ell}_3) = 0.$$

Let us now compute the second variation. By some long computations, it is possible to obtain

$$J''[\delta v, \varepsilon_3, \varepsilon_4]^2 := \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \int_{\widehat{\tau}_1}^s \delta v(s) \delta v(r) L_{g_r^1} \left(\widehat{\psi}_s + L_{g_s^1} \left(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_s^0 \right) \right)(\widehat{q}_1) \mathrm{d}r \mathrm{d}s$$

$$+ \varepsilon_3 \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) L_{g_s^1} L_{k_3} \left(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \right)(\widehat{q}_1) \mathrm{d}s + \varepsilon_4 \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) L_{g_s^1} \left(|\widehat{\psi}_T| + L_{k_4} \widehat{\gamma}_T \right)(\widehat{q}_1) \mathrm{d}s$$

$$+ \frac{\varepsilon_3^2}{2} L_{k_3}^2 \left(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \right)(\widehat{q}_1) + \frac{\varepsilon_4^2}{2} L_{k_4} \left(|\widehat{\psi}_T| + L_{k_4} \widehat{\gamma}_T \right)(\widehat{q}_1)$$

$$+ \varepsilon_3 \varepsilon_4 L_{k_3} \left(|\widehat{\psi}_{\widehat{\tau}_3}| + L_{k_4} \widehat{\gamma}_T \right)(\widehat{q}_1)$$

$$= \int_{\widehat{\tau}_1}^{\widehat{\tau}_2} \delta v(s) L_{\delta \eta_r} \left(\widehat{\psi}_s + L_{g_s^1} \left(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_s^0 \right) \right)(\widehat{q}_1) \mathrm{d}s$$

$$- \varepsilon L_{\delta \eta(\widehat{\tau}_2) - \delta \eta(\widehat{\tau}_1)} L_{k_3} \left(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \right)(\widehat{q}_1) + \varepsilon L_{\delta \eta(\widehat{\tau}_2) - \delta \eta(\widehat{\tau}_1)} \left(|\widehat{\psi}_T| + L_{k_4} \widehat{\gamma}_T \right)(\widehat{q}_1) \mathrm{d}s$$

$$+ \frac{\varepsilon^2}{2} L_{k_3}^2 \left(\widehat{\gamma}_T + \widehat{S}_T^0 - \widehat{S}_{\widehat{\tau}_2}^0 \right)(\widehat{q}_1) + \frac{\varepsilon^2}{2} L_{k_4} \left(|\widehat{\psi}_T| + L_{k_4} \widehat{\gamma}_T \right)(\widehat{q}_1)$$

$$- \varepsilon^2 L_{k_3} \left(|\widehat{\psi}_{\widehat{\tau}_3}| + L_{k_4} \widehat{\gamma}_T \right)(\widehat{q}_1), \tag{A.3}$$

where we set $\varepsilon_4 = -\varepsilon_3 = \varepsilon$. Integrating backward in time system (4.4), and applying the constraints, we see that $\delta \eta_{\widehat{\tau}_2} - \delta \eta_{\widehat{\tau}_1} = -\varepsilon k(\widehat{q}_1)$. Finally, plugging equation (A.2) into (A.3) we obtain equation (4.3).

APPENDIX B. SOME TECHNICAL RESULTS AND PROOFS

Proposition B.1. *For every smooth function $\upsilon(t, \ell)\colon \mathbb{R} \times T^*M \to \mathbb{R}$, the Hamiltonian flow associated with $H_0(\ell) + \upsilon(t, \ell)\Phi^-(\ell)$ preserves $\vec{\Phi}^-$ on Σ^-, that is, if F_t is the flow associated with $H_0(\ell) + \upsilon(t, \ell)\Phi^-(\ell)$ from $\widehat{\tau}_1$ to time t, then*

$$\mathsf{F}_{t*}^{-1}\vec{\Phi}^- \circ \mathsf{F}_t(\ell) = \vec{\Phi}^-(\ell), \qquad \forall \ell \in \mathcal{O}_1. \tag{B.1}$$

Proof. We notice that, for every $\ell \in \Sigma^-$,

$$\frac{\partial}{\partial t}\mathsf{F}_{t*}^{-1}\vec{\Phi}^- \circ \mathsf{F}_t(\ell) = \mathsf{F}_{t*}^{-1}[\vec{H}_0 + \upsilon(t, \cdot)\vec{\Phi}^- + \Phi^-(\cdot)\vec{\upsilon}(t, \cdot), \vec{\Phi}^-] \circ \mathsf{F}_t(\ell)$$
$$= \mathsf{F}_{t*}^{-1}[\vec{H}_0, \vec{\Phi}^-] \circ \mathsf{F}_t(\ell),$$

since $\Phi^-(\mathsf{F}_t(\ell)) = 0$.

Using equation (5.2), for every $\ell \in \Sigma^-$ we have

$$[\vec{H}_0, \vec{\Phi}^-](\ell) = [\exp(-t\vec{\Phi}^-)_*\vec{F}_0 \circ \exp(t(\ell)\vec{\Phi}^-)|_{t=\vartheta(\ell)}, \vec{\Phi}^-]|_\ell$$
$$= [\exp(-t\vec{\Phi}^-)_*\vec{F}_0 \circ \exp(t\vec{\Phi}^-), \vec{\Phi}^-](\ell)|_{t=\vartheta(\ell)}$$
$$\quad + (\langle \mathrm{d}\vartheta, \vec{\Phi}^-\rangle[\vec{F}_0, \vec{\Phi}^-])(\exp(\vartheta(\ell)\vec{\Phi}^-))$$
$$= (1 + \langle \mathrm{d}\vartheta, \vec{\Phi}^-\rangle)[\vec{F}_0, \vec{\Phi}^-]|(\exp(\vartheta(\ell)\vec{\Phi}^-)),$$

which is null by (5.1). $\qquad\square$

Proof of Lemma 6.3

Set $\mathcal{G}_t = \widehat{\mathcal{H}}_t^{-1} \circ \mathcal{K}_t$ and notice that $\mathcal{G}_t(\widehat{\ell}_1) = \widehat{\ell}_1$ for every $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$. For every $\ell \in \Sigma^-$,

$$\frac{\partial}{\partial t}\mathcal{G}_t(\ell) = \left((\nu - \widehat{u}(t))\widehat{\mathcal{H}}_{t*}\vec{\Phi}^-\right) \circ \widehat{\mathcal{H}}_t|_{\mathcal{G}_t(\ell)} = \left(\nu(\widehat{\mathcal{H}}_t \circ \mathcal{G}_t(\ell)) - \widehat{u}(t)\right)\vec{\Phi}^-(\mathcal{G}_t(\ell)),$$

thanks to Proposition B.1. Since $\vec{\Phi}^-$ is tangent to Λ_1, we obtain that $\mathcal{G}_t(\Lambda_1) \subset \Lambda_1$. $\mathcal{G}_{t*}(T_{\widehat{\ell}_1}\Lambda_1) = T_{\widehat{\ell}_1}\Lambda_1$ for every $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$, that is, $\widehat{\mathcal{H}}_{t*}(T_{\widehat{\ell}_1}\Lambda_1) = \mathcal{K}_{t*}(T_{\widehat{\ell}_1}\Lambda_1)$ and claim 1 is proved.

By simple computations, we can prove that $\mathcal{G}_t$ is the Hamiltonian flow associated with the Hamiltonian $G_t = (K - \widehat{H}_t) \circ \widehat{\mathcal{H}}_t$. In particular, from the fact that $\mathrm{d}G_t|_{\widehat{\ell}_1} = 0$, we obtain that

$$G_t'' := \frac{1}{2}\mathrm{D}^2 G_t|_{\widehat{\ell}_1} = \frac{1}{2}\left(\mathrm{d}\Phi^- \otimes \mathrm{d}\nu + \mathrm{d}\nu \otimes \mathrm{d}\Phi^-\right)|_{\widehat{\lambda}(t)}[\widehat{\mathcal{H}}_{t*}\cdot]^2 = \langle \mathrm{d}\Phi^-|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\cdot\rangle\langle \mathrm{d}\nu|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\cdot\rangle \tag{B.2}$$

is a well defined Hamiltonian function on $T_{\widehat{\ell}_1}(T^*M)$, and that $\mathcal{G}_{t*}$ is the Hamiltonian flow associated with G_t'', see [17].

We now restrict ourselves to vectors $\delta\ell \in T_{\widehat{\ell}_1}\Sigma^-$. By definition, and using the fact that $\langle \mathrm{d}\Phi^-|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\delta\ell\rangle = 0$, it follows that $\vec{G}_t''(\delta\ell) = \langle \mathrm{d}\nu|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\delta\ell\rangle\overrightarrow{\langle \mathrm{d}\Phi^-|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\delta\ell\rangle}$. To compute this quantity, we set $\varphi(\delta\ell) = \langle \mathrm{d}\Phi^-|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\delta\ell\rangle$, and choose some vector $X \in T_{\delta\ell}(T_{\widehat{\ell}_1}(T^*M)) \simeq T_{\widehat{\ell}_1}(T^*M)$; then

$$\boldsymbol{\sigma}_{\delta\ell}(X, \vec{\varphi}(\delta\ell)) = \langle \mathrm{d}\varphi|_{\delta\ell}, X\rangle = \langle \mathrm{d}\Phi^-|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}X\rangle = \boldsymbol{\sigma}_{\widehat{\lambda}(t)}\left(\widehat{\mathcal{H}}_{t*}X, \vec{\Phi}^-(\widehat{\mathcal{H}}_t(\widehat{\ell}_1))\right)$$

$$= \boldsymbol{\sigma}_{\widehat{\ell}_1}\big(X, \widehat{\mathcal{H}}_{t*}^{-1}\vec{\Phi}^-(\widehat{\mathcal{H}}_t(\widehat{\ell}_1))\big),$$

so that, by (B.1),

$$\vec{G}_t''(\delta\ell) = \langle d\nu|_{\widehat{\lambda}(t)}, \widehat{\mathcal{H}}_{t*}\delta\ell\rangle \vec{\Phi}^-(\widehat{\ell}_1). \tag{B.3}$$

Let us now assume that $\ker(\pi\widehat{\mathcal{H}}_t)_*|_{T_{\widehat{\ell}_1}\Lambda_1} = 0$ for some $t \in [\widehat{\tau}_1, \widehat{\tau}_2]$.

Set $\lambda(t) := \mathcal{G}_{t*}\delta\ell$, for some $\delta\ell \in T_{\widehat{\ell}_1}\Lambda_1$. Thanks to (3.8), there exist a unique $\delta\ell_S \in T_{\widehat{\ell}_1}S^-$ and a unique $a \in \mathbb{R}$ such that $\delta\ell = \delta\ell_S + a\vec{\Phi}^-(\widehat{\ell}_1)$. From equation (B.3), we obtain that $\lambda(t) = \delta\ell_S + \mu(t)\vec{\Phi}^-(\ell_1)$, for some real function $\mu(\cdot)$ satisfying $\mu(\widehat{\tau}_1) = a$.

In particular, this implies that

$$\mathcal{K}_{t*}\delta\ell = \mathcal{K}_{t*}\delta\ell_S + a\vec{\Phi}^-(\widehat{\lambda}(t)) = \widehat{\mathcal{H}}_{t*}\big(\delta\ell_S + \mu(t)\vec{\Phi}^-(\widehat{\ell}_1)\big).$$

Thus, if $(\pi\mathcal{K}_t)_*\delta\ell = 0$, then $(\pi\widehat{\mathcal{H}}_t)_*\big(\delta\ell_S + \mu(t)\vec{\Phi}^-(\widehat{\ell}_1)\big) = 0$, which implies, by hypothesis, that $\delta\ell_S + \mu(t)\vec{\Phi}^-(\widehat{\ell}_1) = 0$, that is, $\delta\ell_S = 0$ and $\mu(t) = 0$. By (B.2) $\mu(t) = a + \int_{\widehat{\tau}_1}^t \langle d\nu|_{\widehat{\lambda}(s)}, \widehat{\mathcal{H}}_{s*}\delta\ell\rangle ds \equiv a$ since, by construction, ν is constant along the integral lines of $\vec{\Phi}^-$. Thus $a = 0$, so that claim 2 is proved. $\square$

Appendix C. Useful formulas

In this section, we recall classical formulas of differential geometry that we extensively use throughout the paper.

Let f, g be two vector fields on some manifold M. Then, for every t for which $\exp(tf)$ is defined, it holds

$$\frac{d}{dt}\exp(-tf)_*g\circ\exp(tf) = \exp(-tf)_*[f,g]\circ\exp(tf). \tag{C.1}$$

Let $F, G\colon T^*M \to \mathbb{R}$ be some Hamiltonian functions, and denote, as usual, with the script their flow from some time t_0. Then

$$L_{\vec{F}}G = \langle dG, \vec{F}\rangle = \boldsymbol{\sigma}(\vec{F}, \vec{G}), \qquad [\vec{F}, \vec{G}] = \overrightarrow{\boldsymbol{\sigma}(\vec{F}, \vec{G})}, \qquad \mathcal{F}_{t*}^{-1}\vec{G}|_{\mathcal{F}_t(\ell)} = \overrightarrow{G\circ\mathcal{F}_t}|_\ell.$$

References

[1] A.A. Agrachev and Yu. L. Sachkov, Control Theory from the Geometric Viewpoint. Springer-Verlag, Berlin (2004).
[2] A.A. Agrachev, G. Stefani and P. Zezza, An invariant second variation in optimal control. *Internat. J. Control* **71** (1998) 689–7158.
[3] A.A. Agrachev, G. Stefani and P. Zezza, Strong minima in optimal control. *Proc. Steklov Inst. Math.* **220** (1998) 4–26. translation from *Tr. Mat. Inst. Steklova* **220** (1998) 8–22.
[4] A.A. Agrachev, G. Stefani and P. Zezza, Strong optimality for a bang-bang trajectory. *SIAM J. Control Optim.* **41** (2002) 991–1014.
[5] B. Berret, C. Darlot, F. Jean, T. Pozzo, C. Papaxanthis and J.-P. Gauthier, The inactivation principle: Mathematical solutions minimizing the absolute work and biological implications for the planning of arm movements. *PLoS Comput. Biol.* **4** (2008) 10.
[6] N. Boizot and O. Oukacha, Consumption minimisation for an academic vehicle. *Optim. Control Appl. Methods* **41** (2020) 1001–1370.
[7] Z. Chen, J.-B. Caillau and Y. Chitour, L^1 minimization for mechanical systems. *SIAM J. Control. Optim.* **54** (2016) 1245–1265.
[8] Z. Chen, L^1-optimality conditions for the circular restricted three-body problem. *Celestial Mechanics and Dynamical Astronomy* **126** (2016) 461–481.
[9] F.C. Chittaro and L. Poggiolini, Optimality conditions for extremals containing bang and inactivated arcs, in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)* (2017).
[10] F.C. Chittaro and L. Poggiolini, Strong local optimality for generalized L1 optimal control problems. *J. Optim. Theory Appl.* **180** (2019) 207–234.

[11] F.C. Chittaro and G. Stefani, Singular extremals in multi–input time–optimal problem: a sufficient condition. *Control and Cybernetics* **39** (2010) 1029–1068.

[12] F.C. Chittaro and G. Stefani, Minimum-time strong optimality of a singular arc: The multi-input non involutive case. *ESAIM: COCV* **22** (2016) 786–810.

[13] F.H. Clarke, On the inverse function theorem. *Pacific J. Math.* **64** (1976) 97–102.

[14] F.H. Clarke, Optimization and nonsmooth analysis, Unrev. reprinting of the orig., publ. by Wiley, Hoboken 1983. Montréal: Centre de Recherches Mathématiques, Université de Montréal (1989).

[15] M.R. Hestenes, Calculus of Variations and Optimal Control Theory. John Wiley & Sons, New York, New York (1966).

[16] R. Kipka and Yu. Ledyaev, Optimal control on manifolds: Optimality conditions via nonsmooth analysis. *Commun. Appl. Anal.* **18** (2014) 563–590.

[17] J.E. Marsden and T.S. Ratiu, Introduction to Mechanics and Symmetry. Springer-Verlag, New York (1994).

[18] M. Nagahara, D.E. Quevedo and D. Nesìc, Maximum hands-off control: A paradigm of control effort minimization. *IEEE Trans. Automat. Control* **61** (2015) 735–747.

[19] L. Poggiolini, On local state optimality of bang-bang extremals in a free horizon Bolza problem. *Rendiconti del Seminario Matematico dell'Università e del Politecnico di Torino* **64** (2006) 1–23.

[20] L. Poggiolini and G. Stefani, State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *Sys. Control Lett.* **53** (2004) 269–279.

[21] L. Poggiolini and M. Spadini, Sufficient optimality conditions for a bang-bang trajectory in a Bolza problem. Mathematical Control Theory and Finance. Edited by A. Sarychev, A. Shiryaev, M. Guerra, and M. Grossinho. Springer Berlin Heidelberg, (2008) 337–357.

[22] L. Poggiolini and M. Spadini, Strong local optimality for a bang-bang trajectory in a Mayer problem. *SIAM J. Control Optim.* **49** (2011) 140–161.

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

[24] L. Poggiolini and G. Stefani, On the minimum time problem for dodgem car-like bang-singular extremals, Large-Scale Scientific Computing, edited by I. Lirkov, S. Margenov, and J. Wasniewski. Vol. 7116 of *Lecture Notes in Computer Science*. Springer, Berlin/Heidelberg (2012) 147–154.

[25] L. Poggiolini and M. Spadini, Bang–bang trajectories with a double switching time in the minimum time problem. *ESAIM: COCV* **22** (2016) 688–709.

[26] I.M. Ross, Space trajectory optimization and L^1-optimal control problems, Modern Astrodynamics, edited by P. Gurfil. Vol. 1 of Elsevier Astrodynamics Series, Butterworth-Heinemann (2006) 155–VIII.

[27] Y. Sakawa, Trajectory planning of a free-flying robot by using the optimal control. *Optim. Control Appl. Methods* **20** (1999) 235–248.

[28] G. Stefani, Strong optimality of singular trajectories, Geometric Control and Nonsmooth Analysis. Edited by F. Ancona, A. Bressan, P. Cannarsa, F. Clarke, and P. Wolenski. Vol. 76 of *Series on Advances in Mathematics for Applied Sciences*. World Scientific Publishing Co. Pte. Ltd, Hackensack, NJ (2008) 300–326.

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

[30] G. Stefani and P. Zezza, Variational Methods in Imaging and Geometric Control, chapter A Hamiltonian approach to sufficiency in optimal control with minimal regularity conditions: Part I. De Gruyter (2016).

[31] H.J. Sussmann, A nonsmooth hybrid maximum principle. Stability and Stabilization of Nonlinear Systems. Edited by D. Aeyels, F. Lamnabhi-Lagarrigue, and A. van der Schaft. Springer, London (1999) 325–354.