SUFFICIENT OPTIMALITY CONDITIONS FOR A BANG-BANG TRAJECTORY IN A BOLZA PROBLEM

LAURA POGGIOLINI AND MARCO SPADINI

Abstract. This paper gives sufficient conditions for a class of bang-bang extremals with multiple switches to be locally optimal in the strong topology. The conditions are the natural generalizations of the ones considered in [4, 11] and [12]. We require both the *strict bang-bang Legendre condition*, a non degeneracy condition at multiple switching times, and the second order conditions for the finite dimensional problems obtained by moving the switching times of the reference trajectory.

1. Introduction

We consider a Bolza problem on a fixed time interval $[0,T]$, where the control functions are bounded and enter linearly in the dynamics. Namely:

$$\text{minimize} \quad C(\xi,u) := \beta(\xi(T)) + \int_0^T \Big(f_0^0(\xi(t)) + \sum_{i=1}^m u_i f_i^0(\xi(t)) \Big)\, \mathrm{d}t \tag{1a}$$

$$\text{subject to} \quad \dot\xi(t) = f_0(\xi(t)) + \sum_{i=1}^m u_i f_i(\xi(t)) \tag{1b}$$

$$\xi(0) = \widehat{x}_0; \quad u = (u_1,\dots,u_m) \in L^\infty([0,T],[-1,1]^m). \tag{1c}$$

The state space is a n–dimensional manifold M, $\widehat{x}_0$ is a given point, the vector fields $f_0, f_1, \dots, f_m$ and the functions $f_0^0, f_1^0, \dots, f_m^0$, β are C^∞.

Optimal control problems in Economics with the above structure have been considered in [6] and references therein.

The authors aim at giving second order sufficient conditions for a *reference bang-bang couple* $(\widehat{\xi}, \widehat{u})$ *to be a local optimizer in the strong topology*; the strong topology being the one induced by $C([0,T],M)$ on the set of the admissible trajectories. Therefore optimality is with respect to neighboring trajectories, independently of the values of the associated controls.

Recall that a control $\widehat{u}$ (a trajectory $\widehat{\xi}$) is bang-bang if there is a finite number of switching times $0 < \hat t_1 < \cdots < \hat t_r < T$ such that each control function $\widehat{u}_i$ is constantly either -1 or 1 on each interval $(\hat t_k, \hat t_{k+1})$. A switching time $\hat t_k$ is called *simple* if only one control function changes value at $\hat t_k$, while it is called *multiple* if at least two control functions change value.

Second order conditions for the optimality of a bang-bang extremal with simple switches only are given in [4, 8, 11, 12], and references therein, while in [13] the author gives sufficient conditions, in the case of the minimum time problem, for L^1–local optimality of a bang bang extremal having both simple and multiple switches with the extra assumption that the Lie brackets of the switching vector fields is annihilated by the adjoint covector.

Here we consider the problem of local strong optimality in the case of a Bolza problem, when at most one double switch occurs, but there are finitely many simple ones. More precisely we extend the conditions in [4, 11, 12] requiring the sufficient second order conditions for the finite dimensional sub–problems obtained by allowing the switching times to move. We remark that, while in the case of simple switches the only variables are the switching times, each time when a double switch occurs we have to consider the two possible combinations of the switching controls. In order to complete the proof, the investigation of the invertibility of some Lipschitz continuous, piecewise C^1 operators has been done via topological methods described in the Appendix. To apply such methods it is necessary to assume a "non–degeneracy" condition at the double switching time.

2. The result

The result is based on some regularity assumption on the vector fields associated to the problem and on a second order condition for a finite dimensional sub–problem.

2.1. Notation and regularity. Assume we are given an admissible reference couple $(\widehat{\xi}, \widehat{u})$ satisfying Pontryagin maximum principle (PMP) with adjoint covector $\widehat{\lambda}$. Remark that, since no constraint is given on the final point of admissible trajectories, then $(\widehat{\xi}, \widehat{u})$ must satisfy PMP in normal form. We assume the reference control is regular bang–bang with a finite number of switching times $\widehat{t}_1, \dots, \widehat{t}_K$ such that only two kinds of switchings appear:

- *simple switching time*: only one of the control functions $\widehat{u}_1, \dots, \widehat{u}_m$ switches at time $\widehat{t}_i$;
- *double switching time*: two of the control functions $\widehat{u}_1, \dots, \widehat{u}_m$ switch at time $\widehat{t}_i$.

We assume that there is just one double switching time, which we denote by $\widehat{\tau}$. Without loss of generality we may assume that the controls switching at time $\widehat{\tau}$ are $\widehat{u}_1$ and $\widehat{u}_2$. In the interval $(0, \widehat{\tau})$, J_0 simple switches occur (if no simple switch occurs in$(0, \widehat{\tau})$, then $J_0 = 0$), while J_1 simple switches occur in the interval $(\widehat{\tau}, T)$ (if no simple switch occurs in $(\widehat{\tau}, T)$, then $J_1 = 0$). We denote the simple switching times by $\widehat{\theta}_{ij}$, $j = 1, \dots, J_i$, $i = 0, 1$ with a self–evident meaning of the double index. In order to simplify the notation, we also define $\widehat{\theta}_{00} := 0$, $\widehat{\theta}_{0,J_0+1} := \widehat{\theta}_{10} := \widehat{\tau}$, $\widehat{\theta}_{1,J_1+1} := T$, i.e. we have

$$\widehat{\theta}_{00} := 0 < \widehat{\theta}_{01} < \ldots < \widehat{\theta}_{0J_0} < \widehat{\tau} := \widehat{\theta}_{0,J_0+1} := \widehat{\theta}_{10} < \widehat{\theta}_{11} < \ldots < \widehat{\theta}_{1\,J_1} < T$$

For any m–uple $u = (u_1, \dots, u_m) \in \mathbb{R}^m$ let us denote

$$h_u \colon \ell \in T^*M \mapsto \langle \ell\,,\ f_0(\pi\ell) + \sum_{i=1}^m u_i f_i(\pi\ell)\rangle - \Big(f_0^0(\pi\ell) + \sum_{i=1}^m u_i f_i^0(\pi\ell)\Big) \in \mathbb{R}$$

and let $\widehat{f}_t$, $\widehat{f}_t^0$ and $\widehat{H}_t$ be the reference vector field, the reference running cost and the reference Hamiltonian function, respectively, i.e.

$$\widehat{f}_t(x) := f_0(x) + \sum_{i=1}^m \widehat{u}_i(t) f_i(x) \qquad \widehat{f}_t^0(x) := f_0^0(x) + \sum_{i=1}^m \widehat{u}_i(t) f_i^0(x)$$
$$\widehat{H}_t(\ell) := \langle \ell\,,\ \widehat{f}_t(\pi\ell)\rangle - \widehat{f}_t^0(\pi\ell) = h_{\widehat{u}(t)}(\ell)$$

Throughout the paper, for any Hamiltonian function $K\colon T^*M \to \mathbb{R}$, $\overrightarrow{K}$ will denote the associated Hamiltonian vector field. Also, let $\widehat{x}_d := \widehat{\xi}(\widehat{\tau})$ and $\widehat{x}_f := \widehat{\xi}(T)$. In our situation PMP reads as follows:
There exists an absolutely continuous function $\widehat{\lambda}\colon [0,T] \to T^*M$ *such that*

$$\pi\widehat{\lambda}(t) = \widehat{\xi}(t) \ \forall t \in [0,T] \qquad \widehat{\lambda}(T) = -\,\mathrm{d}\beta(\widehat{x}_f)$$

$$\dot{\widehat{\lambda}}(t) = \overrightarrow{\widehat{H}}_t(\widehat{\lambda}(t)) \ \ a.e. \ t \in [0,T],$$

$$\widehat{H}_t(\widehat{\lambda}(t)) = \max\left\{h_u(\widehat{\lambda}(t)) \colon u \in [-1,1]^m\right\} \quad \forall t \in [0,T] \tag{2}$$

Maximality condition (2) implies $\widehat{u}_i(t)\big(\langle\widehat{\lambda}(t)\,,\ f_i(\widehat{\xi}(t))\rangle - f_i^0(\widehat{\xi}(t))\big) \geq 0$ for any $t \in [0,T]$. We assume the following regularity condition holds:
Regularity. *If t is not a switching time for the control $\widehat{u}_i$, then*

$$\widehat{u}_i(t)\big(\langle\widehat{\lambda}(t)\,,\ f_i(\widehat{\xi}(t))\rangle - f_i^0(\widehat{\xi}(t))\big) > 0. \tag{3}$$

Notice that this implies that $\operatorname{argmax} h_u(\widehat{\lambda}(t)) = \widehat{u}(t)$ for any t that is not a switching time. Let

$$k_{ij} = \widehat{f}_t|_{(\widehat{\theta}_{ij},\widehat{\theta}_{i,j+1})}, \quad k_{ij}^0 = \widehat{f}_t^0|_{(\widehat{\theta}_{ij},\widehat{\theta}_{i,j+1})} \qquad i = 0,1 \quad j = 0,\ldots,J_i$$

be the restrictions of $\widehat{f}_t$, and $\widehat{f}_t^0$ to each of the time intervals where the reference control $\widehat{u}$ is constant. Let $K_{ij}(\ell) := \langle \ell\,,\ k_{ij}(\pi\ell)\rangle - k_{ij}^0(\pi\ell)$ be the associated Hamiltonian function. Then, from maximality condition (2) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(K_{10} - K_{0J_0}) \circ \widehat{\lambda}_{|\widehat{\tau}} \geq 0 \quad \text{ and } \quad \frac{\mathrm{d}}{\mathrm{d}t}\,(K_{ij} - K_{i,j-1}) \circ \widehat{\lambda}_{|\widehat{\theta}_{ij}} \geq 0$$

for any $i = 0,1, \quad j = 1,\ldots,J_i$. We assume that the strong inequality holds at each simple switching time $\widehat{\theta}_{ij}$:
Strong bang–bang Legendre condition for simple switching times.

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(K_{ij} - K_{i,j-1}) \circ \widehat{\lambda}_{|\widehat{\theta}_{ij}} > 0 \qquad i = 0,1, \quad j = 1,\ldots,J_i. \tag{4}$$

We make a stronger assumption at the double switching time $\widehat{\tau}$. Denoting by $\Delta_\nu := \widehat{u}_\nu(\widehat{\tau}{+}0) - \widehat{u}_\nu(\widehat{\tau}{-}0)$, $\nu = 1,2$, the jumps at $\widehat{\tau}$ of the two switching components, we have

$$k_{10} = k_{0J_0} + \Delta_1 f_1 + \Delta_2 f_2\,, \qquad k_{10}^0 = k_{0J_0}^0 + \Delta_1 f_1^0 + \Delta_2 f_2^0$$

Define the new vector fields and functions

$$k_\nu := k_{0J_0} + \Delta_\nu f_\nu\,, \qquad k_\nu^0 := k_{0J_0}^0 + \Delta_\nu f_\nu^0\,, \qquad \nu = 1,2,$$

with associated hamiltonian functions $K_\nu(\ell) := \langle \ell\,,\ k_\nu(\pi\ell)\rangle - k_\nu^0(\pi\ell)$. We assume that all the following one–side derivatives are strictly positive:
Strong bang–bang Legendre condition for double switching times.

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(K_\nu - K_{0J_0}) \circ \widehat{\lambda}_{|\widehat{\tau}-0} > 0, \quad \frac{\mathrm{d}}{\mathrm{d}t}\,(K_{10} - K_\nu) \circ \widehat{\lambda}_{|\widehat{\tau}+0} > 0, \quad \nu = 1,2. \tag{5}$$

Equivalently, conditions (4) and (5) can be expressed in terms of the canonical symplectic structure $\boldsymbol{\sigma}(\cdot,\cdot)$ on T^*M:

$$\boldsymbol{\sigma}\left(\overrightarrow{K}_{i,j-1}, \overrightarrow{K}_{ij}\right)(\widehat{\lambda}(\widehat{\theta}_{ij})) > 0 \qquad i = 0,1, \quad j = 1,\ldots,J_i, \tag{6}$$

$$\boldsymbol{\sigma}\left(\overrightarrow{K}_{0J_0}, \overrightarrow{K}_\nu\right)(\widehat{\lambda}(\widehat{\tau})) > 0, \quad \boldsymbol{\sigma}\left(\overrightarrow{K}_\nu, \overrightarrow{K}_{10}\right)(\widehat{\lambda}(\widehat{\tau})) > 0 \quad \nu = 1,2. \tag{7}$$

We also assume the following condition holds at the double switching time:
Non degeneracy.

$$\frac{\Delta_1 f_1(\widehat{x}_d)}{\boldsymbol{\sigma}\left(\overrightarrow{K}_{0J_0}, \overrightarrow{K}_1\right)(\widehat{\lambda}(\widehat{\tau}))} \neq \frac{\Delta_2 f_2(\widehat{x}_d)}{\boldsymbol{\sigma}\left(\overrightarrow{K}_{0J_0}, \overrightarrow{K}_2\right)(\widehat{\lambda}(\widehat{\tau}))} \tag{8}$$

2.2. **The finite dimensional sub–problem.** By allowing the switching times of the reference control function to move we can define a finite dimensional sub–problem of the given one. In doing so we must distinguish between the simple switching times and the double switching time. Moving a simple switching time $\widehat{\theta}_{i\,j}$ to time $\theta_{i\,j} := \widehat{\theta}_{i\,j} + \delta_{i\,j}$ amounts to using the values $\widehat{u}|_{(\widehat{\theta}_{i,j-1}, \widehat{\theta}_{i,j})}$ and $\widehat{u}|_{(\widehat{\theta}_{i,j}, \widehat{\theta}_{i,j+1})}$ of the control function in the time intervals $\left(\widehat{\theta}_{i,j-1}, \theta_{i\,j}\right)$ and $\left(\theta_{i\,j}, \widehat{\theta}_{i,j+1}\right)$, respectively. On the other hand, when we move the double switching time $\widehat{\tau}$ we change the switching time of two different components of the reference control function and we must allow for each of them to change its switching time independently of the other. This means that between the values of $\widehat{u}|_{(\widehat{\theta}_{0J_0}, \widehat{\tau})}$ and $\widehat{u}|_{(\widehat{\tau}, \widehat{\theta}_{01})}$ we introduce a value of the control function which is not assumed by the reference one at least in a neighborhood of $\widehat{\tau}$, and which may assume two different values according to which component switches first between the two available ones. Let $\tau_\nu := \widehat{\tau} + \varepsilon_\nu$, $\nu = 1, 2$. We move the switching time of $\widehat{u}_1$ from $\widehat{\tau}$ to $\tau_1 := \widehat{\tau} + \varepsilon_1$, and the switching time of $\widehat{u}_2$ from $\widehat{\tau}$ to $\tau_2 := \widehat{\tau} + \varepsilon_2$.

Defining $\theta_{ij} := \widehat{\theta}_{ij} + \delta_{ij}$, $\;j = 1, \dots, J_i$, $i = 0, 1$; $\theta_{0,J_0+1} := \min\{\tau_\nu,\ \nu = 1, 2\}$, $\theta_{10} := \max\{\tau_\nu,\ \nu = 1, 2\}$, $\theta_{00} := 0$ and $\theta_{1,J_1+1} := T$, we have two finite–dimensional sub–problems P_ν, $\nu = 1, 2$ given by

$$\text{minimize}\quad \beta(\xi(T)) + \sum_{i=0}^{1} \sum_{j=0}^{J_i} \int_{\theta_{ij}}^{\theta_{i,j+1}} k^0_{ij}(\xi(t))\,\mathrm{d}t + \int_{\theta_{0,J_0+1}}^{\theta_{10}} k^0_\nu(\xi(t))\,\mathrm{d}t \tag{P_νa}$$

$$\text{subject to}\quad \dot{\xi}(t) = \begin{cases} k_{0j}(\xi(t)) & t \in (\theta_{0j}, \theta_{0,j+1}) \quad j = 0, \dots, J_0, \\ k_\nu(\xi(t)) & t \in (\theta_{0,J_0+1}, \theta_{10}), \\ k_{1j}(\xi(t)) & t \in (\theta_{0j}, \theta_{0,j+1}) \quad j = 0, \dots, J_1 \end{cases} \tag{P_νb}$$

$$\text{and}\quad \xi(0) = \widehat{x}_0. \tag{P_νc}$$

where $k_\nu = k_1$, $k^0_\nu = k^0_1$ if $\theta_{0,J_0+1} = \tau_1$, and $k_\nu = k_2$, $k^0_\nu = k^0_2$ if $\theta_{0,J_0+1} = \tau_2$.

We shall denote the solution, evaluated at time t, of (P_νb) emanating from a point $x \in \mathbb{R}^n$ at time 0 as $S_t(x, \delta, \varepsilon)$. We remark that $S_t(x, 0, 0)$ is the flow associated to the reference control. We shall denote it by $\widehat{S}_t(x)$.

Notice that P_1 is defined only for $\varepsilon_1 \le \varepsilon_2$, while P_2 is defined only for $\varepsilon_2 \le \varepsilon_1$, and the reference control is the one we obtain when every δ_{ij} and ε_k is zero, i.e. in a point on the boundary of the domain of P_ν. From PMP we get that the first variation of both these problems at $\delta_{ij} = 0$, $\varepsilon_1 = \varepsilon_2 = 0$ is null, hence we can consider the second variation for the constrained problems P_1 and P_2. We shall ask for their second order variations to be positive and prove the following theorem:

Theorem 2.1. *Let $(\widehat{\xi}, \widehat{u})$ be a bang–bang regular extremal* (3) *for problem* (1) *with associated covector $\widehat{\lambda}$. Assume all the switching times of $(\widehat{\xi}, \widehat{u})$ but one are simple, while the only non–simple switching time is double.*

Assume the Legendre conditions (6) *and* (7) *hold. Also, assume the non degeneracy condition* (8) *holds at the double switching time. Assume also that each second* *variation J''_ν is positive definite on the kernel of the first variation of problem P_ν. Then $(\widehat{\xi}, \widehat{u})$ is a strict strong local optimizer for problem* (1).

3. Proof of the result

The proof will be carried out by means of Hamiltonian methods. Namely we shall define a time–dependent maximized Hamiltonian H in T^*M with flow $\mathcal{H}\colon [0,T] \times T^*M \to T^*M$ and consider the restriction of $\mathcal{H}$ to a suitable Lagrangian manifold Λ_0 containing $\widehat{\ell}_0 := \widehat{\lambda}(0)$. We shall prove that $\psi := \mathrm{id} \times \pi \circ \mathcal{H}\colon (t, \ell) \in [0,T] \times \Lambda_0 \mapsto (t, \pi\mathcal{H}_t(\ell)) \in [0,T] \times M$ is locally invertible around $[0,T] \times \{\widehat{\ell}_0\}$ and we will take advantage of the exactness of $\omega := \mathcal{H}^*\left(p\,\mathrm{d}q - H\,\mathrm{d}t\right)$ on $\{(t, \mathcal{H}_t(\ell)),\ \ell \in \Lambda_0\}$ (see Section 3.4) to reduce our problem to a local optimization problem for a function F defined in a neighborhood of $\widehat{x}_T$. Finally we shall conclude the proof of Theorem 2.1 showing that such problem has a local minimum in $\widehat{x}_T$. In proving both the invertibility of ψ and the minimality of $\widehat{x}_T$ for F we shall exploit the positivity of the second variations J''_ν. See [1, 2, 3] for a general introduction to Hamiltonian methods.

3.1. **The maximized flow.** We are now going to define the maximized Hamiltonian and the flow of its associated Hamiltonian vector field. Such flow will turn out to be Lipschitz continuous and piecewise–C^1. Define

$$\theta_{00}(\ell) := 0 \qquad \varphi_{00}(\ell) := \ell$$

for $j = 1, \ldots, J_0$

$$\theta_{0j}(\ell) := \begin{cases} \theta_{0j}(\widehat{\ell}_0) = \widehat{\theta}_{0j} \\ (K_{0j} - K_{0,j-1}) \circ \exp\theta_{0j}(\ell)\overrightarrow{K}_{0,j-1}\left(\varphi_{0,j-1}(\ell)\right) = 0 \end{cases} \tag{10a}$$

$$\varphi_{0j}(\ell) := \exp\left(-\theta_{0j}(\ell)\overrightarrow{K}_{0j}\right) \circ \exp\theta_{0j}(\ell)\overrightarrow{K}_{0,j-1}\left(\varphi_{0,j-1}(\ell)\right) \tag{10b}$$

for $\nu = 1, 2$

$$\tau_\nu(\ell) := \begin{cases} \tau_\nu(\widehat{\ell}_0) = \widehat{\tau} \\ (K_\nu - K_{0J_0}) \circ \exp\tau_\nu(\ell)\overrightarrow{K}_{0J_0}(\varphi_{0J_0}(\ell)) = 0 \end{cases} \tag{10c}$$

$$\theta_{0,J_0+1}(\ell) := \min\left\{\tau_1(\ell), \tau_2(\ell)\right\} \tag{10d}$$

$$K'(\ell) := \begin{cases} K_1(\ell) & \text{if } \theta_{0,J_0+1}(\ell) = \tau_1(\ell) \\ K_2(\ell) & \text{if } \theta_{0,J_0+1}(\ell) = \tau_2(\ell) \end{cases}$$

$$\varphi'(\ell) := \exp\left(-\theta_{0,J_0+1}(\ell)\overrightarrow{K}'\right) \circ \exp\theta_{0,J_0+1}(\ell)\overrightarrow{K}_{0J_0}\left(\varphi_{0J_0}(\ell)\right) \tag{10e}$$

$$\theta_{10}(\ell) := \begin{cases} \theta_{10}(\widehat{\ell}_0) = \widehat{\theta}_{10} = \widehat{\tau} \\ (K_{10} - K') \circ \exp\theta_{10}(\ell)\overrightarrow{K}'\left(\varphi'(\ell)\right) = 0 \end{cases} \tag{10f}$$

$$\varphi_{10} := \exp\left(-\theta_{10}(\ell)\overrightarrow{K}_{10}\right)\exp\theta_{10}(\ell)\overrightarrow{K}'\left(\varphi'(\ell)\right) \tag{10g}$$

for $j = 1, \ldots, J_1$

$$\theta_{1j}(\ell) := \begin{cases} \theta_{1j}(\widehat{\ell}_0) = \widehat{\theta}_{1j} \\ (K_{1j} - K_{1,j-1}) \circ \exp\theta_{1j}(\ell)\overrightarrow{K}_{1,j-1}\left(\varphi_{1,j-1}(\ell)\right) = 0 \end{cases} \tag{10h}$$

$$\varphi_{1j}(\ell) := \exp\left(-\theta_{1j}(\ell)\overrightarrow{K}_{1j}\right)\exp\theta_{1j}(\ell)\overrightarrow{K}_{1,j-1}\left(\varphi_{i,j-1}(\ell)\right) \tag{10i}$$

$$\theta_{1,J_1+1}(\ell) = T \tag{10j}$$

To prove that such flow is well defined, we need to show that the switching times $\theta_{ij}(\ell)$, $\tau(\ell)$ are themselves well defined and that they are ordered as follows

$$\theta_{0,j-1}(\ell) < \theta_{0j}(\ell) \ldots < \theta_{0J_0}(\ell) < \theta_{0,J_0+1}(\ell) \leq \theta_{10}(\ell) < \theta_{11}(\ell) < \ldots$$

The proof that the switching times θ_{0j}, are well defined can be carried out as in [4]. Here we show that θ_{0,J_0+1} and θ_{10} are also well defined. Let

$$\Psi_\nu(t,\ell) = (K_\nu - K_{0J_0}) \circ \exp t\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell)$$

then $\left.\dfrac{\partial \Psi_\nu}{\partial t}\right|_{(\widehat{\tau},\widehat{\ell}_0)} = \boldsymbol{\sigma}\left(\overrightarrow{K}_{0J_0}, \overrightarrow{K}_\nu\right)(\widehat{\lambda}(\widehat{\tau}))$ which is positive by (7). Now, let

$$\Phi_{10}(t,\ell) = (K_{10} - K') \circ \exp t\overrightarrow{K}' \circ \varphi_i(\ell)$$

then $\left.\dfrac{\partial \Phi_{10}}{\partial t}\right|_{(\widehat{\tau},\widehat{\ell}_0)} = \boldsymbol{\sigma}\left(\overrightarrow{K}', \overrightarrow{K}_{10}\right)(\widehat{\lambda}(\widehat{\tau}))$ which is positive by (7).

Since, by assumption $\widehat{\theta}_{i,j-1} < \widehat{\theta}_{ij}$ and $\widehat{\theta}_{0J_0} < \widehat{\tau}$, then, by continuity, $\theta_{i,j-1}(\ell) < \theta_{ij}(\ell)$ and $\theta_{0J_0}(\ell) < \theta_{0,J_0+1}(\ell)$ for any ℓ in a sufficiently small neighborhood of $\widehat{\ell}_0$. Therefore, it suffices to show that $\theta_{0,J_0+1}(\ell) \leq \theta_{10}(\ell)$. Notice that if $\tau_1(\ell) = \tau_2(\ell)$, then $\theta_{10}(\ell) = \theta_{0,J_0+1}(\ell)$, so there is nothing to prove and the choice of $K'(\ell)$ either as $K_1(\ell)$ or as $K_2(\ell)$ gives no contribution to the flow of such ℓ's, since for these ℓ's the interval $(\theta_{0,J_0+1}(\ell), \theta_{10}(\ell))$ is empty.

Let us assume $\theta_{0,J_0+1}(\ell) = \tau_1(\ell) < \tau_2(\ell)$; at time $\theta_{0,J_0+1}(\ell)$ we have

$$0 = (K_1 - K_{0J_0}) \circ \exp \theta_{0,J_0+1}(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell) \tag{11}$$

$$0 > (K_2 - K_{0J_0}) \circ \exp \theta_{0,J_0+1}(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell). \tag{12}$$

Since $K_2 - K_{0J_0} = K_{10} - K_1$, equation (12) can be written as

$$0 > (K_{10} - K_1) \circ \exp 0\overrightarrow{K}_1 \circ \exp \theta_{0,J_0+1}(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell),$$

i.e. $\theta_{10}(\ell) - \tau_1(\ell) > 0$. Analogous proof holds if $\theta_{0,J_0+1}(\ell) = \tau_2(\ell) < \tau_1(\ell)$. The proof for the θ_{1j}'s can again be done as in [4].

The maximized flow is thus defined as follows:

$$\mathcal{H} \colon (t,\ell) \in [0,T] \times T^*M \mapsto \mathcal{H}_t(\ell) \in T^*M$$

$$\mathcal{H}_t(\ell) := \begin{cases} \exp t\overrightarrow{K}_{0j}(\varphi_{0j}(\ell)) & t \in (\theta_{0j}(\ell), \theta_{0,j+1}(\ell)], \quad j = 0, \ldots, J_0 \\ \exp t\overrightarrow{K}'(\varphi'(\ell)) & t \in (\theta_{0,J_0+1}(\ell), \theta_{10}(\ell)] \\ \exp t\overrightarrow{K}_{1j}(\varphi_{1j}(\ell)) & t \in (\theta_{1j}(\ell), \theta_{1,j+1}(\ell)], \quad j = 0, \ldots, J_1 \end{cases} \tag{13}$$

3.2. **The second variations.** In order to write the second variations of the finite dimensional sub–problems P_ν we write them in Mayer form introducing an auxiliary variable x^0, as in [11]. The new state space is $\mathbb{R} \times M$ whose elements we denote by $\widetilde{x} := (x^0, x)$. Let

$$\widetilde{k}_{ij} := \begin{pmatrix} k^0_{ij} \\ k_{ij} \end{pmatrix} \quad i = 0,1, \quad j = 0, \ldots, J_i, \qquad \widetilde{k}_\nu := \begin{pmatrix} k^0_\nu \\ k_\nu \end{pmatrix} \qquad \nu = 1,2.$$

Then problem P_ν is equivalent to

(14a) minimize $\beta(\xi(T)) + \xi^0(T)$ subject to

(14b) $$\dot{\widetilde{\xi}}(t) = \begin{pmatrix} \dot{\xi}^0(t) \\ \dot{\xi}(t) \end{pmatrix} = \begin{cases} \widetilde{k}_{0j}(\xi(t)) & t \in (\theta_{0j}, \theta_{0,j+1}) \quad j = 0, \ldots, J_0, \\ \widetilde{k}_\nu(\xi(t)) & t \in (\tau, \theta_{10}), \\ \widetilde{k}_{1j}(\xi(t)) & t \in (\theta_{0j}, \theta_{0,j+1}) \quad j = 0, \ldots, J_1 \end{cases}$$

(14c) $\widetilde{\xi}(0) = (0, \widehat{x}_0)$.

We denote the solutions of (14b) evaluated at time t, emanating from a point $\widetilde{x} = (x^0, x)$ at time 0, as $\widetilde{S}_t(\widetilde{x}, \delta, \varepsilon) = \big(S^0_t(x^0, x, \delta, \varepsilon), S_t(x, \delta, \varepsilon)\big)$ and by $\widehat{\widetilde{S}}_t(\widetilde{x}) = \Big(\widehat{S}^0_t(x^0, x), \widehat{S}_t(x)\Big) = \big(S^0_t(x^0, x, 0, 0), S_t(x, 0, 0)\big)$ we denote the flow associated to the reference control. Define

$$a_{00} := \delta_{01}; \qquad a_{ij} := \delta_{ij+1} - \delta_{ij} \qquad i = 0, 1 \quad j = 1, \ldots, J_i - 1;$$
$$a_{0J_0} := \varepsilon_1 - \delta_{0J_0}; \qquad b := \varepsilon_2 - \varepsilon_1; \qquad a_{10} := \delta_{11} - \varepsilon_2; \qquad a_{1J_1} := -\delta_{1J_1}.$$

Then $b + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} = 0$. Let

$$g_{ij}(x) = \big(D\widehat{S}_{\widehat{\theta}_{ij}}\big)^{-1} k_{ij} \circ \widehat{S}_{\widehat{\theta}_{ij}}(x), \qquad g^0_{ij}(x) = k^0_{ij} \circ \widehat{S}_{\widehat{\theta}_{ij}}(x) - g_{ij} \cdot S^0_{\widehat{\theta}_{ij}}(x),$$
$$h_\nu(x) = \big(D\widehat{S}_{\widehat{\tau}}\big)^{-1} k_\nu \circ \widehat{S}_{\widehat{\tau}}(x), \qquad h^0_\nu(x) = f^0_\nu \circ \widehat{S}_{\widehat{\tau}}(x) - h_\nu \cdot \widehat{S}^0_{\widehat{\tau}}(x)$$

and put $\widehat{\beta}(x) := \beta \circ \widehat{S}_T(x)$, $\widehat{B}_0(x) := \int_0^T \widehat{f}_t(\widehat{S}_t(x))\,\mathrm{d}t$, $\alpha := -\widehat{\beta}$ and $\widehat{\gamma} := \alpha + \widehat{\beta} + \widehat{B}_0$. Also define $\Lambda_0 := \{\mathrm{d}\alpha(x),\ x \in M\}$. Let $\widetilde{\zeta}_t(\widetilde{x}, \delta, \varepsilon) := \Big(\widehat{\widetilde{S}}_t\Big)^{-1} \circ \widetilde{S}_t(x, \delta, \varepsilon)$. We consider the second–order variations of

$$J_\nu(x, a, b) := \alpha(x) + \widehat{\beta}(\zeta^\nu_T(x, a, b)) + \widehat{S^0_T}\big(\widetilde{\zeta}_T(0, x, a, b)\big)$$

at the reference triplet $(x, a, b) = (\widehat{x}_0, 0, 0)$. By assumption, for each $\nu = 1, 2$, J''_ν is positive definite on

$$\mathcal{N}^\nu_0 := \Big\{(\delta x, a, b) \in T_{\widehat{x}_0}M \times \mathbb{R}^{J_0 + J_1} \times \mathbb{R} \colon \delta x = 0, \quad b + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} = 0\Big\}.$$

Possibly redefining α by adding a suitable second–order penalty at $\widehat{x}_0$, we may assume that each second variation J''_ν is positive definite on

$$\mathcal{N}^\nu := \Big\{(\delta x, a, b) \in T_{\widehat{x}_0}M \times \mathbb{R}^{J_0 + J_1} \times \mathbb{R} \colon b + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} = 0\Big\}.$$

Let G_{ij}, H_ν be the Hamiltonian functions associated to (g_{ij}, g^0_{ij}) and (h_ν, h^0_ν) respectively, and introduce the anti–symplectic isomorphism i as in [4],

(15) $$i \colon (\delta p, \delta x) \in T^*_{\widehat{x}_0}M \times T_{\widehat{x}_0}M \mapsto -\delta p + \mathrm{d}(-\widehat{\beta} - \widehat{B}_0)_* \delta x \in T\,(T^*M)\,.$$

Defining $\overrightarrow{G}''_{ij} = i^{-1}\Big(\overrightarrow{G}_{ij}(\widehat{\ell}_0)\Big)$, $\overrightarrow{H}''_\nu = i^{-1}\Big(\overrightarrow{H}_\nu(\widehat{\ell}_0)\Big)$, we have that $\overrightarrow{G}''_{ij}$ and $\overrightarrow{H}''_\nu$ are the Hamiltonian vector fields associated to the following linear Hamiltonian functions defined in $T^*_{\widehat{x}_0}M \times T_{\widehat{x}_0}M$

(16) $$G''_{ij}(\omega, \delta x) = \langle \omega, g_{ij}(\widehat{x}_0)\rangle + \delta x \cdot \Big(g_{ij} \cdot \big(\widehat{\beta} + \widehat{B}_0 - \widehat{S}^0_{\widehat{\theta}_{ij}}\big) + g^0_{ij} \circ \widehat{S}_{\widehat{\theta}_{ij}}\Big)(\widehat{x}_0)$$

$$H''_\nu(\omega,\delta x) = \langle \omega\,,\ h_\nu(\widehat{x}_0)\rangle + \delta x \cdot \Big(h_\nu \cdot \big(\widehat{\beta} + \widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}}\big) +\ h^0_\nu \circ \widehat{S}_{\widehat{\tau}}\Big)\,(\widehat{x}_0). \tag{17}$$

Moreover $L''_0 := i^{-1}T_{\widehat{\ell}_0}\Lambda_0 = \big\{\delta\ell = \big(-D^2\widehat{\gamma}(\widehat{x}_0)(\delta x,\cdot),\delta x\big) : \delta x \in T_{\widehat{x}_0}M\big\}$ and the bilinear form J''_ν associated to the second variation can be written in a rather compact form: for any $\delta e := (\delta x, a, b) \in \mathcal{N}^\nu$ let

$$\omega_0 := -D^2\widehat{\gamma}(\widehat{x}_0)(\delta x,\cdot), \quad \delta\ell := (\omega_0,\delta x) = i^{-1}\left(\mathrm{d}\alpha_*\delta x\right),$$
$$(\omega_\nu,\delta x_\nu) := \delta\ell + \sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij}\,\overrightarrow{G}''_{ij} + b\,\overrightarrow{H}''_\nu \ \text{ and } \ \delta\ell_\nu := (\omega_\nu,\delta x_\nu).$$

Then J''_ν can be written as

$$\begin{aligned} J''_\nu\Big((\delta x,a,b),(\delta y,c,d)\Big) = &-\langle\omega_\nu\,,\ \delta y + \sum_{s=0}^{J_0} c_{0s}g_{0s} + d\,h_\nu + \sum_{s=0}^{J_1} c_{1s}g_{1s}\rangle \\ &+ \sum_{j=0}^{J_0} c_{0j}\,G''_{0j}\Big(\delta\ell + \sum_{s=0}^{j-1} a_{0s}\overrightarrow{G}''_{0s}\Big) + d\,H''_\nu\Big(\delta\ell + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}''_{0s}\Big) \\ &+ \sum_{j=0}^{J_1} c_{1j}G''_{1j}\Big(\delta\ell + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}''_{0s} + b\overrightarrow{H}''_\nu + \sum_{s=0}^{j-1} a_{1s}\overrightarrow{G}''_{1s}\Big) \end{aligned} \tag{18}$$

We shall study the positivity of J''_ν as follows: consider the increasing sequence of sub–spaces of

$$V^\nu := \Big\{(\delta x,a,b) \in \mathcal{N}^\nu \colon \delta x + \sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij}g_{ij}(\widehat{x}_0) + b\,h_\nu(\widehat{x}_0) = 0\Big\}.$$

defined as

$$\begin{aligned} V^\nu_{0j} &:= \{(\delta x,a,b) \in V^\nu \colon a_{0s} = 0 \quad \forall s = j+1,\dots,J_0,\ a_{1s} = 0 \\ &\qquad\qquad \forall s = 0,\dots,J_1, \quad b = 0\}, \\ V^\nu_{1j} &:= \{(\delta x,a,b) \in V^\nu \colon a_{1s} = 0 \quad \forall s = j+1,\dots,J_1\}. \end{aligned}$$

Then $V^1_{0j} = V^2_{0j}$ for any $j = 0,\dots,J_0$, so we denote these sets as V_{0j}. Moreover

$$\dim\Big(V_{0j} \cap V_{0,j-1}^{\perp_{J''_\nu}}\Big) = \dim\Big(V^\nu_{1k} \cap V_{1,k-1}^{\perp_{J''_\nu}}\Big) = 1, \quad \dim\Big(V^\nu_{10} \cap V_{0J_0}^{\perp_{J''_\nu}}\Big) = 2$$

for any $j = 2,\dots,J_0$ and any $k = 0,\dots,J_1$.

Using the first order approximations of the quantities $\theta_{ij}(\ell)$, $\varphi_{ij}(\ell)$, defined in equations (10) and proceeding as in [4] we can prove the following lemmata

Lemma 3.1. *$\delta e = (\delta x,a,b) \in V_{0j} \cap V_{0,j-1}^{\perp_{J''_\nu}}$ if and only if $\delta e \in V_{0j}$ and*

$$G''_{0s}(\delta\ell + \sum_{\mu=0}^{s-1} a_{0\mu}\overrightarrow{G}''_{0\mu}) = G''_{0,j-1}(\delta\ell + \sum_{s=0}^{j-2} a_{0s}\overrightarrow{G}''_{0s}), \quad \forall\, s = 0,\dots,j-2 \tag{19}$$

i.e. $a_{0s} = \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\mathrm{d}\alpha_*\delta x) \quad \forall s = 0,\dots,j-2.$

In this case $J''[\delta e]^2 = a_{0j}\,\boldsymbol{\sigma}\Big(\delta\ell + \sum_{s=0}^{j-1} a_{0s}\overrightarrow{G}''_{0s}, \overrightarrow{G}''_{0j} - \overrightarrow{G}''_{0,j-1}\Big).$

Lemma 3.2. $\delta e = (\delta x, a, b) \in V_{10}^{\nu} \cap V_{0J_0}^{\perp_{J''_\nu}}$ *if and only if* $\delta e \in V_{10}^{\nu}$ *and*

$$G''_{0s}\Big(\delta\ell + \sum_{\mu=0}^{s-1} a_{0\mu}\,\overrightarrow{G}''_{0\mu}\Big) = G''_{0,J_0}\Big(\delta\ell + \sum_{s=0}^{J_0-1} a_{0s}\,\overrightarrow{G}''_{0s}\Big)\,, \quad \forall\, s = 0, \dots, J_0 - 1 \tag{20}$$

i.e. $a_{0s} = \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\mathrm{d}\alpha_*\delta x) \quad \forall s = 0, \dots, J_0 - 1.$

In this case

$$J''[\delta e]^2 = b\,\boldsymbol{\sigma}\Big(\delta\ell + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}''_{0s}, \overrightarrow{H}''_{\nu} - \overrightarrow{G}''_{0,J_0}\Big) + a_{10}\,\boldsymbol{\sigma}\Big(\delta\ell + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}''_{0s} + b\overrightarrow{H}''_{\nu}, \overrightarrow{G}''_{10} - \overrightarrow{H}''_{\nu}\Big).$$

Lemma 3.3. $\delta e = (\delta x, a, b) \in V_{1j}^{\nu} \cap V_{1,j-1}^{\perp_{J''_\nu}}$ *if and only if* $\delta e \in V_{1j}^{\nu}$ *and*

$$\begin{aligned} G''_{0s}\Big(\delta\ell + \sum_{i=0}^{s-1} a_{0i}\,\overrightarrow{G}''_{0i}\Big) &= H''_{\nu}\Big(\delta\ell + \sum_{i=0}^{J_0} a_{0i}\,\overrightarrow{G}''_{0i}\Big) \\ &= G''_{1k}\Big(\delta\ell + \sum_{i=0}^{J_0} a_{0i}\,\overrightarrow{G}''_{0i} + b\overrightarrow{H}''_{\nu} + \sum_{i=0}^{k-1} a_{1i}\,\overrightarrow{G}''_{1i}\Big) \\ &\qquad \forall\, s = 0, \dots, J_0 \quad \forall\, k = 0, \dots, j-2 \end{aligned}$$

i.e. if and only if $\delta e \in V_{1j}^{\nu}$ *and*

$$\begin{aligned} a_{0s} &= \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\mathrm{d}\alpha_*\delta x) \quad \forall s = 0, \dots, J_0 \\ b &= \mathrm{d}(\theta_{0,J_0+1} - \theta_{0J_0})\,(\mathrm{d}\alpha_*\delta x) \\ a_{1s} &= \mathrm{d}(\theta_{1,s+1} - \theta_{1s})\,(\mathrm{d}\alpha_*\delta x) \quad \forall s = 0, \dots, j-2. \end{aligned}$$

In this case

$$J''[\delta e]^2 = a_{1j}\,\boldsymbol{\sigma}\left(\delta\ell + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}''_{0s} + b\overrightarrow{H}''_{\nu} + \sum_{i=0}^{j-1} a_{1i}\overrightarrow{G}''_{1i}, \overrightarrow{G}''_{1j} - \overrightarrow{G}''_{1,j-1}\right).$$

Lemma 3.4. $\delta e = (\delta x, a, b) \in \mathcal{N}^{\nu} \cap V_{1J_1}^{\perp_{J''_\nu}}$ *if and only if* $\delta e \in \mathcal{N}^{\nu}$ *and*

$$\begin{aligned} G''_{0s}\Big(\delta\ell + \sum_{i=0}^{s-1} a_{0i}\,\overrightarrow{G}''_{0i}\Big) &= H''_{\nu}\Big(\delta\ell + \sum_{i=0}^{J_0} a_{0i}\,\overrightarrow{G}''_{0i}\Big) \\ &= G''_{1k}\Big(\delta\ell + \sum_{i=0}^{J_0} a_{0i}\,\overrightarrow{G}''_{0i} + b\overrightarrow{H}''_{\nu} + \sum_{i=0}^{k-1} a_{1i}\,\overrightarrow{G}''_{1i}\Big) \\ &\qquad \forall\, s = 0, \dots, J_0 \quad \forall\, k = 0, \dots, J_1 \end{aligned}$$

i.e. if and only if $\delta e \in \mathcal{N}^{\nu}$ *and*

$$\begin{aligned} a_{0s} &= \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\mathrm{d}\alpha_*\delta x) \quad \forall s = 0, \dots, J_0 \\ b &= \mathrm{d}(\theta_{0,J_0+1} - \theta_{0J_0})\,(\mathrm{d}\alpha_*\delta x) \\ a_{1s} &= \mathrm{d}(\theta_{1,s+1} - \theta_{1s})\,(\mathrm{d}\alpha_*\delta x) \quad \forall s = 0, \dots, J_1 - 1. \end{aligned}$$

In this case

$$J''[\delta e]^2 = -\langle \omega_\nu\,,\ \delta x + \sum_{i=0}^{1}\sum_{s=0}^{J_i} a_{is} g_{is} + b\, h_\nu \rangle$$
$$= \boldsymbol{\sigma}\left(\left(0, \delta x + \sum_{i=0}^{1}\sum_{s=0}^{J_i} a_{is} g_{is} + b h_\nu \right), -D^2\widehat{\gamma}(\delta x, \cdot) + \sum_{i=0}^{1}\sum_{s=0}^{J_i} a_{is}\overrightarrow{G}''_{is} + b\overrightarrow{H}''_\nu \right).$$

3.3. **The invertibility of the flow.** Lemma 3.1 allows us to prove the following property (whose proof can be found in [4]) for the linearization of the maximized flow:

Lemma 3.5. *Let* $j \in \{1, \dots, J_0\}$ *and* $\delta x_1,\ \delta x_2 \in T_{\widehat{x}_0}M$ *such that* $\mathrm{d}\theta_{0j}(\delta x_2) < 0 < \mathrm{d}\theta_{0j}(\delta x_1)$. *Then* $\left(\pi \circ \mathcal{H}_{\widehat{\theta}_{0j}}\right)_* \mathrm{d}\alpha_* \delta x_1 \neq \left(\pi \circ \mathcal{H}_{\widehat{\theta}_{0j}}\right)_* \mathrm{d}\alpha_* \delta x_2$.

Lemma 3.5 implies that the application

$$\psi\colon (t, \ell) \in [0, T] \times \Lambda_0 \mapsto (t, \pi \circ \mathcal{H}_t(\ell)) \in [0, T] \times M \tag{21}$$

is locally invertible around $[0, \widehat{\tau} - \varepsilon] \times \left\{\widehat{\ell}_0\right\}$. In fact, ψ is locally one–to–one if and only if $\pi \circ \mathcal{H}_t$ is locally one–to–one in $\widehat{\ell}_0$ for any t. On the other hand $\pi \circ \mathcal{H}_t$ is locally one–to–one for any $t < \widehat{\tau}$ if and only if it is one–to–one at any $\widehat{\theta}_{0j}$. This property is granted by Lemma 3.5.

We now want to show that such procedure can be carried out also on $[\widehat{\tau} - \varepsilon, T] \times \left\{\widehat{\ell}_0\right\}$, so that ψ will turn out to be locally invertible from a neighborhood $[0, T] \times \mathcal{O} \subset [0, T] \times \Lambda_0$ of $[0, T] \times \left\{\widehat{\ell}_0\right\}$ onto a neighborhood $\mathcal{U} \subset [0, T] \times M$ of the graph $\widehat{\Xi}$ of $\widehat{\xi}$.

The first step will be proving the invertibility of $\pi \circ \mathcal{H}_{\widehat{\tau}}$ at $\widehat{\ell}_0$. In a neighborhood of $\widehat{\ell}_0$, $\pi \circ \mathcal{H}_{\widehat{\tau}}$ has the following piecewise representation

M1	$\min\{\tau_1(\ell),\ \tau_2(\ell)\} \geq \widehat{\tau}$	$\pi \exp \widehat{\tau}\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell)$
M2	$\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_1(\ell) \leq \widehat{\tau} \leq \theta_{10}(\ell)$	$\pi \exp \widehat{\tau}\overrightarrow{K}_1 \circ \exp(-\tau_1(\ell)\overrightarrow{K}_1) \circ \exp \tau_1(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell)$
M3	$\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_2(\ell) \leq \widehat{\tau} \leq \theta_{10}(\ell)$	$\pi \exp \widehat{\tau}\overrightarrow{K}_2 \circ \exp(-\tau_2(\ell)\overrightarrow{K}_2) \circ \exp \tau_2(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell)$
M4	$\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_1(\ell) \leq \theta_{10}(\ell) \leq \widehat{\tau}$	$\pi \exp\left(\widehat{\tau} - \theta_{10}(\ell)\right)\overrightarrow{K}_{10} \circ \exp \theta_{10}(\ell)\overrightarrow{K}_1 \circ \exp(-\tau_1(\ell)\overrightarrow{K}_1) \circ \exp \tau_1(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell)$
M5	$\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_2(\ell) \leq \theta_{10}(\ell) \leq \widehat{\tau}$	$\pi \exp\left(\widehat{\tau} - \theta_{10}(\ell)\right)\overrightarrow{K}_{10} \circ \exp \theta_{10}(\ell)\overrightarrow{K}_2 \circ \exp(-\tau_2(\ell)\overrightarrow{K}_2) \circ \exp \tau_2(\ell)\overrightarrow{K}_{0J_0} \circ \varphi_{0J_0}(\ell)$

The invertibility of $\pi \circ \mathcal{H}_{\widehat{\tau}}$ will be proved by the means of Theorem 4.1 in the Appendix. Notice that the non degeneracy condition (8) implies that the second order penalty on α can be chosen so that $\mathrm{d}\tau_1(\mathrm{d}\alpha_*(\cdot)) \neq \mathrm{d}\tau_2(\mathrm{d}\alpha_*(\cdot))$. In order to apply Theorem 4.1 we write the piecewise linearized map $(\pi \circ \mathcal{H}_{\widehat{\tau}})_*$.
$(\pi \circ \mathcal{H}_{\widehat{\tau}})_*\delta\ell$ is given by

M1′ if $\min\{\mathrm{d}\tau_1(\delta\ell),\ \mathrm{d}\tau_2(\delta\ell)\} \geq 0$

$$\exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0}\delta\ell \tag{22a}$$

M2′ if $\mathrm{d}\tau_1(\delta\ell) \leq 0 \leq \mathrm{d}\theta_{10}(\ell)$, $\mathrm{d}\tau_1(\delta\ell) \leq \mathrm{d}\tau_2(\delta\ell)$

$$-\mathrm{d}\tau_1(\delta\ell)\Big[\exp(\widehat{\tau}\overrightarrow{K}_1)_*\overrightarrow{K}_1 - \overrightarrow{K}_{0J_0}\Big] + \exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0*}\delta\ell \tag{22b}$$

M3′ if $\mathrm{d}\tau_2(\delta\ell) \le 0 \le \mathrm{d}\theta_{10}(\delta\ell)$, $\mathrm{d}\tau_2(\delta\ell) \le \mathrm{d}\tau_1(\delta\ell)$

$$-\,\mathrm{d}\tau_2(\delta\ell)\Big[\exp(\widehat{\tau}\overrightarrow{K}_2)_*\overrightarrow{K}_2 - \overrightarrow{K}_{0J_0}\Big] + \exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0*}\delta\ell \tag{22c}$$

M4′ if $\mathrm{d}\tau_1(\delta\ell) \le \mathrm{d}\theta_{10}(\delta\ell) \le 0$, $\mathrm{d}\tau_1(\delta\ell) \le \mathrm{d}\tau_2(\delta\ell)$

$$\begin{aligned} &\mathrm{d}\theta_{10}(\delta\ell)\left(\overrightarrow{K}_{10} + \mathrm{d}\theta_{10}(\delta\ell)\exp(\widehat{\tau}\overrightarrow{K}_{10})_*\overrightarrow{K}_1\right) \\ &\quad + \exp(\widehat{\tau}\overrightarrow{K}_1)_*\Big(-\,\mathrm{d}\tau_1(\delta\ell)\overrightarrow{K}_1 + \exp(-\widehat{\tau}_1\overrightarrow{K}_1)_*\,\mathrm{d}\tau_1(\delta\ell)\overrightarrow{K}_{0J_0} \\ &\qquad + \exp(-\widehat{\tau}\overrightarrow{K}_1)_*\exp(\widehat{\tau}\overrightarrow{K}_{0J_0})\varphi_{0J_0*}\delta\ell\Big) \end{aligned} \tag{22d}$$

M5′ if $\mathrm{d}\tau_2(\delta\ell) \le \mathrm{d}\theta_{10}(\delta\ell) \le 0$, $\mathrm{d}\tau_2(\delta\ell) \le \mathrm{d}\tau_1(\delta\ell)$

$$\begin{aligned} &\mathrm{d}\theta_{10}(\delta\ell)\left(\overrightarrow{K}_{10} + \mathrm{d}\theta_{10}(\delta\ell)\exp(\widehat{\tau}\overrightarrow{K}_{10})_*\overrightarrow{K}_2\right) \\ &\quad + \exp(\widehat{\tau}\overrightarrow{K}_2)_*\Big(-\,\mathrm{d}\tau_2(\delta\ell)\overrightarrow{K}_2 + \exp(-\widehat{\tau}_2\overrightarrow{K}_2)_*\,\mathrm{d}\tau_2(\delta\ell)\overrightarrow{K}_{0J_0} \\ &\qquad + \exp(-\widehat{\tau}\overrightarrow{K}_2)_*\exp(\widehat{\tau}\overrightarrow{K}_{0J_0})\varphi_{0J_0*}\delta\ell\Big) \end{aligned} \tag{22e}$$

According to Theorem 4.1, in order to prove the invertibility of our map it is sufficient to prove that both the map and its linearization are continuous in a neighborhood of $\widehat{\ell}_0$ and of 0 respectively, that they mantain the orientation and that there exists a point $\delta\overline{x}$ whose preimage is a singleton not belonging to the above boundaries.

Notice that the continuity of $\pi \circ \mathcal{H}_{\widehat{\tau}}$ follows from the very definition of the maximized flow. Discontinuities of $(\pi \circ \mathcal{H}_{\widehat{\tau}})_*$ may occur only at the boundaries described above. A direct computation in formulas (22) shows that this is not the case. Let us now prove the last assertion.

Throughout the rest of the section, all the Hamiltonian vector fields $\overrightarrow{G}_{ij}$ and $\overrightarrow{H}_\nu$ are computed in $\widehat{\ell}_0$. For "symmetry" reasons it is convenient to look for $\delta\overline{x}$ among those which belong to the image of the set $\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda_0 : 0 < \mathrm{d}\tau_1(\delta\ell) = \mathrm{d}\tau_2(\delta\ell)\}$. Observe that this implies that $\mathrm{d}\theta_{10}(\delta\ell) = \mathrm{d}\tau_1(\delta\ell) = \mathrm{d}\tau_2(\delta\ell)$: Introducing the quantity $\eta_{\widehat{\tau}}(\delta\ell) := \widehat{\mathcal{H}}_{\widehat{\tau}*}^{-1}\exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0*}(\delta\ell)$, the assertion $\mathrm{d}\tau_1(\delta\ell) = \mathrm{d}\tau_2(\delta\ell)$ can be written as

$$\frac{\boldsymbol{\sigma}\left(\eta_{\widehat{\tau}}(\delta\ell), \overrightarrow{H}_1 - \overrightarrow{G}_{0J_0}\right)}{\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1\right)} = \frac{\boldsymbol{\sigma}\left(\eta_{\widehat{\tau}}(\delta\ell), \overrightarrow{H}_2 - \overrightarrow{G}_{0J_0}\right)}{\sigma(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_2)}$$

thus, $\mathrm{d}\theta_{10}(\delta\ell)\boldsymbol{\sigma}\left(\overrightarrow{H}_2, \overrightarrow{G}_{10}\right)$ is given by

$$-\boldsymbol{\sigma}\left(\eta_{\widehat{\tau}}(\delta\ell), \overrightarrow{G}_{10} - \overrightarrow{H}_2\right) + \mathrm{d}\tau_2(\delta\ell)\boldsymbol{\sigma}\left(\overrightarrow{H}_2 - \overrightarrow{G}_{0J_0}, \overrightarrow{G}_{10} - \overrightarrow{H}_2\right)$$

and, using that $\mathrm{d}\tau_1(\delta\ell) = \mathrm{d}\tau_2(\delta\ell)$ and $\overrightarrow{G}_{10} - \overrightarrow{H}_2 = \overrightarrow{H}_1 - \overrightarrow{G}_{0J_0}$, this is equal to $\mathrm{d}\tau_1(\delta\ell)\boldsymbol{\sigma}\left(\overrightarrow{H}_2, \overrightarrow{G}_{10}\right)$.

Thus, we consider $\delta\overline{x} = \pi_*\exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0*}\delta\ell_1$ with $0 < \mathrm{d}\tau_1(\delta\ell_1) = \mathrm{d}\tau_2(\delta\ell_1)$. Clearly $\delta\overline{x}$ has at most one preimage per each of the above sectors. Let us prove that actually its preimage is the singleton $\{\delta\ell_1\}$.

Assume by contradiction that there is $\delta\ell_2$ in sector M2$'$ such that

$$\begin{aligned}\pi_* \exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0*}(\delta\ell_1) =&\pi\big(\exp \widehat{\tau}\overrightarrow{K}_1)_*\big(-\mathrm{d}\tau_1(\delta\ell_2)\overrightarrow{K}_1\big)\\ &+\pi_*\,\mathrm{d}\tau_1(\delta\ell_2)\overrightarrow{K}_{0J_0}+\pi_*\exp(\widehat{\tau}\overrightarrow{K}_{0J_0})_*\varphi_{0J_0*}(\delta\ell_2)\end{aligned}$$

Taking the pull-back we get

$$\begin{aligned}&\delta x_1-\delta x_2+\sum_{s=1}^{J_0-1}\mathrm{d}(\theta_{0,s+1}-\theta_{0s})\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)\big)g_{0s}\\ &\qquad-\Big(\mathrm{d}\theta_{0J_0}\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)\big)+\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_2)\Big)\,g_{0J_0}+\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_2)h_1=0.\end{aligned}$$

Consider $\delta e := (\delta x_1-\delta x_2, a, b)$, where, for $j=0,\dots,J_1$, $a_{1j}=0$ and, for $s=0,\dots,J_0$,

$$a_{0s}=\begin{cases}\mathrm{d}(\theta_{0,s+1}-\theta_{0s})\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)\big) & s=0,\dots,J_0-1\\ -\mathrm{d}\theta_{0J_0}\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)\big)+\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_2) & s=J_0,\end{cases}$$

and $b=\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_2)<0$. Thus $\delta e\in V^1_{10}\cap V_{0J_0}^{\perp_{J_1''}}$, therefore Lemma 3.2 applies:

$$\boldsymbol{\sigma}\Big(\delta\ell+\sum_{s=0}^{J_0}a_{0s}\overrightarrow{G}''_{0s},\overrightarrow{H}''_1-\overrightarrow{G}''_{0J_0}\Big)<0$$

where $\delta\ell=\big(-D^2\widehat{\gamma}(x_0)(\delta x_1-\delta x_2,\cdot),\delta x_1-\delta x_2\big)$. Thus, applying i,

$$\boldsymbol{\sigma}\Big(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)+\sum_{s=0}^{J_0}a_{0s}\overrightarrow{G}_{0s},\overrightarrow{H}_1-\overrightarrow{G}_{0J_0}\Big)>0$$

or, linearizing the formula for $\tau_1(\delta\ell)$ in (10),

$$\boldsymbol{\sigma}\left(\eta_{\widehat{\tau}}(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)),\overrightarrow{H}_1-\overrightarrow{G}_{0J_0}\right)-\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_2)\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0},\overrightarrow{H}_1\right)>0$$

which implies $-\mathrm{d}\tau_1\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_2)\big)-\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_2)>0$ or $\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_1)<0$ a contradiction.

Let us now assume by contradiction that there is $\delta\ell_4$ in sector M4$'$ whose image under the linearized map coincides with $\delta\overline{x}$.

Thus, proceeding in the same way as between sectors M1$'$ and M2$'$, we get

$$\begin{aligned}&\delta x_1-\delta x_4+\sum_{s=1}^{J_0-1}\mathrm{d}(\theta_{0,s+1}-\theta_{0s})\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_4)\big)g_{0s}-\Big(\mathrm{d}\theta_{0J_0}\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_4)\big)\\ &\qquad+\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_4)\Big)g_{0J_0}-\mathrm{d}(\theta_{10}-\tau_1)(\mathrm{d}\alpha_*\delta x_4)h_1+\mathrm{d}\theta_{10}(\mathrm{d}\alpha_*\delta x_4)g_{10}=0.\end{aligned}$$

Consider $\delta e:=(\delta x_1-\delta x_4,a,b)$, where, for $j=1,\dots,J_1$, $a_{1j}=0$, $a_{10}=\mathrm{d}\theta_{10}(\delta x_4)<0$ and, for $s=0,\dots,J_0$,

$$a_{0s}=\begin{cases}\mathrm{d}(\theta_{0,s+1}-\theta_{0s})\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_4)\big) & s=0,\dots,J_0-1\\ -\mathrm{d}\theta_{0J_0}\big(\mathrm{d}\alpha_*(\delta x_1-\delta x_4)\big)+\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_4) & s=J_0,\end{cases}$$

and $b = \mathrm{d}(\theta_{10} - \tau_1)(\mathrm{d}\alpha_*\delta x_4) < 0$. Thus $\delta e \in V_{10}^1 \cap V_{0J_0}^{\perp_{J_1''}}$, and Lemma 3.3 applies

$$b\,\boldsymbol{\sigma}\left(\mathrm{d}\alpha_*(\delta x_1 - \delta x_4) + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}_{0s}, \overrightarrow{H}_1 - \overrightarrow{G}_{0J_0}\right) \\ + a_{10}\,\boldsymbol{\sigma}\left(\mathrm{d}\alpha_*(\delta x_1 - \delta x_4) + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}_{0s} + b\overrightarrow{H}_1, \overrightarrow{G}_{10} - \overrightarrow{H}_1\right) < 0$$

The coefficient of b is equal to

$$\begin{aligned} &\boldsymbol{\sigma}\left(\eta_{\widehat{\tau}}(\mathrm{d}\alpha_*(\delta x_1 - \delta x_4)) - \mathrm{d}\tau_1(\delta x_4)\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1 - \overrightarrow{G}_{0J_0}\right) \\ &= -\mathrm{d}\tau_1\big(\mathrm{d}\alpha_*(\delta x_1 - \delta x_4)\big)\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1\right) - \mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_4)\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1\right) \\ &= -\mathrm{d}\tau_1(\mathrm{d}\alpha_*\delta x_1)\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1,)\right) < 0 \end{aligned}$$

On the other hand, taking the first order approximations in (10), one can show that the coefficient of a_{10} is:

$$\Big(-\mathrm{d}\tau_1(\mathrm{d}\alpha_*(\delta x_1)) + \mathrm{d}\theta_{10}(\mathrm{d}\alpha_*(\delta x_4))\big(-\mathrm{d}\tau_1(\mathrm{d}\alpha_*(\delta x_1))\big)\Big)\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_2\right) \\ - \mathrm{d}(\theta_{10} - \tau_1)(\mathrm{d}\alpha_*(\delta x_4))\big(-\mathrm{d}\tau_1(\mathrm{d}\alpha_*(\delta x_1))\big)\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1\right) < 0$$

which is impossible.

The orientation preserving condition can be proved by the means of Lemma 4.1: consider any pair of adjacent cones M_i' and M_j'. They are separated by a hyperplane. A similar argument to the one used above shows that any pair of points lying on opposite sides of the separating hyperplane have different images under the maps used in M_i' and M_j', extended to the corresponding half space.

This proves the invertibility of $\pi \circ \mathcal{H}_{\widehat{\tau}}$, hence ψ is one–to–one in a neighborhood of $[0, \widehat{\theta}_{10} - \varepsilon] \times \{\widehat{\ell}_0\}$.

We only sketch the idea of the proof of the invertibility of $\pi \circ \mathcal{H}_{\widehat{\theta}_{1j}}$, $j = 1, \dots, J_1$. Given j, there are four regions $N_{1j}, \dots, N_{4j}$ in Λ_0, characterized by the following properties

$$\begin{aligned} &(N_{1j})\ \theta_{1j}(\ell) \geq \widehat{\theta}_{1j} \text{ and } \theta_{0,J_0+1}(\ell) = \tau_1(\ell), \\ &(N_{2j})\ \theta_{1j}(\ell) \geq \widehat{\theta}_{1j} \text{ and } \theta_{0,J_0+1}(\ell) = \tau_2(\ell), \\ &(N_{3j})\ \theta_{1j}(\ell) < \widehat{\theta}_{1j} \text{ and } \theta_{0,J_0+1}(\ell) = \tau_1(\ell), \\ &(N_{4j})\ \theta_{1j}(\ell) < \widehat{\theta}_{1j} \text{ and } \theta_{0,J_0+1}(\ell) = \tau_2(\ell); \end{aligned}$$

as for $\pi \circ \mathcal{H}_{\widehat{\tau}}$, $\pi \circ \mathcal{H}_{\widehat{\theta}_{1j}}$ turns out to be a Lipschitz continuous, piecewise C^1 application. Its invertibility can be proved applying again Theorem 4.1. We will consider first the case $j = 1$ and the following linearization of $\pi \circ \mathcal{H}_{\widehat{\theta}_{11}}$. Here, for the sake of brevity we have already passed to the pullback

N_{1j}' where $\mathrm{d}\theta_{11}(\delta\ell) \geq 0$ and $\mathrm{d}\tau_1(\delta\ell) \leq \mathrm{d}\tau_2(\delta\ell)$,

$$-\mathrm{d}\theta_{10}(\delta\ell)g_{10} + \mathrm{d}(\theta_{10} - \tau_1)(\delta\ell)h_1 + \mathrm{d}\tau_1(\delta\ell)g_{0J_0} + \eta_{\widehat{\tau}}(\delta\ell)$$

N_{2j}' where $\mathrm{d}\theta_{11}(\delta\ell) \geq 0$ and $\mathrm{d}\tau_2(\delta\ell) \leq \mathrm{d}\tau_1(\delta\ell)$,

$$-\mathrm{d}\theta_{10}(\delta\ell)g_{10} + \mathrm{d}(\theta_{10} - \tau_2)(\delta\ell)h_2 + \mathrm{d}\tau_2(\delta\ell)g_{0J_0} + \eta_{\widehat{\tau}}(\delta\ell)$$

N'_{3j} where $\mathrm{d}\theta_{11}(\delta\ell) \le 0$ and $\mathrm{d}\tau_1(\delta\ell) \le \mathrm{d}\tau_2(\delta\ell)$,

$-\mathrm{d}\theta_{11}(\delta\ell)g_{11} + \mathrm{d}(\theta_{11}-\theta_{10})(\delta\ell)g_{10} + \mathrm{d}(\theta_{10}-\tau_1)(\delta\ell)h_1 + \mathrm{d}\tau_1(\delta\ell)g_{0J_0} + \eta_{\widehat{\tau}}(\delta\ell)$

N'_{4j} where $\mathrm{d}\theta_{11}(\delta\ell) \le 0$ and $\mathrm{d}\tau_2(\delta\ell) \le \mathrm{d}\tau_2(\delta\ell)$,

$-\mathrm{d}\theta_{11}(\delta\ell)g_{11} + \mathrm{d}(\theta_{11}-\theta_{10})(\delta\ell)g_{10} + \mathrm{d}(\theta_{10}-\tau_2)(\delta\ell)h_2 + \mathrm{d}\tau_2(\delta\ell)g_{0J_0} + \eta_{\widehat{\tau}}(\delta\ell)$

As above, according to Theorem 4.1, we only have to prove that both the map and its linearization are continuous in a neighborhood of $\widehat{\ell}_0$ and of 0 respectively, that the the linearized pieces are orientation preserving and that there exists a point $\delta\overline{x}$ whose preimage is a singleton. The only nontrivial part is the last statement which can be proved by picking $\delta\overline{x} \in N'_{11} \cap N'_{12}$.

3.4. **Reduction to a finite–dimensional problem.** In this section, in order to shorten the notation, for any $(t,\ell) \in [0,T]\times\Lambda_0$, let us define $\psi_t(\ell) := \pi\circ\mathcal{H}_t(\ell)$. Also we recall that the maximized Hamiltonian function is a lift: $H_t(\ell) = \langle \ell\,,\ f(t,\pi\ell)\rangle - f^0(t,\pi\ell)$

In the product space $[0,T]\times M$ consider the path obtained with the concatenation of the graph of a generic trajectory, $\Xi := \{(t,\xi(t))\colon t \in [0,T]\}$ (ran backward) contained in $\mathcal{U}$ and the graph of the reference trajectory $\widehat{\Xi} := \{(t,\widehat{\xi}(t))\colon t \in [0,T]\}$. We can obtain a close circuit with a path γ from $(T,\widehat{x}_T)$ to $(T,\xi(T))$ whose image is contained in $\{T\}\times M$.
Consider the following sets in $[0,T]\times T^*M$:

$$\mathcal{O}_{0j} = \{(t,\ell)\colon \ell\in\mathcal{O},\quad t\in[\theta_{0,j-1}(\ell),\theta_{0j}(\ell)]\} \qquad j=1,\dots,J_0$$

and, for $\nu = 1,2$ define

$$\begin{aligned}
\mathcal{O}^{\nu}_{0,J_0+1} &= \{(t,\ell)\colon \ell\in\mathcal{O},\ \theta_{0,J_0+1}(\ell)=\tau_\nu(\ell),\ t\in[\theta_{0J_0}(\ell),\theta_{0,J_0+1}(\ell)]\}\\
\mathcal{O}^{\nu}_{10} &= \{(t,\ell)\colon \ell\in\mathcal{O},\ \theta_{0,J_0+1}(\ell)=\tau_\nu(\ell),\ t\in[\theta_{0,J_0+1}(\ell),\theta_{10}(\ell)]\}\\
\mathcal{O}^{\nu}_{ij} &= \{(t,\ell)\colon \ell\in\mathcal{O},\ \theta_{0,J_0+1}(\ell)=\tau_\nu(\ell),\\
&\qquad t\in[\theta_{1,j-1}(\ell),\theta_{1j}(\ell)]\} \qquad j=1,\dots,J_1+1.
\end{aligned}$$

The one–form $\omega := \mathcal{H}^*(p\,\mathrm{d}q - H_t\,\mathrm{d}t)$ is closed on each of these sets, it is continuous on $[0,T]\times\mathcal{O}$ hence it is exact on $[0,T]\times\mathcal{O}$ (without loss of generality we may assume $\mathcal{O}$ to be simply connected) and we have

$$0 = \oint \omega = \int_{\psi^{-1}(\widehat{\Xi})}\omega + \int_{\psi^{-1}(\gamma)}\omega - \int_{\psi^{-1}(\Xi)}\omega.$$

From the maximality properties of H we get

$$\int_{\psi^{-1}(\widehat{\Xi})}\omega = \int_0^T \widehat{f}^0_t(\widehat{\xi}(t))\,\mathrm{d}t \qquad \int_{\psi^{-1}(\Xi)}\omega \le \int_0^T f^0(\xi(t),u(t))\,\mathrm{d}t; \tag{23}$$

so that

$$\int_0^T f^0(\xi(t),u(t))\,\mathrm{d}t - \int_0^T \widehat{f}^0_t(\widehat{\xi}(t))\,\mathrm{d}t \ge \int_{\psi^{-1}(\gamma)}\omega$$

If we now evaluate the difference of the costs associated to the generic pair (ξ,u) and to the reference pair $(\widehat{\xi},\widehat{u})$ we have

$$C(\xi,u) - C(\widehat{\xi},\widehat{u}) \ge \beta(\xi(T)) - \beta(\widehat{x}_T) + \int_{\psi^{-1}(\gamma)}\omega \tag{24}$$

Evaluating this last integral we get

$$\int_{\psi^{-1}(\gamma)} \omega = \alpha(\pi\psi_T^{-1}(\xi(T)) - \alpha(\pi\psi_T^{-1}(\widehat{x}_T)) \\ + \int_0^T f^0(\psi(r, \psi_T^{-1}\xi(T)))\,\mathrm{d}r - \int_0^T f^0(\psi(r, \psi_T^{-1}\widehat{x}_T))\,\mathrm{d}r.$$

Defining $F\colon y \in M \mapsto \alpha(\pi \circ \psi_T^{-1}(y)) + \beta(y) + \int_0^T f^0(\psi(r, \psi_T^{-1}(y)))\,\mathrm{d}r$ equation (24) simplifies to $C(\xi, u) - C(\widehat{\xi}, \widehat{u}) \geq F(\xi(T)) - F(\widehat{x}_T)$ i.e. we have reduced optimal control problem (1) to a finite–dimensional one. Thus in order to prove that $(\widehat{\xi}, \widehat{u})$ is a minimum it now suffices to prove that F has a local minimum in $\widehat{x}_T$.

Theorem 3.1. *F has a strict local minimum in $\widehat{x}_T$.*

Proof. It suffices to prove that

$$\mathrm{d}F(\widehat{x}_T) = 0 \qquad \mathrm{D}^2\,F(\widehat{x}_T) > 0\,. \tag{25}$$

The first equality in (25) is an immediate consequence of the definition of α and PMP. Let us prove that also the second one holds.
Since $\mathrm{d}\Big(\alpha \circ \pi \circ \psi_T^{-1} + \int_0^T f(r, \psi_r \circ \psi_T^{-1})\,\mathrm{d}r\Big) = \mathcal{H}_T \circ \psi_T^{-1}$, we also have

$$\mathrm{d}F = \mathcal{H}_T \circ \psi_T^{-1} + \mathrm{d}\beta \tag{26}$$

$$\begin{aligned} \mathrm{D}^2\,F(\widehat{x}_T)[\delta x_T]^2 &= \big((\mathcal{H}_T \circ \psi_T^{-1})_* + \mathrm{D}^2\,\beta\big)\,(\widehat{x}_T)[\delta x_T]^2 \\ &= \boldsymbol{\sigma}\,\big((\mathcal{H}_T \circ \psi_T^{-1})_*\delta x_T, \mathrm{d}(-\beta)_*\delta x_T\big) \end{aligned} \tag{27}$$

From Lemma 3.4 we get

$$0 < \boldsymbol{\sigma}\Bigg(\Big(0, \delta x + \sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij}g_{ij} + bh_\nu\Big), -\mathrm{D}^2\,\widehat{\gamma}(\delta x, \cdot) + \sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij}\,\overrightarrow{G}''_{ij} + b\overrightarrow{H}''_\nu\Bigg) \tag{28}$$

Applying $\widehat{\mathcal{H}}_{T*} \circ i^{-1}$ we get $0 < \boldsymbol{\sigma}\,(\mathcal{H}_{T*}\,\mathrm{d}\alpha_*\delta x, \mathrm{d}(-\beta)_*(\psi_{T*}\,\mathrm{d}\alpha_*\delta x))$ which is exactly (27) with $\delta x := \pi_*\psi_{T*}^{-1}\delta x_T$. Since $\pi_*\psi_{T*}^{-1}$ is one–to–one, such a choice is always possible. □

To conclude the proof of Theorem 2.1 we have to prove that $\widehat{\xi}$ is a strict minimizer. Assume $C(\xi, u) = C(\widehat{\xi}, \widehat{u})$. Since $\widehat{x}_T$ is a strict minimizer for F, then $\xi(T) = \widehat{x}_T$ and equality must hold in (23):

$$\langle \mathcal{H}_s(\psi_s^{-1}(\xi(s)))\,,\, \dot{\xi}(s)\rangle - f^0(\xi(s), u(s)) = H_s(\mathcal{H}_s(\psi_s^{-1}(\xi(s)))).$$

By regularity assumption, $u(s) = \widehat{u}(s)$ for any s at least in a left neighborhood of T, hence $\xi(s) = \widehat{\xi}(s)$ and $\psi_s^{-1}(\xi(s)) = \widehat{\ell}_0$ for any s in such neighborhood. u takes the value $\widehat{u}_{|(\widehat{\theta}_{1J_1}, T)}$ until $\mathcal{H}_s\psi_s^{-1}(\xi(s)) = \mathcal{H}_s(\widehat{\ell}_0) = \widehat{\lambda}(s)$ hits the hyper-surface $K_{1,J_1} = K_{1,J_1-1}$, which happens at time $s = \widehat{\theta}_{1,J_1}$. At such time, again by regularity assumption, u must switch to $\widehat{u}_{(\widehat{\theta}_{1,J_1-1}, \widehat{\theta}_{1,J_1})}$, so that $\xi(s) = \widehat{\xi}(s)$ also for s in a left neighborhood of $\widehat{\theta}_{1,J_1}$. Proceeding backward in time, with an induction argument we finally get $(\xi(s), u(s)) = (\widehat{\xi}(s), \widehat{u}(s))$ for any $s \in [0, T]$.

4. Appendix: Invertibility of piecewise C^1 maps

The straightforward proof of the following fact is left to the reader.

Lemma 4.1. *Let A and B be linear automorphisms of $\mathbb{R}^n$. Assume that for some $v \in \mathbb{R}^n$, A and B coincide on the space $\{x \in \mathbb{R}^n : \langle x, v\rangle = 0\}$. Then, the map $\mathcal{L}_{AB}$ defined by $x \mapsto Ax$ if $\langle x, v\rangle \geq 0$, and by $x \mapsto Bx$ if $\langle x, v\rangle \leq 0$, is a homeomorphism if and only if* $\det(A) \cdot \det(B) > 0$.

Let $G : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous, piecewise linear map at 0, in the sense that G is continuous and there exists a decomposition $S_1, \dots, S_k$ of $\mathbb{R}^n$ in closed polyhedral cones (intersection of half spaces, hence convex) with common vertex in the origin and such that $\partial S_i \cap \partial S_j = S_i \cap S_j$, $i \neq j$, and linear maps $L_1, \dots, L_k$ with

$$G(x) = L_i x \qquad x \in S_i$$

with $L_i x = L_j x$ for any $x \in S_i \cap S_j$.

It is easily shown that G is proper, and therefore $\deg(G, \mathbb{R}^n, p)$ is well-defined for any $p \in \mathbb{R}^n$ (the construction of [9] is still valid if the assumption on the compactness of the manifolds is replaced with the assumption that G is proper). Moreover $\deg(G, \mathbb{R}^n, p)$ is constant with respect to p. So we shall denote it by $\deg(G)$.

We shall also assume that $\det L_i > 0$ for any $i = 1, \dots, k$.

Lemma 4.2. *If G is as above, then* $\deg(G) > 0$. *In particular, if there exists $q \neq 0$ such that its preimage belongs to at most two of the convex polyhedral cones S_i and $G^{-1}(q)$ is a singleton, then* $\deg(G) = 1$.

Proof. Let us assume in addition that $q \notin \cup_{i=1}^k G\big(\partial S_i\big)$. Observe that the set $\cup_{i=1}^k G\big(\partial S_i\big)$ is nowhere dense hence $A_1 := G(S_1) \setminus \cup_{i=1}^k G\big(\partial S_i\big)$ is non–empty.

Take $x \in A_1$ and observe that if $y \in G^{-1}(x)$ then $y \notin \cup_{i=1}^k \partial S_i$. Thus

$$\deg(G) = \sum_{y \in G^{-1}(x)} \operatorname{sign} \det \mathrm{d}G(y) = \# G^{-1}(x). \tag{29}$$

Since $G^{-1}(x) \neq \emptyset$ the first part of the assertion is proved. The second part of the assertion follows taking $x = q$ in (29).

Let us now remove the additional assumption. Let $\{p\} = G^{-1}(q)$ be such that $p \in \partial S_i \cap \partial S_j$ for some $i \neq j$. Thus one can find a neighborhood V of p, with $V \subset \operatorname{int}(S_i \cup S_j \setminus \{0\})$. By the excision property of the topological degree $\deg(G) = \deg(G, V, p)$. Let $\mathcal{L}_{L_i L_j}$ be a map as in Lemma 4.1. Observe that, the assumption on the signs of the determinants of L_i and L_j imply that $\mathcal{L}_{L_i L_j}$ is orientation preserving. Also notice that $\mathcal{L}_{L_i L_j}|_{\partial V} = G|_{\partial V}$. The multiplicativity, excision and boundary dependence properties of the degree yield $1 = \deg(\mathcal{L}_{L_i L_j}) = \deg(\mathcal{L}_{L_i L_j}, V, p) = \deg(G, V, p)$. Thus, $\deg(G) = 1$, as claimed. $\square$

Let $\sigma_1, \dots, \sigma_r$ be a family of C^1–regular pairwise transversal hyper-surfaces in $\mathbb{R}^n$ with $\cap_{i=1}^r \sigma_i = \{x_0\}$ and let $U \subset \mathbb{R}^n$ be an open and bounded neighborhood of x_0. Clearly, if U is sufficiently small, $U \setminus \cup_{i=1}^r \sigma_i$ is partitioned into a finite number of open sets $U_1, \dots, U_k$.

Let $f : \overline{U} \to \mathbb{R}^n$ be a continuous map such that there exist Fréchet differentiable functions $f_1, \dots, f_k$ in $\overline{U}$ with the property that

$$f(x) = f_i(x) \qquad x \in \overline{U}_i. \tag{30}$$

with $f_i(x) = f_j(x)$ for any $x \in \overline{U}_i \cap \overline{U}_j$. Notice that such a function is $PC^1(\overline{U})$, hence locally Lipschitz continuous (see [7]).

Let $S_1, \dots, S_k$ be the tangent cones at x_0 to the sets $U_1, \dots, U_k$, (by the transversality assumption on the hyper–surfaces σ_i each S_i is a convex polyhedral cone with non empty interior) and assume $\mathrm{d}f_i(x_0)x = \mathrm{d}f_j(x_0)x$ for any $x \in S_i \cap S_j$. Define

$$F(x) = \mathrm{d}f_i(x_0)x \qquad x \in S_i. \tag{31}$$

so that F is a continuous piecewise linear map (compare [7]).

One can see that f is Bouligand differentiable and that its B-derivative is the map F (compare [7, 10]). Let $y_0 := f(x_0)$. There exists a continuous function ε, with $\varepsilon(0) = 0$, such that $f(x) = y_0 + F(x - x_0) + |x - x_0|\varepsilon(x - x_0)$.

Lemma 4.3. *Let f and F be as in* (30)–(31)*, then there exists $\rho > 0$ such that* $\deg\big(f, B(x_0, \rho), y_0\big) = \deg\big(F, B(0, \rho), 0\big)$. *In particular, if* $\det \mathrm{d}f_i(x_0) > 0$*, then F is proper and* $\deg\big(f, B(x_0, \rho), y_0\big) = \deg(F)$.

Proof. Consider the homotopy $H(x, \lambda) = F(x - x_0) + \lambda\,|x - x_0|\,\varepsilon(x - x_0)$, $\lambda \in [0, 1]$ and observe that $m := \inf\{|F(v)| : |v| = 1\} > 0$, F being invertible. Thus,

$$|H(x, \lambda)| \geq \big(m - |\varepsilon(x - x_0)|\big)\,|x - x_0|\,.$$

This shows that in a conveniently small ball centered at x_0, homotopy H is admissible. The assertion follows from the homotopy invariance property of the degree. □

Let f and F be as in (30)–(31) and assume $\det \mathrm{d}f_i(x_0) > 0$. Assume also that there exists $p \in \mathbb{R}^n \setminus \cup_{i=1}^k F(\partial S_i)$ such that $F^{-1}(p)$ is a singleton. From Lemmas 4.2–4.3, it follows that $\deg(f, B(x_0, \rho), y_0) = 1$ for sufficiently small $\rho > 0$. By Theorem 4 in [10], we immediately obtain

Theorem 4.1. *Let f and F be as in* (30)–(31) *and assume* $\det \mathrm{d}f_i(x_0) > 0$*. Assume also that there exists $p \in \mathbb{R}^n \setminus \cup_{i=1}^k F(\partial S_i)$ such that $F^{-1}(p)$ is a singleton. Then f is a Lipschitzian homeomorphism in a sufficiently small neighborhood of x_0.*

References

[1] Andrei A. Agrachev and Revaz V. Gamkrelidze. Symplectic geometry for optimal control. In H. J. Sussmann, editor, *Nonlinear Controllability and Optimal Control*, volume 133 of *Pure and Applied Mathematics*. Marcel Dekker, 1990.

[2] Andrei A. Agrachev and Revaz V. Gamkrelidze. Symplectic methods for optimization and control. In B. Jacubczyk and W. Respondek, editors, *Geometry of Feedback and Optimal Control*, Pure and Applied Mathematics, pages 1–58, New York, New York, 1997. Marcel Dekker.

[3] Andrei A. Agrachev and Yu. L. Sachkov. *Control Theory from the Geometric Viewpoint*. Springer-Verlag, 2004.

[4] Andrei A. Agrachev, Gianna Stefani, and PierLuigi Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control Optimization*, 41(4):991–1014, 2002.

[5] Vladimir I. Arnold. *Mathematical Methods in Classical Mechanics*. Springer, New York, 1980.

[6] B. Koslik and M. H. Breitner. In optimal control problem in economics with four linear control. *J. Optim. Theory and App.*, 94(3):619–634, 1997.

[7] Ludwig Kuntz and Stefan Scholtes. Structural analysis of nonsmooth mappings, inverse functions and metric projections. *Journal of Mathematical Analysis and Applications*, 188:346–386, 1994.

[8] Helmut Maurer and Nikolai P. Osmolovskii. Second order optimality conditions for bangbang control problems. *Control and Cybernetics*, 3(32):555–584, 2003.

[9] John Milnor. *Topology from the Differentiable Viewpoint*. The University Press of Virginia, 1965.

[10] Jong-Shi Pang and Daniel Ralph. Piecewise smoothness, local invertibility, and parametric analysis of normal maps. *Mathematics of Operations Research*, 21(2):401–426, 1996.

[11] Laura Poggiolini. On local state optimality of bang-bang extremals in a free horizon bolza problem. *Rendiconti del Seminario Matematico dell'Universitá e del Politecnico di Torino*, 2006.

[12] Laura Poggiolini and Gianna Stefani. State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *Systems & Control Letters*, 53:269–279, 2004.

[13] Andrei V. Sarychev. First and second order sufficient optimality conditions for bang-bang controls. *SIAM J. Control Optimization*, 1(35):315–340, 1997.

Dipartimento di Matematica Applicata "G. Sansone", Università di Firenze, Italy laura.poggiolini@math.unifi.it and Dipartimento di Matematica Applicata "G. Sansone", Università di Firenze, Italy marco.spadini@math.unifi.it