Rend. Sem. Mat. Univ. Pol. Torino - Vol. 64, 1 (2006)
Control Theory and Stabil., II

L. Poggiolini

ON LOCAL STATE OPTIMALITY OF BANG-BANG EXTREMAL

Abstract. Free horizon Optimal Control Problems are studied where the cost functional is given by

$$C(T, \xi, u) = c_0(\xi(0)) + c_f(\xi(T)) + \int_0^T f^0(\xi(t), u(t))\, dt$$

Sufficient second order conditions are given for the trajectory $\widehat{\xi}$ of a bang-bang regular Pontryagin extremal $(\widehat{T}, \widehat{\xi}, \widehat{u})$ to be state locally optimal.
The control system is control-affine and the controls take values in a polyhedron. The state space and the end points constraints are smooth finite-dimensional manifolds. The hypotheses made concern the positivity of the second variation of the finite-dimensional sub-problem obtained by perturbation of the switching times only and the injectivity of the reference trajectory $\widehat{\xi}$.

1. Introduction

We consider an optimal control problem on a n-dimensional manifold M. The cost to be minimized is in the general Bolza form i.e. it depends on the end-points of the trajectory and on the integral of a running cost depending both on the state $\xi(\cdot)$ and on the control $u(\cdot)$:

$$\text{(1)} \qquad \text{minimize } C(T, \xi, u) = c_0(\xi(0)) + c_f(\xi(T)) + \int_0^T f^0(\xi(t), u(t))\, \mathrm{d}t\,.$$

subject to

$$\text{(2)} \qquad \dot{\xi}(t) = X_0(\xi(t)) + \sum_{k=1}^{m} u_k(t) X_k(\xi(t))\,, \qquad \text{a.e. } t \in (0, T)\,,$$

$$\text{(3)} \qquad \xi(0) \in N_0 \text{ and } \xi(T) \in N_f\,,$$

$$\text{(4)} \qquad u \in L^\infty([0, T], \Delta)\,,$$

where N_0 and N_f are two given sub-manifolds of M, c_0 and c_f are given regular functions defined on the initial manifold N_0 and on the final manifold N_f, respectively, while

$$f^0 \colon (x, u) \in M \times \Delta \to F_0(x) + \sum_{k=1}^{m} u_k F_k(x) \in \mathbb{R}$$

is a control affine regular function where $F_i : M \to \mathbb{R}$, $i = 0, \dots, m$ are given regular (at least C^2) functions. $X_0, X_1, \dots, X_m \in Vec(M)$ are $m + 1$ given regular (at least C^2) vector fields and Δ is a polyhedron of $\mathbb{R}^m$. We explicitly remark that the horizon T is free i.e. it is not a priori given but it is an unknown of the problem.

In this paper we assume we are given an admissible bang-bang trajectory $\widehat{\xi}$ defined on a time interval $[0, \widehat{T}]$ and give sufficient conditions for $\widehat{\xi}$ to be a strong local minimizer. Since the problem is a free horizon one, we must be more precise on the meaning of the expression "local optimality", namely we can distinguish between two kinds of such optimality:

- $\widehat{\xi}$ is optimal with respect to a neighborhood of its graph $\{(t, \widehat{\xi}(t)) : t \in [0, \widehat{T}]\}$ in $\mathbb{R} \times M$ and hence local with respect to both state and final time. We call this type of local optimality *(state, time)-local*.

- $\widehat{\xi}$ is optimal with respect to the trajectories $\xi : [0, T] \to M$ whose range is contained in a neighborhood of the range $\widehat{\Xi} = \{\widehat{\xi}(t) : t \in [0, \widehat{T}]\}$ of $\widehat{\xi}$ in M, whose initial point $\xi(0)$ is contained in a neighborhood of $\widehat{\xi}(0)$ in N_0 and whose final point $\xi(T)$ is contained in a neighborhood of $\widehat{\xi}(\widehat{T})$ in N_f and hence locally only with respect to the state. We call this type of local optimality *state-local*.

Of course (state, time)-local optimality is a weaker condition and in fact it could be proved in less strong hypotheses, namely with a suitable time-reparametrization the problem fits with the results of [1]. An example of a (state, time)-locally optimal trajectory for the minimum time problem ($c_0 \equiv 0$, $c_f \equiv 0$, $f^0 \equiv 1$ in (1)) with fixed end-points ($N_0 = \{x_0\}$, $N_f = \{x_f\}$), associated to a Van der Pool oscillator which is not state-locally optimal was provided in [4].

Another kind of local optimality was studied in [3]. Here the author studies (control, time)-local optimality for the minimum time problem with fixed end-points. Namely the author proves that given a reference control $\widehat{u}$ which steers the initial point to the final point in time $\widehat{T}$, then, under certain assumptions, there exists a positive δ_0 such that no control in a $L^1[0, \widehat{T}]$-neighborhood of $\widehat{u}$ can steer one point to another in time $\widehat{T} - \delta$, for any $\delta \in (0, \delta_0)$. In [2] the authors address the same problem and prove state-local optimality. The hypotheses in [2] are the same we make here for our more general problem. Namely our sufficient optimality conditions involve Pontryagin Maximum Principle, regularity of the maximized Hamiltonian and the positivity of the second variation of a finite-dimensional sub-problem of the given one. Moreover injectivity of the reference trajectory $\widehat{\xi}$ is required.

The choice of our cost functional (1) being so general allows us to recover the minimum time problem (as we already said it suffices to choose $c_0 \equiv 0$, $c_f \equiv 0$, $f^0 \equiv 1$) and the Mayer problem, for which it suffices to choose $f^0 \equiv 0$.

2. Notation and statement of the problem

Let the reference trajectory be defined on the interval $[0, \widehat{T}]$ (to which, henceforth, we shall refer to as to the *reference interval*) and let $\widehat{u} : [0, \widehat{T}] \to \Delta$ be the control associated to $\widehat{\xi}$. We assume $\widehat{u}$ to be bang-bang i.e. there exists a partition of the reference interval

$$0 = \widehat{\tau}_0 < \widehat{\tau}_1 < \cdots < \widehat{\tau}_{r-1} < \widehat{\tau}_r = \widehat{T}$$

such that

$$\widehat{u}(t) = \widehat{u}^i \in \text{Vertexes}(\Delta), \quad \forall\, t \in (\widehat{\tau}_{i-1}, \widehat{\tau}_i), \quad i = 1, \dots, r.$$

For each $i = 1, \dots, r$ let us define the vector field h_i and the function f_i^0 associated to $\widehat{u}^i$:

$$\begin{aligned} & h_i : x \in M \to X_0(x) + \sum_{k=1}^{m} \widehat{u}_k^i X_k(x) \in T_x M \\ & f_i^0 : x \in M \to f^0(x, \widehat{u}^i) = F_0(x) + \sum_{k=1}^{m} \widehat{u}_k^i F_k(x) \in \mathbb{R}, \end{aligned} \qquad i = 1, \dots, r.$$

Moreover we define the time-dependent reference vector field $\widehat{h}$ and the function $\widehat{f}$ in the following way

$$\begin{aligned} & \widehat{h} : (t, x) \in [0, \widehat{T}] \times M \to \widehat{h}_t(x) \equiv h_i(x) \in T_x M \\ & \widehat{f} : (t, x) \in [0, \widehat{T}] \times M \to \widehat{f}_t(x) \equiv f_i^0(x) \in \mathbb{R} \end{aligned} \qquad t \in (\widehat{\tau}_{i-1}, \widehat{\tau}_i) \quad i = 1, \dots, r.$$

With such notation the reference trajectory satisfies the differential equation

$$\dot{\widehat{\xi}}(t) = \widehat{h}_t(\xi(t)), \qquad \text{a.e. } t \in [0, \widehat{T}]. \tag{5}$$

For each $t \in [0, \widehat{T}]$ we denote by $\widehat{S}_t : M \to M$ the flow of (5) at time t.

To each vector field h_i and to each function f_i^0 we associate an Hamiltonian function H_i, depending on a parameter $p_0 \in \{0, 1\}$. If we let $\pi : T^*M \to M$ be the canonical projection, H_i is defined as

$$H_i : \ell \in T^*M \to \langle \ell\,,\ h_i(\pi\ell) \rangle - p_0 f_i^0(\pi\ell) \in \mathbb{R}.$$

We call reference Hamiltonian $\widehat{H}$ the time-dependent Hamiltonian function associated to $\widehat{h}_t$ and to $\widehat{f}_t$

$$\widehat{H} : (t, \ell) \in [0, \widehat{T}] \times T^*M \to \widehat{H}_t(\ell) \equiv \langle \ell\,,\ \widehat{h}_t(\pi\ell) \rangle - p_0 \widehat{f}_t(\pi\ell) \in \mathbb{R}.$$

We are now able to state the first part of our assumptions

ASSUMPTION 1. $(\widehat{\xi}, \widehat{u})$ is a Pontryagin extremal: there exists a lift $\widehat{\lambda}(t)$ of the reference trajectory $\widehat{\xi}(t)$ to the cotangent bundle T^*M such that $\widehat{\lambda}$ is a Lipschitz solution of the Hamiltonian system

$$\dot{\lambda}(t) = \overrightarrow{\widehat{H}}_t(\lambda(t)). \tag{6}$$

Moreover

$$
\begin{aligned}
(7)\qquad & H_i(\widehat{\lambda}(t)) = 0 = \max_{u\in\Delta}\Big\{\langle\widehat{\lambda}(t)\,,\ X_0(\widehat{\xi}(t)) + \sum_{k=1}^{m} u_k X_k(\widehat{\xi}(t))\rangle \\
& \quad - p_0\Big(F_0(\widehat{\xi}(t)) + \sum_{k=1}^{m} u_k F_k(\widehat{\xi}(t))\Big)\Big\} \qquad t\in(\widehat{\tau}_{i-1},\widehat{\tau}_i)\ , \quad i=1,\dots,r\ ,
\end{aligned}
$$

$$(8)\qquad p_0 + \|\widehat{\lambda}(0)\| \neq 0\,,$$

$$(9)\qquad \widehat{\lambda}(0) = p_0\,\mathrm{d}c_0(\widehat{x}_0) \text{ on } T_{\widehat{x}_0}N_0\,,$$

$$(10)\qquad \widehat{\lambda}(\widehat{T}) = -p_0\,\mathrm{d}c_f(\widehat{x}_{\widehat{T}}) \text{ on } T_{\widehat{x}_{\widehat{T}}}N_f\,,$$

where $\widehat{x}_0 = \widehat{\xi}(0)$ and $\widehat{x}_{\widehat{T}} = \widehat{\xi}(\widehat{T})$.

REMARK 1. If, in coordinates, we let $\widehat{\lambda}(t) = \big(\widehat{p}(t), \widehat{\xi}(t)\big)$, equation (6) means that $\widehat{p}(t)$ is a Lipschitz solution to the adjoint equation

$$(11)\quad \dot{p}(t) = \frac{-\partial\widehat{H}}{\partial x}(p(t),\widehat{\xi}(t)) = p_0\mathrm{d}\widehat{f}_t(\widehat{\xi}(t)) - \langle p(t)\,,\ D\widehat{h}_t(\widehat{\xi}(t))\rangle \quad \text{a.e. } t\in[0,\widehat{T}].$$

and we can write $\widehat{p}(t)$ explicitly:

$$\widehat{p}(t) = \Big(\widehat{p}(0) + p_0\int_0^t \mathrm{d}(\widehat{f}_s\circ\widehat{S}_s)(\widehat{x}_0)\,\mathrm{d}s\Big)\,\widehat{S}_{t*}^{-1}\,.$$

ASSUMPTION 2. The extremal is regular:

$$
\begin{aligned}
(12)\qquad & \langle\widehat{\lambda}(t)\,,\ X_0(\widehat{\xi}(t)) + \sum_{k=1}^{m} u_k X_k(\widehat{\xi}(t))\rangle - p_0\Big(F_0(\widehat{\xi}(t)) + \sum_{k=1}^{m} u_k F_k(\widehat{\xi}(t))\Big) < 0 \\
& \qquad\qquad \forall\, u\in\Delta\,, \quad u\neq\widehat{u}\,, \quad t\neq\widehat{\tau}_1,\dots,\widehat{\tau}_r\,.
\end{aligned}
$$

ASSUMPTION 3. All the switching points $\widehat{x}_i \equiv \widehat{\xi}(\widehat{\tau}_i)$, $i = 1,\dots,r-1$, are simple:

$$
\begin{aligned}
(13)\qquad & \langle\widehat{\lambda}(\widehat{\tau}_i)\,,\ X_0(\widehat{x}_i) + \sum_{k=1}^{m} u_k X_k(\widehat{x}_i)\rangle - p_0\,f^0(\widehat{x}_i, u) < 0 \\
& \qquad\qquad \forall\, u\in \text{Vertexes}(\Delta)\,, \quad u\neq\widehat{u}_i\,,\ \widehat{u}_{i+1}\,.
\end{aligned}
$$

If we let $\widehat{\ell}_i \equiv \widehat{\lambda}(\widehat{\tau}_i)$, $i = 1,\dots,r$, from equation (7) in Pontryagin Maximum Principle one easily gets

$$\langle\,\mathrm{d}\,(H_{i+1} - H_i)\,,\ \overrightarrow{H}_{i+1}\rangle(\widehat{\ell}_i) \geq 0 \qquad i = 1,\dots,r-1\,.$$

We strengthen this inequality

ASSUMPTION 4 (Strong bang-bang Legendre condition).

$$\boldsymbol{\sigma}\left(\overrightarrow{H}_i, \overrightarrow{H}_{i+1}\right) = \langle \mathrm{d}(H_{i+1} - H_i)\,,\ \overrightarrow{H}_{i+1}\rangle(\widehat{\ell}_i) > 0 \qquad i = 1, \ldots, r-1. \tag{14}$$

REMARK 2. In coordinates, inequality (14) reads

$$\begin{aligned}&\langle \widehat{p}(\widehat{\tau}_i)\,,\ [h_i, h_{i+1}](\widehat{x}_i)\rangle \\ &\quad - p_0\left(\langle d f_{i+1}^0(\widehat{x}_i)\,,\ h_i(\widehat{x}_i)\rangle - \langle d f_i^0(\widehat{x}_i)\,,\ h_{i+1}(\widehat{x}_i)\rangle\right) > 0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad i = 1, \ldots, r-1\,.\end{aligned}$$

ASSUMPTION 5. The reference trajectory $\widehat{\xi}$ has no self-intersection.

3. A finite-dimensional sub-problem

In order to introduce our last assumption, we now consider the finite-dimensional sub-problem that one gets allowing only piece-wise constant controls having exactly the same number of switches of the reference control $\widehat{u}$

$$0 < \tau_1 < \tau_2 < \cdots < \tau_{r-1} < \tau_r = T$$

and whose value in the i-th interval of this partition of $[0, T]$ is $\widehat{u}^i$, the value of the reference control in the i-th interval of the reference partition $0 < \widehat{\tau}_1 < \widehat{\tau}_2 < \cdots < \widehat{\tau}_{r-1} < \widehat{\tau}_r = \widehat{T}$ of the reference interval $[0, \widehat{T}]$ i.e.

$$u|_{(\tau_{i-1}, \tau_i)} = \widehat{u}^i\,.$$

The initial point of the trajectory is still allowed to be any point of N_0.

The unknowns of this sub-problem are the initial point of the trajectory $\xi(0) \in N_0$ and the lengths $\theta_i \equiv \tau_i - \tau_{i-1}$, $i = 1, \ldots, r$ of each interval of the partition. We call $\theta \equiv (\theta_1, \ldots, \theta_r) \in \mathbb{R}^r_+$ the vector of such lengths.

We have therefore restricted ourselves to consider only state equations of the following kind

$$\dot{\xi}(t) = h_i(\xi(t))\,, \quad t \in (\tau_{i-1}, \tau_i) \quad i = 1, \ldots, r \tag{15}$$

while the minimization problem becomes

$$\begin{aligned}\min\Bigg\{ & c_0(\xi(0)) + c_f(\xi(\tau_r)) + \sum_{i=1}^r \int_{\tau_{i-1}}^{\tau_i} f_i^0(\xi(t))\, \mathrm{d}t : \theta \in \mathbb{R}^r_+\,,\ \xi(0) \in N_0\,, \\ & \qquad \xi(\tau_r) \in N_f\,,\ \tau_i = \sum_{j=1}^i \theta_j\,, \quad i = 1, \ldots, r\Bigg\}.\end{aligned} \tag{16}$$

The reference trajectory is identified by the couple $(\xi(0), \theta) = (\widehat{x}_0, \widehat{\theta})$ where $\widehat{\theta} = \left(\widehat{\theta}_1, \ldots, \widehat{\theta}_r\right)$ and $\widehat{\theta}_i \equiv \widehat{\tau}_i - \widehat{\tau}_{i-1}$, $i = 1, \ldots, r$.

ASSUMPTION 6. The second variation of the finite-dimensional sub-problem (which we shall call *second variation at the switching points*) at $(\widehat{x}_0, \widehat{\theta})$ is positive definite on the linearization of the constraints

$$\mathcal{N} \equiv \left\{ (\delta x, \varepsilon) \in T_{\widehat{x}_0} N_0 \times \mathbb{R}^r : \widehat{S}_{\widehat{T}*}(\delta x, \varepsilon) \in T_{\widehat{x}_{\widehat{T}}} N_f \right\}. \tag{17}$$

We can now state our main result

THEOREM 1. *Assume that the given bang-bang trajectory is not self-intersecting (Assumption 5), it is a regular Pontryagin extremal (Assumptions 1 and 2), it has simple switching points (Assumption 3) and that the strict bang-bang Legendre condition (Assumption 4) holds. Finally assume that the second variation at the switching points is positive definite (Assumption 6). Then $\widehat{\xi}$ is a state-local strict minimizer of the optimal control problem* (1) - (2). *If it is an abnormal extremal, this implies that the reference trajectory is isolated among the solution of* (2) *with constraints* (3) - (4).

4. The second variation

In order to write the second variation at the switching points, it is better to consider both the original problem (1) - (2) and the sub-problem (15) - (16) as equivalent Mayer problems in the state space $\mathbb{R} \times M$. Let us denote by $\widetilde{x} = (x^0, x)$ the elements of $\mathbb{R} \times M$ and by $\widetilde{\ell} = (\widetilde{p}, \widetilde{x}) = (-p^0, p, x^0, x)$ the elements of its cotangent bundle $T^*(\mathbb{R} \times M)$. Consider the vector fields

$$\widetilde{h}_i : \widetilde{x} \in \mathbb{R} \times M \to \begin{pmatrix} f_i^0(x) \\ h_i(x) \end{pmatrix} \in T_{\widetilde{x}}(\mathbb{R} \times M) \qquad i = 1, \dots, r .$$

To each vector field $\widetilde{h}_i$ an Hamiltonian function $\widetilde{H}_i$ is associated. In coordinates:

$$\widetilde{H}_i : \widetilde{\ell} = (-p^0, p, x^0, x) \in T^*(\mathbb{R} \times M) \to \langle \widetilde{\ell}\,,\ \widetilde{h}_i(\pi \widetilde{\ell}) \rangle \in \mathbb{R} \tag{18}$$

PROPOSITION 1. *The minimization problem* (1) - (2) *is equivalent to*

$$\text{minimize } c_0(\xi(0)) + c_f(\xi(T)) + \xi^0(T)$$

subject to

$$\begin{aligned}
&\dot{\widetilde{\xi}}(t) = \begin{pmatrix} \dot{\xi}^0(t) \\ \dot{\xi}(t) \end{pmatrix} = \begin{pmatrix} F_0(\xi(t)) + \sum_{k=1}^m u_k F_k(\xi(t)) \\ X_0(\xi(t)) + \sum_{k=1}^m u_k(t) X_k(\xi(t)) \end{pmatrix}, \ \textit{a.e. } t \in [0, T], \\
&\widetilde{\xi}(0) = \begin{pmatrix} \xi^0(0) \\ \xi(0) \end{pmatrix} \in \{0\} \times N_0, \\
&\widetilde{\xi}(T) = \begin{pmatrix} \xi^0(T) \\ \xi(T) \end{pmatrix} \in \mathbb{R} \times N_f,
\end{aligned} \tag{19}$$

while the minimization sub-problem (15) - (16) *is equivalent to*

$$\textit{minimize } c_0(\xi(0)) + c_f(\xi(\tau_r)) + \xi^0(\tau_r)$$

subject to

$$
\begin{aligned}
&\dot{\widetilde{\xi}}(t) = \widetilde{h}_i(\widetilde{\xi}(t)), \qquad \textit{a.e. } t \in (\tau_{i-1}, \tau_i), \\
&\widetilde{\xi}(0) = \begin{pmatrix} \xi^0(0) \\ \xi(0) \end{pmatrix} \in \{0\} \times N_0, \\
&\widetilde{\xi}(\tau_r) = \begin{pmatrix} \xi^0(\tau_r) \\ \xi(\tau_r) \end{pmatrix} \in \mathbb{R} \times N_f, \\
&\theta \in \mathbb{R}^r_+, \qquad \tau_i = \sum_{j=1}^{i} \theta_i, \qquad i = 1, \ldots, r.
\end{aligned} \tag{20}
$$

The extended reference trajectory $\widehat{\widetilde{\xi}}(t) = (\widehat{\xi}^0(t), \widehat{\xi}(t))$ *is the one associated to the initial point* $(0, \widehat{x}_0)$ *and to the vector* $\widehat{\theta}$*.*

We can restate Assumption 1 as follows:

Assumption 1' $(\widehat{\widetilde{\xi}}, \widehat{u})$ is a Pontryagin extremal: there exists a lift $\widehat{\widetilde{\lambda}}(t)$ to the cotangent bundle $T^*(\mathbb{R} \times M)$ of the reference extended trajectory $\widehat{\widetilde{\xi}}(t)$ such that $\widehat{\widetilde{\lambda}}$ is a Lipschitz solution of the Hamiltonian system

$$\dot{\widetilde{\lambda}}(t) = \overrightarrow{\widetilde{H}_i}(\widetilde{\lambda}(t)) \quad t \in (\widehat{\tau}_{i-1}, \widehat{\tau}_i)$$

such that

$$
\begin{aligned}
\widetilde{H_i}(\widehat{\widetilde{\lambda}}(t)) = 0 = \max_{u \in \Delta} \Bigg\{ &\langle \widehat{p}(t), \ X_0(\widehat{\xi}(t)) + \sum_{k=1}^{m} u_k X_k(\widehat{\xi}(t)) \rangle \\
&- p_0 \left(F_0(\widehat{\xi}(t)) + \sum_{k=1}^{m} u_k F_k(\widehat{\xi}(t)) \right) \Bigg\} \qquad t \in (\widehat{\tau}_{i-1}, \widehat{\tau}_i), \quad i = 1, \ldots, r.
\end{aligned} \tag{21}
$$

$$\|\widehat{\widetilde{\lambda}}(0)\| \neq 0 \tag{22}$$

$$\widehat{\widetilde{\lambda}}(0) = p_0(0, \mathrm{d}c_0(\widehat{x}_0)) \text{ on } T_{(0,\widehat{x}_0)}(\{0\} \times N_0) \tag{23}$$

$$\widehat{\widetilde{\lambda}}(\widehat{T}) = -p_0(1, \mathrm{d}c_f(\widehat{x}_{\widehat{T}})) \text{ on } T_{(\widehat{I},\widehat{x}_{\widehat{T}})}(\mathbb{R} \times N_f) \tag{24}$$

where $\widehat{I} \equiv \int_0^{\widehat{T}} \widehat{f}_t(\widehat{\xi}(t))\,\mathrm{d}t$ is the integral cost on the reference trajectory. Let us denote $S_t(\cdot, \theta)$ the flow of (15) at time t and $\widetilde{S}_t(\widetilde{x}, \theta) = \left(S^0_t(x^0, x, \theta), S_t(x, \theta)\right)$ the extended flow of the state equation in (20), and let

$$\widehat{\widetilde{S}}_t(\widetilde{x}) = \left(\widehat{S}^0_t(x^0, x), \widehat{S}_t(x)\right) = \left(S^0_t(x^0, x, \widehat{\theta}), S_t(x, \widehat{\theta})\right)$$

be the extended reference flow associated to the reference vector $\widehat{\theta}$.

Following [1] and [2] let us define the pull back of the vector fields h_i and $\widetilde{h}_i$ through the reference flows $\widehat{S}_{\widehat{\tau}_{i-1}}$ and $\widehat{\widehat{S}}_{\widehat{\tau}_{i-1}}$, respectively:

$$g_i(x) \equiv \left[D\widehat{S}_{\widehat{\tau}_{i-1}}(x)\right]^{-1} h_i(\widehat{S}_{\widehat{\tau}_{i-1}}(x))\,;$$

$$\begin{aligned}\widetilde{g}_i(x) = \begin{pmatrix} g_i^0(x) \\ g_i(x)\end{pmatrix} &\equiv \left[D\widehat{\widehat{S}}_{\widehat{\tau}_{i-1}}(x)\right]^{-1} \widetilde{h}_i(\widehat{\widehat{S}}_{\widehat{\tau}_{i-1}}(x)) \\ &= \begin{pmatrix} \left(f_i^0 \circ \widehat{S}_{\widehat{\tau}_{i-1}}\right)(x) - L_{g_i}\widehat{S}^0_{\widehat{\tau}_{i-1}}(x) \\ g_i(x)\,. \end{pmatrix}\end{aligned}$$

REMARK 3. $\widehat{S}^0_{\widehat{\tau}_{i-1}}(\widetilde{x}) = x^0 + \sum_{j=1}^{i-1} \int_{\widehat{\tau}_{j-1}}^{\widehat{\tau}_j} f_j^0(\widehat{S}_t(x))\,\mathrm{d}t$ depends on x^0, but its Lie derivative, $L_{g_i}\widehat{S}^0_{\widehat{\tau}_{i-1}}(x)$, does not depend on x^0 hence $\widetilde{g}_i$ actually depends only on x.

If we define $\varepsilon_i = \theta_i - \widehat{\theta}_i$, $i = 1, \ldots, r$ the variation of the length of the i-th interval of the partition and we let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_r)$ be the vector of such variations, it is a straightforward calculation to prove the following proposition (the proof of an analoguous result can be found in section 3 of [1]):

PROPOSITION 2. *The minimization sub-problem is equivalent to*

$$\text{(25)} \qquad \textit{mininimize}\ \ c_0(\zeta(0)) + (c_f \circ \widehat{S}_{\widehat{T}})(\zeta(\widehat{T})) + \widehat{S}^0_{\widehat{T}}(\widetilde{\zeta}(\widehat{T}))$$

subject to

$$\text{(26)} \qquad \begin{aligned} &\dot{\widetilde{\zeta}}(t) = \frac{\varepsilon_i}{\widehat{\theta}_i}\,\widetilde{g}_i(\zeta(t))\,, \quad t \in (\widehat{\tau}_{i-1}, \widehat{\tau}_i), \\ &\widetilde{\zeta}(0) \in \{0\} \times N_0\,, \\ &\widetilde{\zeta}(\widehat{T}) \in \mathbb{R} \times \widehat{S}_{\widehat{T}}^{-1} N_f\,. \end{aligned}$$

Call $\widetilde{\zeta}_t(x, \varepsilon)$ the flow of the state equation in (26)*. Then the extended reference trajectory $\widehat{\widetilde{\xi}}(t)$ corresponds to the constant trajectory $\widehat{\widetilde{\zeta}}(t) = (0, \widehat{x}_0)$ associated to $(\widetilde{x}, \varepsilon) = (x^0, x, \varepsilon) = (0, \widehat{x}_0, 0)$.*

Moreover the linearization of the constraint $\mathcal{N}$, defined in (17)*, can be expressed through the vector fields g_i:*

$$\mathcal{N} = \left\{ (\delta x, \varepsilon) \in T_{\widehat{x}_0}N_0 \times \mathbb{R}^r : \delta x + \sum_{i=1}^r \varepsilon_i g_i(\widehat{x}_0) \in T_{\widehat{x}_0}\widehat{S}_{\widehat{T}}^{-1} N_f \right\}\,.$$

We must write the first and the second variation at the switching points. Inspired by (23) and (24), let $\widetilde{\alpha}\colon \mathbb{R} \times M \to \mathbb{R}$ be a C^∞ extension of $p_0 c_0\colon N_0 \to \mathbb{R}$ such that

$$\left(\frac{\partial \widetilde{\alpha}}{\partial x^0}, \mathrm{d}\widetilde{\alpha}\right)\Big|_{(0,\widehat{x}_0)} = (-p_0, \widehat{p}(0)) \quad \text{on} \quad T_{(0,\widehat{x}_0)}(\mathbb{R} \times M) \simeq \mathbb{R} \times T_{\widehat{x}_0}M\,.$$

We choose $\widetilde{\alpha}$ of the following form

$$\widetilde{\alpha}(\widetilde{x}) = -p_0 x^0 + \alpha(x)$$

where $\alpha \colon M \to \mathbb{R}$ it is a C^∞ extension of $p_0 c_0$ to the state space manifold M and it is such that

$$\mathrm{d}\alpha(\widehat{x}_0) = \widehat{p}(0) \text{ on } T_{\widehat{x}_0} M \,.$$

Analogously let $\widetilde{\beta} \colon \mathbb{R} \times M \to \mathbb{R}$ be a C^∞ extension of $p_0 \left(x^0 + c_f \right) \colon \mathbb{R} \times N_f \to \mathbb{R}$ such that

$$\left(\frac{\partial \widetilde{\beta}}{\partial x^0}, \mathrm{d}\widetilde{\beta} \right) \Bigg|_{(\widehat{\Gamma}, \widehat{x}_{\widehat{T}})} = \left(p_0, \widehat{p}(\widehat{T}) \right) \quad \text{on} \quad T_{(\widehat{\Gamma}, \widehat{x}_{\widehat{T}})}(\mathbb{R} \times M) \simeq \mathbb{R} \times T_{\widehat{x}_{\widehat{T}}} M \,.$$

We choose

$$\widetilde{\beta}(\widetilde{x}) = p_0 x^0 + \beta(x)$$

where $\beta \colon M \to \mathbb{R}$ is a C^∞ extension of $p_0 c_f$ to M such that

$$\mathrm{d}\beta(\widehat{x}_{\widehat{T}}) = -\widehat{p}(\widehat{T}) \text{ on } T_{\widehat{x}_{\widehat{T}}} M \,.$$

Let us define $\widetilde{J} \colon \mathbb{R} \times M \times \mathbb{R}^r_+ \to \mathbb{R}$ as follows

$$\widetilde{J}(x^0, x, \theta) = \alpha(x) + \beta(S_t(x, \theta)) + p_0 \left(x^0 + \sum_{i=1}^{r} \int_{\tau_{i-1}}^{\tau_i} f_i^0(S_t(x, \theta))\, \mathrm{d}t \right)$$

and $J \colon (x, \theta) \in M \times \mathbb{R}^r_+ \to \widetilde{J}(0, x, \theta) \in \mathbb{R}$ where, as usual, $\tau_i = \sum_{j=1}^{i} \theta_j$, $i = 1, \ldots, r$.

PROPOSITION 3. *If the reference trajectory is a normal extremal (i.e.* $p_0 = 1$*) then* $J|_{N_0 \times \mathbb{R}^r_+}$ *coincides with the cost in* (16). *If the reference trajectory is an abnormal extremal (i.e.* $p_0 = 0$*) then* J *is identically* 0.

With $\widehat{\beta} \equiv \beta \circ \widehat{S}_{\widehat{T}}$, we can write

$$J(x, \theta) = \alpha(x) + \widehat{\beta}(\zeta_{\widehat{T}}(x, \varepsilon)) + p_0 \widehat{S}^0_{\widehat{T}}(\widetilde{\zeta}_{\widehat{T}}(0, x, \varepsilon)) \,.$$

Finally let us define

$$\widehat{B}_0 \colon x \in M \to \int_0^{\widehat{T}} \left(\widehat{f}_t \circ \widehat{S}_t \right)(x)\, \mathrm{d}t \in \mathbb{R}.$$

From (9) and (10) in Assumption 1 (or, equivalently, from (23) and (24) in Assumption 1') and the choice of α, $\widehat{\beta}$, $\widehat{B}_0$ we get

$$\mathrm{d}(\alpha + \widehat{\beta} + p_0 \widehat{B}_0)(\widehat{x}_0) = 0.$$

Let us define $\widehat{\gamma}(x) \equiv \alpha(x) + \widehat{\beta}(x) + p_0 \widehat{B}_0(x)$ so that $D^2 \widehat{\gamma}(\widehat{x}_0)$ is a well defined bilinear functional on $T_{\widehat{x}_0} M$.

PROPOSITION 4. *For any $\delta e = (\delta x, \varepsilon) \in \mathcal{N}$ the second variation at the switching points is*

$$
\begin{aligned}
J''[\delta e]^2 &= \frac{1}{2}\mathrm{D}^2\widehat{\gamma}(\delta x, \delta x) \\
&+ \sum_{j=1}^{r} \varepsilon_j \left[L_{\delta x} L_{g_j} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{j-1}} \right) \right)(\widehat{x}_0) + p_0 \langle \mathrm{d}(f_j^0 \circ \widehat{S}_{\widehat{\tau}_{j-1}})(\widehat{x}_0)\,,\ \delta x \rangle \right] \\
&+ \frac{1}{2} \sum_{j=1}^{r} \sum_{i=1}^{r} \varepsilon_i \varepsilon_j \Big[L_{g_j} L_{g_i} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{i-1}} \right) \right)(\widehat{x}_0) \\
&+ p_0 \langle \mathrm{d}(f_i^0 \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0)\,,\ g_j(\widehat{x}_0) \rangle \Big] + \frac{1}{2} \sum_{j=1}^{r} \sum_{i=1}^{j-1} \varepsilon_i \varepsilon_j \Big[L_{[g_i\,,\, g_j]} \left(\widehat{\beta} + p_0 \widehat{B}_0 \right)(\widehat{x}_0) \\
&+ p_0 \Big(L_{g_j} L_{g_i} \widehat{S}^0_{\widehat{\tau}_{i-1}}(\widehat{x}_0) - L_{g_i} L_{g_j} \widehat{S}^0_{\widehat{\tau}_{j-1}}(\widehat{x}_0) \\
&+ \langle \mathrm{d}(f_j^0 \circ \widehat{S}_{\widehat{\tau}_{j-1}})(\widehat{x}_0)\,,\ g_i(\widehat{x}_0) \rangle - \langle \mathrm{d}(f_i^0 \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0)\,,\ g_j(\widehat{x}_0) \rangle \Big) \Big]\,.
\end{aligned}
$$

REMARK 4. In the particular case $f^0 = f^0(u)$ (which includes the two model cases: the Mayer problem $f^0 \equiv 0$ and the minimum time problem $f^0 \equiv 1$), both $\widehat{B}_0$ and $\widehat{S}^0_{\widehat{\tau}_i}$ do not depend on x so the expression of J'' simplifies to

$$
\begin{aligned}
J''[\delta e]^2 &= \frac{1}{2}\mathrm{D}^2\widehat{\gamma}(\delta x, \delta x) + \sum_{j=1}^{r} \varepsilon_j L_{\delta x} L_{g_j} \widehat{\beta}(\widehat{x}_0) \\
&\quad + \frac{1}{2} \sum_{j=1}^{r} \sum_{i=1}^{r} \varepsilon_i \varepsilon_j L_{g_j} L_{g_i} \widehat{\beta}(\widehat{x}_0) + \frac{1}{2} \sum_{j=1}^{r} \sum_{i=1}^{j-1} \varepsilon_i \varepsilon_j L_{[g_i\,,\, g_j]} \widehat{\beta}(\widehat{x}_0)
\end{aligned}
$$

LEMMA 1. *The bilinear quadratic form associated to the second variation at the switching points (and which we still denote J'') is*

$$
\begin{aligned}
J''(\delta e\,,\ \delta f) &= \mathrm{D}^2\widehat{\gamma}(\delta x, \delta y) \\
&+ \sum_{i=1}^{r} \varepsilon_i \left[L_{\delta y} L_{g_i} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{i-1}} \right) \right)(\widehat{x}_0) + p_0 \langle \mathrm{d}(f_i^0 \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0)\,,\ \delta y \rangle \right] \\
&+ \sum_{j=1}^{r} \eta_j \left[L_{\delta x} L_{g_j} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{j-1}} \right) \right)(\widehat{x}_0) + p_0 \langle \mathrm{d}(f_j^0 \circ \widehat{S}_{\widehat{\tau}_{j-1}})(\widehat{x}_0)\,,\ \delta x \rangle \right] \\
&+ \sum_{j=1}^{r} \sum_{i=1}^{r} \varepsilon_i \eta_j \Big[L_{g_j} L_{g_i} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{i-1}} \right) \right)(\widehat{x}_0) \\
&+ p_0 \langle \mathrm{d}(f_i^0 \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0)\,,\ g_j(\widehat{x}_0) \rangle \Big] + \sum_{j=1}^{r} \sum_{i=1}^{j-1} \varepsilon_i \eta_j \left[L_{[g_i\,,\, g_j]} \left(\widehat{\beta} + p_0 \widehat{B}_0 \right)(\widehat{x}_0) \right.
\end{aligned}
$$

$$+ p_0 \Big(L_{g_j} L_{g_i} \widehat{S}^0_{\widehat{\tau}_{i-1}}(\widehat{x}_0) \; - L_{g_i} L_{g_j} \widehat{S}^0_{\widehat{\tau}_{j-1}}(\widehat{x}_0)$$

$$+ \langle \mathrm{d}(f^0_j \circ \widehat{S}_{\widehat{\tau}_{j-1}})(\widehat{x}_0) \,,\; g_i(\widehat{x}_0)\rangle \; - \langle \mathrm{d}(f^0_i \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0) \,,\; g_j(\widehat{x}_0)\rangle \Big) \Big] \,.$$

for any $\delta e = (\delta x, \varepsilon)$, $\delta f = (\delta y, \eta) \in \mathcal{N}$.

If we define the linear Hamiltonian functions

$$\begin{aligned} G''_i \colon (\omega, \delta x) \in & T^*_{\widehat{x}_0} M \times T_{\widehat{x}_0} M \; \to \\ \to \quad & \langle \omega \,,\; g_i(\widehat{x}_0)\rangle + L_{\delta x} L_{g_i} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{i-1}} \right)\right)(\widehat{x}_0) \\ & + p_0 \langle \mathrm{d}(f^0_i \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0) \,,\; \delta x \rangle \in \mathbb{R}, \end{aligned}$$

to each G''_i a constant Hamiltonian vector field $\overrightarrow{G}''_i$ is associated, namely

$$\overrightarrow{G}''_i = \left(-L_{(\cdot)} L_{g_i} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{i-1}} \right)\right)(\widehat{x}_0) - p_0 \mathrm{d}(f^0_i \circ \widehat{S}_{\widehat{\tau}_{i-1}})(\widehat{x}_0), g_i(\widehat{x}_0) \right) \,.$$

We define the subspaces of the initial and final transversality conditions as

$$\begin{aligned} L''_0 \equiv \Big\{ \delta\ell = & \left(-\mathrm{D}^2 \widehat{\gamma}(\widehat{x}_0)(\delta x, \cdot) + \omega, \delta x \right) : \delta x \in T_{\widehat{x}_0} N_0 \,, \\ & \omega \in \left(T_{\widehat{x}_0} N_0 \right)^{\perp} \text{ and } \boldsymbol{\sigma}(\delta\ell, \overrightarrow{G}''_1) = 0 \Big\} \,, \\ L''_{\widehat{T}} \equiv & \left(T_{\widehat{x}_0} \widehat{S}^{-1}_{\widehat{T}}(N_f) \right)^{\perp} \times T_{\widehat{x}_0} \widehat{S}^{-1}_{\widehat{T}}(N_f). \end{aligned}$$

For any $\delta e = (\delta x, \varepsilon) \in \mathcal{N}$ let ω_0 be such that $\delta\ell = (\omega_0, \delta x) \in L''_0$, and let

$$\delta\ell_{\widehat{T}}(\omega_0, \delta e) = (\omega_{\widehat{T}}(\omega_0, \delta e), \delta x_{\widehat{T}}(\delta e)) = \delta\ell + \sum_{i=1}^{r} \varepsilon_i \overrightarrow{G}''_i.$$

PROPOSITION 5. *With the notation established above*

$$\begin{aligned} J''(\delta e \,,\; \delta f) = \mathrm{D}^2 \widehat{\gamma}(\delta x, \delta y) - \langle \omega_{\widehat{T}}(\omega_0, \delta e) \,,\; \delta y + \sum_{j=1}^{r} \eta_j g_j(\widehat{x}_0)\rangle + \langle \omega_0 \,,\; \delta y \rangle \\ + \sum_{j=1}^{r} \eta_j G''_j \left(\delta\ell + \sum_{i=1}^{j-1} \varepsilon_i \overrightarrow{G}''_i \right) \qquad \forall \, \delta e = (\delta x, \varepsilon) \,,\; \delta f = (\delta y, \eta) \in \mathcal{N} \,. \end{aligned}$$

We want to gain some piece of information from the positivity of J'' on $\mathcal{N}$ (Assumption 6). To do so we study its meaning on an increasing sequence of subspaces of $\mathcal{N}$. Let us define

$$\begin{aligned} V & \equiv \left\{ (\delta x, \varepsilon) \in \mathcal{N} \colon \delta x + \sum_{i=1}^{r} \varepsilon_i g_i(\widehat{x}_0) = 0 \right\} , \\ V_k & \equiv \{ (\delta x, \varepsilon) \in V \colon \varepsilon_i = 0 \quad \forall \, i = k+1, \ldots, r \} \quad \; k = 1, \ldots, r \,. \end{aligned}$$

If W is any subspace of $\mathcal{N}$, we denote by $W^{\perp}$ the orthogonal sub-space to W with respect to J''.

LEMMA 2. *$\delta e = (\delta x, \varepsilon) \in V_k \cap V_{k-1}^{\perp}$ if and only if*

1. *$\delta x \in T_{\widehat{x}_0} N_0$;*
2. *$\delta x + \sum_{i=1}^{k} \varepsilon_i g_i(\widehat{x}_0) = 0$ and $\varepsilon_i = 0 \quad \forall\, i = k+1, \dots, r$;*
3. *for any $\omega_0 \in (T_{\widehat{x}_0} N_0)^{\perp}$, $\delta\ell = (-D^2\widehat{\gamma}(\widehat{x}_0)(\delta x, \cdot) + \omega_0, \delta x)$ is in L_0'' and,*

$$G_j''(\delta\ell + \sum_{i=1}^{j-1} \varepsilon_i \overrightarrow{G}_i'') = 0\,, \quad \forall\, j = 2, \dots, k-1. \tag{27}$$

In this case $J''[\delta e]^2 = \varepsilon_k G_k''(\delta\ell + \sum_{i=1}^{k-1} \varepsilon_i \overrightarrow{G}_i'')$.

LEMMA 3. *$\delta e \in \mathcal{N} \cap V^{\perp}$ if and only if*

1. *$\delta x \in T_{\widehat{x}_0} N_0$;*
2. *$\delta x + \sum_{i=1}^{r} \varepsilon_i g_i(\widehat{x}_0) \in T_{\widehat{x}_0} \widehat{S}_{\widehat{T}}^{-1} N_f$;*
3. *$\delta\ell = (-D^2\widehat{\gamma}(\widehat{x}_0)(\delta x, \cdot) + \omega, \delta x) \in L_0''$ and $G_j''(\delta\ell + \sum_{i=1}^{j-1} \varepsilon_i \overrightarrow{G}_i'') = 0$, for every $j = 2, \dots, r$.*

In this case

$$J''[\delta e]^2 = -\Big\langle \omega_0 - \sum_{i=1}^{r} \varepsilon_i \left(L_{(\cdot)} L_{g_i} \left(\widehat{\beta} + p_0 \left(\widehat{B}_0 - \widehat{S}^0_{\widehat{\tau}_{i-1}} \right) \right) (\widehat{x}_0) \right.$$
$$\left. + \; p_0 \mathrm{d}(f_i^0 \circ \widehat{S}^0_{\widehat{\tau}_{i-1}})(\widehat{x}_0) \right)\,,\ \delta x + \sum_{j=1}^{r} \varepsilon_j g_j(\widehat{x}_0) \Big\rangle.$$

For a proof of Lemma 2 and Lemma 3 see Lemma 3.1 in [1].

5. Discrete Jacobi system

For $(\delta p, \delta x) \in T^*_{\widehat{x}_0} M \times T_{\widehat{x}_0} M$ we recursively define

$$\mathcal{G}_0''(\delta p, \delta x) = (\delta p, \delta x),$$
$$\omega_i(\delta p, \delta x) = \frac{-\boldsymbol{\sigma}\left(\mathcal{G}_{i-1}''(\delta p, \delta x), \overrightarrow{G}_{i+1}'' \right)}{\boldsymbol{\sigma}\left(\overrightarrow{G}_i'', \overrightarrow{G}_{i+1}'' \right)}\,, \qquad i = 1, \dots, r-1,$$
$$\mathcal{G}_i''(\delta p, \delta x) = \mathcal{G}_{i-1}''(\delta p, \delta x) + \omega_i(\delta p, \delta x) \overrightarrow{G}_i''\,, \qquad i = 1, \dots, r-1\,.$$

REMARK 5. From the very definition the following equality follows:

$$\boldsymbol{\sigma}\left(\mathcal{G}''_i(\delta p, \delta x),\ \overrightarrow{G}''_{i+1}\right) = 0\,.$$

PROPOSITION 6. *$(\delta p, \delta x) \in L''_0$ and equations (27), $j = 2, \ldots, k-1$, are satisfied if and only if*

$$\boldsymbol{\sigma}\left((\delta p, \delta x),\ \overrightarrow{G}''_1\right) = 0 \quad \textit{and} \quad \varepsilon_i = \omega_i(\delta p, \delta x) \quad \forall\, i = 1, \ldots, k-2\,.$$

More precisely we have

1. *$\delta e \in V_1$ if and only if*

$$\begin{aligned}
&\delta x + \varepsilon_1 g_1(\widehat{x_0}) = 0\,, \quad \varepsilon_i = 0\,, \quad \forall\, i = 2, \ldots, r,\\
&\delta x \in T_{\widehat{x_0}} N_0,\\
&\exists\, \delta p \ \ \textit{such that} \ \ (\delta p, \delta x) \in L''_0\,.
\end{aligned}$$

   In this case

$$J''[\delta e]^2 = \varepsilon_1 G''_1(\delta \ell) \tag{28}$$

2. *For $k = 2, \ldots, r$, $\delta e \in V_k \cap V_{k-1}^{\perp}$ if and only if*

$$\begin{aligned}
&\delta x + \sum_{i=1}^{k} \varepsilon_i g_i(\widehat{x_0}) = 0\,, \quad \varepsilon_i = 0\,, \quad \forall\, i = k+1, \ldots, r,\\
&\delta x \in T_{\widehat{x_0}} N_0,\\
&\exists\, \delta p \ \ \textit{such that} \ \ (\delta p, \delta x) \in L''_0\,, \quad \varepsilon_i = \omega_i(\delta p, \delta x)\,, \quad \forall\, i = 1, \ldots, k-2,
\end{aligned}$$

   and in this case

$$(\delta p, \delta x) + \sum_{i=1}^{k} \varepsilon_i \overrightarrow{G}''_i = \mathcal{G}''_{k-1}(\delta p, \delta x) + (\varepsilon_{k-1} - \omega_{k-1}(\delta p, \delta x))\overrightarrow{G}''_{k-1} + \varepsilon_k \overrightarrow{G}''_k$$

$$J''[\delta e]^2 = \varepsilon_k\,(\varepsilon_{k-1} - \omega_{k-1}(\delta p, \delta x))\,\boldsymbol{\sigma}\left(\overrightarrow{G}''_{k-1}\,\overrightarrow{G}''_k\right). \tag{29}$$

3. *$\delta e \in V_k \cap V_k^{\perp}$ if and only if*

$$\begin{aligned}
&\delta x + \sum_{i=1}^{k} \varepsilon_i g_i(\widehat{x_0}) = 0\,, \quad \varepsilon_i = 0\,, \quad \forall\, i = k+1, \ldots, r,\\
&\delta x \in T_{\widehat{x_0}} N_0,\\
&\exists\, \delta p \ \ \textit{such that} \ \ (\delta p, \delta x) \in L''_0\,, \quad \varepsilon_i = \omega_i(\delta p, \delta x)\,, \quad \forall\, i = 1, \ldots, k-1\,.
\end{aligned}$$

   In this case

$$(\delta p, \delta x) + \sum_{i=1}^{k} \varepsilon_i \overrightarrow{G}''_i = \mathcal{G}''_{k-1}(\delta p, \delta x) + \varepsilon_k \overrightarrow{G}''_k$$

4. $\delta e \in \mathcal{N} \cap V^{\perp}$ *if and only if*

$$\delta x \in T_{\widehat{x}_0} N_0 \quad e \quad \delta x + \sum_{i=1}^{r} \varepsilon_i g_i(\widehat{x}_0) \in T_{\widehat{x}_0} \widehat{S}_{\widehat{T}}^{-1} N_f$$
$$\exists\, \delta p \ \text{ such that } \ (\delta p, \delta x) \in L_0'', \quad \varepsilon_i = \omega_i(\delta p, \delta x)\,, \quad \forall\, i = 1, \dots, r-1\,.$$

In this case

$$(\delta p, \delta x) + \sum_{i=1}^{r} \varepsilon_i \overrightarrow{G}_i'' = \mathcal{G}_{r-1}''(\delta p, \delta x) + \varepsilon_r \overrightarrow{G}_r''$$

and, for any $\omega_{\widehat{T}} \in \left(T_{\widehat{x}_0} \widehat{S}_{\widehat{T}}^{-1} N_f\right)^{\perp}$

$$J''[\delta e]^2 = \boldsymbol{\sigma}\left(\left(\omega_{\widehat{T}}, \pi_*\left(\mathcal{G}_{r-1}''(\delta p, \delta x) + \varepsilon_r \overrightarrow{G}_r''\right)\right), \mathcal{G}_{r-1}''(\delta p, \delta x) + \varepsilon_r \overrightarrow{G}_r''\right) \tag{30}$$

PROPOSITION 7. *If* $J''|_V > 0$*, then* $J''|_{\mathcal{N} \cap V^{\perp}} > 0$ *if and only if for any* $(\delta p, \delta x) \in L_0''$*, any* $\delta t \in \mathbb{R}$ *such that* $\pi_*\left(\mathcal{G}_{r-1}''(\delta p, \delta x) + \delta t \overrightarrow{G}_r''\right)$ *is different from zero and is in* $\pi_* L_{\widehat{T}}''$ *and for any* $(\delta p_{\widehat{T}}, \delta x_{\widehat{T}}) \in L_{\widehat{T}}''$ *such that* $\delta x_{\widehat{T}} \equiv \pi_*(\delta p_{\widehat{T}}, \delta x_{\widehat{T}})$ $= \pi_*\left(\mathcal{G}_{r-1}''(\delta p, \delta x) + \delta t \overrightarrow{G}_r''\right)$ *we have*

$$\boldsymbol{\sigma}\left((\delta p_{\widehat{T}}, \delta x_{\widehat{T}}), \mathcal{G}_{r-1}''(\delta p, \delta x) + \delta t \overrightarrow{G}_r''\right) > 0\,. \tag{31}$$

Proof. 1. Assume $J''|_{\mathcal{N} \cap V^{\perp}} > 0$ and let $(\delta p, \delta x) \in L_0''$, $(\delta p_{\widehat{T}}, \delta x_{\widehat{T}}) \in L_{\widehat{T}}''$ and $\delta t \in \mathbb{R}$ satisfy the hypotheses.
Then $\delta e \equiv (\pi_*(\delta p, \delta x), \omega_1(\delta p, \delta x), \dots, \omega_{r-1}(\delta p, \delta x), \delta t)$ is in $\mathcal{N} \cap V^{\perp}$ and $J''[\delta e]^2$ is strictly positive. Moreover $\boldsymbol{\sigma}\left((\delta p_{\widehat{T}}, \delta x_{\widehat{T}}), \mathcal{G}_{r-1}''(\delta p, \delta x) + \delta t \overrightarrow{G}_r''\right) = J''[\delta e]^2 > 0$
2. Vice versa: assume $\boldsymbol{\sigma}\left((\delta p_{\widehat{T}}, \delta x_{\widehat{T}}), \mathcal{G}_{r-1}''(\delta p, \delta x) + \delta t \overrightarrow{G}_r''\right) > 0$ for any $(\delta p, \delta x)$ $\in L_0''$, $\delta t \in \mathbb{R}$ and $(\delta p_{\widehat{T}}, \delta x_{\widehat{T}}) \in L_{\widehat{T}}''$ satisfying our hypotheses.
Let $\delta e = (\delta x, \varepsilon_1, \dots, \varepsilon_r) \in \mathcal{N} \cap V^{\perp}$. Since $J''|_V > 0$, there exists $(\delta p, \delta x) \in L_0''$ such that $\pi_*(\delta p, \delta x) = \delta x$ and $\delta e = (\delta x, \omega_1(\delta p, \delta x), \dots, \omega_{r-1}(\delta p, \delta x), \varepsilon_r)$. We now use (30) with $\omega_{\widehat{T}}$ such that $\left(\omega_{\widehat{T}}, \pi_*\left(\mathcal{G}_{r-1}''(\delta p, \delta x) + \varepsilon_r \overrightarrow{G}_r''\right)\right) = (\delta p_{\widehat{T}}, \delta x_{\widehat{T}})$ and we are done. $\square$

6. Hamiltonian methods

Let

$$\Sigma_0 \equiv \left\{\ell = (p, x) \in T^*M \colon x \in N_0\,, \quad H_1(\ell) = 0\,, \quad p \in \mathrm{d}\alpha(x) + (T_x N_0)^{\perp}\right\}$$

and

$$\Omega \equiv \left\{(t, \ell) \colon t \in [-\delta, \widehat{T} + \delta]\,, \quad \ell \in \Sigma_0\right\}\,.$$

We want to define an one-to-one map from a neighborhood of $[0, \widehat{T}] \times \{\widehat{\ell_0}\}$ in Ω to a neighborhood of $\widehat{\Xi}$ in M. It is evident that Ω may not be "big" enough. More precisely its dimension may be lower than the dimension of M if N_0 is a proper sub-manifold of M. To overcome this difficulty we define an equivalent problem with free initial point.

Let $Q\colon T_{\widehat{x}_0}M \times T_{\widehat{x}_0}M \to \mathbb{R}$ be a positive semi-definite quadratic form whose kernel is $T_{\widehat{x}_0}N_0$. Since, by assumption, J'' is positive definite on $\mathcal{N}$, there exists $\rho > 0$ such that $J''[\delta x, \varepsilon]^2 + \rho Q[\delta x]^2$ is positive definite on

$$\mathcal{N}_\rho \equiv \left\{(\delta x, \varepsilon) \in T_{\widehat{x}_0}M \times \mathbb{R}^r : S_{\widehat{T}*}(\delta x, \varepsilon) \in T_{\widehat{x}_{\widehat{T}}} N_f\right\} .$$

Let $\alpha_\rho\colon M \to \mathbb{R}$ be a C^∞ function such that

1. $\alpha_\rho = \alpha$ on N_0;
2. $\mathrm{d}\alpha_\rho = \mathrm{d}\alpha$ on $T_{\widehat{x}_0}M$;
3. $\mathrm{D}^2\alpha_\rho(\widehat{x}_0) = \mathrm{D}^2\alpha(\widehat{x}_0) + \rho Q$

and study the minimization problem

$$\text{minimize } \alpha_\rho(\xi(0)) + \beta(\xi(T_f)) + p_0 \int_0^T f^0(\xi(t), u(t))\,\mathrm{d}t$$

with the constraints

$$\begin{aligned}
&\dot{\xi}(t) = X_0(\xi(t)) + \sum_{k=1}^m u_k(t) X_k(\xi(t)) , \qquad \text{a.e. } t \in (0, T) , \\
&\xi(T) \in N_f , \ u \in L^\infty([0, T], \Delta) .
\end{aligned}$$

If we prove that the reference trajectory $\widehat{\xi}$ is optimal for this new problem then it is also optimal for problem (1) - (2) since it satisfies the additional constraint $\xi(0) \in N_0$. Therefore, without loss of generality, we can assume $N_0 = M$, $\alpha = \alpha_\rho$ and $\mathcal{N} = \mathcal{N}_\rho$. Also Σ_0 must be redefined accordingly, so from now on

$$\begin{aligned}
&\Sigma_0 \equiv \left\{\ell \in T^*M\colon \ell = (\mathrm{d}\alpha(x), x) , \ x \in M , \ H_1(\ell) = 0\right\} , \\
&\pi\,\Sigma_0 \equiv \left\{x \in M\colon \left(L_{h_1}\alpha - p_0 f_1^0\right)(x) = 0\right\}
\end{aligned}$$

From the positivity of J'' on V_1 we immediately get:

PROPOSITION 8. *Σ_0 and its projection $\pi\,\Sigma_0$ on the manifold M are $(n-1)$-dimensional. More precisely*

$$\begin{aligned}
T_{\widehat{\ell}_0}\Sigma_0 &= \left\{\mathrm{d}\alpha_*\delta x \in T_{\widehat{\ell}_0}\left(T^*M\right) : L_{\delta x}\left(L_{h_1}\alpha - p_0 f_1^0\right)(\widehat{x}_0) = 0\right\} , \\
T_{\widehat{x}_0}\pi\,\Sigma_0 &= \left\{\delta x \in T_{\widehat{x}_0}M\colon L_{\delta x}\left(L_{h_1}\alpha - p_0 f_1^0\right)(\widehat{x}_0) = 0\right\}
\end{aligned}$$

and $h_1(\widehat{x}_0) = g_1(\widehat{x}_0) \notin T_{\widehat{x}_0}\pi\,\Sigma_0$.

Proof. Σ_0 is defined as the intersection of the Lagrangian sub-manifold

$$\Phi_\alpha \equiv \{\ell \in T^*M \colon \ell = (\mathrm{d}\alpha(x), x)\,,\; x \in M\}$$

with the manifold

$$\Phi_1 = \{\ell \in T^*M \colon H_1(\ell) = 0\}.$$

We have to prove that they intersect transversally in $\widehat{\ell}_0$. Their tangent spaces are given, respectively, by

$$T_{\widehat{\ell}_0}\Phi_\alpha = \{d\alpha_* \delta x \colon \delta x \in T_{\widehat{x}_0}M\}\,,$$
$$T_{\widehat{\ell}_0}\Phi_1 = \{\delta\ell \colon \langle \delta\ell\,,\; h_1(\widehat{x}_0)\rangle + \Big(d\alpha(\widehat{x}_0)Dh_1(\widehat{x}_0) - p_0 d f_1^0(\widehat{x}_0)\Big)\,\pi_*\delta\ell = 0\}\,.$$

Hence

$$T_{\widehat{\ell}_0}\Phi_\alpha \cap T_{\widehat{\ell}_0}\Phi_1 = \{\delta\ell = d\alpha_*\delta x \colon \delta x \in T_{\widehat{x}_0}M\,,\; L_{\delta x}\left(L_{h_1}\alpha - p_0 f_1^0\right)(\widehat{x}_0) = 0\}$$

Equation (28) yields

$$0 < J''[(g_1(\widehat{x}_0), -1)]^2 = L_{g_1}\left(L_{g_1}\alpha - p_0 f_1^0\right)(\widehat{x}_0)$$

so, since $h_1 = g_1$, we have that $T_{\widehat{\ell}_0}\Phi_\alpha \cap T_{\widehat{\ell}_0}\Phi_1$ is $(n-1)$-dimensional and its projection

$$\pi_*(T_{\widehat{\ell}_0}\Phi_\alpha \cap T_{\widehat{\ell}_0}\Phi_1) = \{\delta x \in T_{\widehat{x}_0}M \colon L_{\delta x}\left(L_{h_1}\alpha - p_0 f_1^0\right)(\widehat{x}_0) = 0\}$$

is also $(n-1)$-dimensional i.e. $\pi\,\Sigma_0$ is a $(n-1)$-dimensional sub-manifold of M. □

6.1. The maximized flow

For any ℓ in a sufficiently small neighborhood $\mathcal{I}$ of $\widehat{\ell}_0$ in Σ_0 and $k = 1, \dots, r-1$, let us recursively define

$$\varphi_0(\ell) = \ell\,;$$
$$\theta_k(\ell)\colon \begin{cases} \theta_k(\widehat{\ell}_0) = \widehat{\theta}_k\,; \\ H_{k+1}\left(\exp\left(\theta_k(\ell)\overrightarrow{H}_k\right)\varphi_{k-1}(\ell)\right) = 0\,; \end{cases} \tag{32}$$
$$\varphi_k(\ell) = \exp\left(\theta_k(\ell)\overrightarrow{H}_k\right)\varphi_{k-1}(\ell)$$

and let

$$\tau_k(\ell) = \sum_{i=1}^{k}\theta_i(\ell).$$

Therefore the maximized Hamiltonian flow $\mathcal{H}$ can be defined on a neighborhood $\mathcal{O}$ of $[0, \widehat{T}] \times \{\widehat{\ell}_0\}$ in $\mathbb{R} \times \Sigma_0$ by the recursive rule

$$\mathcal{H}_t(\ell) = \begin{cases} \exp\left(t\overrightarrow{H}_k\right)(\ell) & t \in (\tau_0(\ell), \tau_1(\ell)), \\ \exp\left((t - \tau_{k-1}(\ell))\overrightarrow{H}_k\right)\varphi_{k-1}(\ell) & t \in (\tau_{k-1}(\ell), \tau_k(\ell))\,, \\ & k = 2, \dots, r, \end{cases} \tag{33}$$

where, by definition $\tau_0(\ell) = -\delta$, $\tau_r(\ell) = \widehat{T} + \delta$ for some sufficiently small positive δ. For $k = 1, \ldots, r$, let

$$\Sigma_k \equiv \{(\tau_k(\ell), \ell) \colon \ell \in \mathcal{I}\},$$
$$\Omega_k \equiv \{(t, \ell) \colon \ell \in \mathcal{I}, \quad \tau_{k-1}(\ell) < t < \tau_k(\ell)\}.$$

In order to link the results of the previous section to the maximized Hamiltonian flow $\mathcal{H}$ we define an isomorphism from $T^*_{\widehat{x}_0} M \times T_{\widehat{x}_0} M$ to $T_{\widehat{\ell}_0}(T^*M)$.

DEFINITION 1. *Define* $i \colon T^*_{\widehat{x}_0} M \times T_{\widehat{x}_0} M \to T(T^*M)$ *as the linear application*

$$i \colon (\delta p, \delta x) \to -\delta p + \mathrm{d}(-p_0 \widehat{B}_0 - \widehat{\beta})_* \delta x.$$

PROPERTY 1. 1. i is an antisymplectic isomorphism;

2. $i(\overrightarrow{G}''_i) = \widehat{\mathcal{H}}^{-1}_{\widehat{\tau}_i *} \overrightarrow{H}_i(\widehat{\ell}_i) = \widehat{\mathcal{H}}^{-1}_{\widehat{\tau}_{i-1} *} \overrightarrow{H}_i(\widehat{\ell}_{i-1})$,
3. $i|_{L''_0}$ is an isomorphism from L''_0 onto $T_{\widehat{\ell}_0} \Sigma_0$.

Differentiating (32) and (33) and recalling Property 1 we get the following proposition:

PROPOSITION 9. *For any* $\delta\ell \in T_{\widehat{\ell}_0} \Sigma_0$ *and* $k = 1, \ldots, r-1$, *the following equalities hold*

$$\langle \mathrm{d}\theta_k(\widehat{\ell}_0),\ \delta\ell \rangle = \omega_k(i^{-1}\delta\ell),$$
$$\frac{\mathrm{d}}{\mathrm{d}\ell} \mathcal{H}_{\tau_k(\ell)}(\ell) \Big|_{\widehat{\ell}_0} \delta\ell = \widehat{\mathcal{H}}_{\widehat{\tau}_k *} \delta\ell + \sum_{i=1}^{k-1} \omega_i(i^{-1}\delta\ell) \exp\left(\widehat{\theta}_k \overrightarrow{H}_k\right)_* \circ \ldots$$
$$\cdots \circ \exp\left(\widehat{\theta}_{i+1} \overrightarrow{H}_{i+1}\right)_* \overrightarrow{H}_i(\widehat{\ell}_i) + \omega_k(i^{-1}\delta\ell) \overrightarrow{H}_k(\widehat{\ell}_k),$$

and

$$T_{\ell_k} \mathcal{H}(\Sigma_k) = \Bigg\{ \widehat{\mathcal{H}}_{\widehat{\tau}_k *} \delta\ell + \sum_{i=1}^{k-1} \omega_i(i^{-1}\delta\ell) \exp\left(\widehat{\theta}_k \overrightarrow{H}_k\right)_* \circ \ldots$$
$$\cdots \circ \exp\left(\widehat{\theta}_{i+1} \overrightarrow{H}_{i+1}\right)_* \overrightarrow{H}_i(\widehat{\ell}_i) + \omega_k(i^{-1}\delta\ell) \overrightarrow{H}_k(\widehat{\ell}_k), \quad \delta\ell \in T_{\widehat{\ell}_0} \Sigma_0 \Bigg\}.$$

PROPOSITION 10. $\dim T_{\widehat{x}_k} \pi \mathcal{H}(\Sigma_k) = n - 1$ *for any* $k = 1, \ldots, r-1$.

Proof. 1) $\mathbf{k = 1}$: The elements of $T_{\widehat{x}_1} \pi \mathcal{H}(\Sigma_1)$ can be written as

$$\delta y_1 \equiv \widehat{S}_{\widehat{\tau}_1 *} \delta x + \omega_1(i^{-1}\delta\ell) h_1(\widehat{x}_1), \qquad \delta\ell \in T_{\widehat{\ell}_0} \Sigma_0, \quad \delta x = \pi_* \delta\ell$$

therefore $\delta y_1 = 0$ if and only if $\delta x + \omega_1(i^{-1}\delta\ell) g_1(\widehat{x}_0) = 0$. From Proposition 8 we know that $g_1(\widehat{x}_0)$ is not in $T_{\widehat{x}_0} \pi \Sigma_0$, while δx is, so we must have

$$\delta x = 0, \qquad \omega_1(i^{-1}\delta\ell) = 0.$$

Therefore $\dim T_{\widehat{x}_1}\pi\mathcal{H}(\Sigma_1) = \dim T_{\widehat{\ell}_0}\Sigma_0 = n-1$.
2) $\mathbf{k} = \mathbf{2}, \dots, \mathbf{r}-\mathbf{1}$: The elements of $T_{\widehat{x}_k}\pi\mathcal{H}(\Sigma_k)$ can be written as

$$\begin{aligned}\delta y_k \equiv\ & \widehat{S}_{\widehat{\tau}_k *}\delta x + \omega_k(i^{-1}\delta\ell)h_k(\widehat{x}_k) \\ & + \sum_{i=1}^{k-1}\omega_i(i^{-1}\delta\ell)\exp\left(\widehat{\theta}_k h_k\right)_* \dots \exp\left(\widehat{\theta}_{i+1}h_{i+1}\right)_* h_i(\widehat{x}_i)\,.\end{aligned}$$

So $\delta y_k = 0$ if and only if $\delta x + \sum_{i=1}^{k}\omega_i(i^{-1}\delta\ell)g_i(\widehat{x}_0) = 0$. i.e. if and only if $\pi_*\mathcal{G}''_k(i^{-1}\delta\ell) = 0$. But $J'' > 0$ implies $V_k \cap V_k^{\perp} = \{0\}$ hence $\delta x = 0$ and $\omega_i(i^{-1}(\delta\ell)) = 0$. □

PROPOSITION 11. *Let $k = 1, \dots, r-1$. Then $h_k(\widehat{x}_k)$, $h_{k+1}(\widehat{x}_k)$ are not in the tangent space $T_{\widehat{x}_k}\pi\mathcal{H}(\Sigma_k)$, moreover there exist $\delta y_k \in T_{\widehat{x}_k}\pi\mathcal{H}(\Sigma_k)$ and $C_k > 0$ such that $h_{k+1}(\widehat{x}_k) = \delta y_k + C_k h_k(\widehat{x}_k)$.*

Proof. $h_k(\widehat{x}_k) \in T_{\widehat{x}_k}\pi\mathcal{H}(\Sigma_k)$ if and only if there exists $\delta\ell \in T_{\widehat{\ell}_0}(\Sigma_0)$ such that

$$\begin{aligned}h_k(\widehat{x}_k) =\ & \omega_k(i^{-1}\delta\ell)h_k(\widehat{x}_k) + \widehat{S}_{\widehat{\tau}_k *}\pi_*\delta\ell \\ & + \sum_{i=1}^{k-1}\omega_i(i^{-1}\delta\ell)\exp(\widehat{\theta}_k h_k)_* \circ \dots \circ \exp(\widehat{\theta}_{i+1}h_{i+1})_* h_i(\widehat{x}_i)\,.\end{aligned} \tag{34}$$

If we let $\delta x \equiv \pi_* i^{-1}\delta\ell = \pi_*\delta\ell$, equation (34) is equivalent to

$$\delta x + \sum_{i=1}^{k-1}\omega_i(i^{-1}\delta\ell)g_i(\widehat{x}_0) + \Big[\omega_k(i^{-1}\delta\ell) - 1\Big]g_k(\widehat{x}_0) = 0$$

i.e.

$$\Big(\delta x, \omega_1(i^{-1}\delta\ell), \dots, \omega_{k-1}(i^{-1}\delta\ell), \omega_k(i^{-1}\delta\ell) - 1\Big) \in V_k \cap V_k^{\perp}\,.$$

Since $J'' > 0$, $V_k \cap V_k^{\perp} = \{0\}$ and we should have

$$\delta x = 0\,, \qquad \omega_i(i^{-1}\delta\ell) = 0 \quad i = 1, \dots, k-1\,, \qquad \omega_k(i^{-1}\delta\ell) = 1$$

which is of course impossible.

$h_{k+1}(\widehat{x}_k) \in T_{\widehat{x}_k}\pi\mathcal{H}(\Sigma_k)$ if and only if there exists $\delta\ell \in T_{\widehat{\ell}_0}\Sigma_0$ such that

$$\begin{aligned}h_{k+1}(\widehat{x}_k) =\ & \omega_k(i^{-1}\delta\ell)h_k(\widehat{x}_k) + \widehat{S}_{\widehat{\tau}_k *}\pi_*\delta\ell \\ & + \sum_{i=1}^{k-1}\omega_i(i^{-1}\delta\ell)\exp(\widehat{\theta}_k h_k)_* \circ \dots \circ \exp(\widehat{\theta}_{i+1}h_{i+1})_* h_i(\widehat{x}_i)\,.\end{aligned} \tag{35}$$

If we let $\delta x \equiv \pi_* i^{-1}\delta\ell = \pi_*\delta\ell$ equation (35) is equivalent to

$$\delta x + \sum_{i=1}^{k}\omega_i(i^{-1}\delta\ell)g_i(\widehat{x}_0) + (-1)g_{k+1}(\widehat{x}_0) = 0$$

i.e.

$$\left(\delta x, \omega_1(i^{-1}\delta\ell), \ldots, \omega_k(i^{-1}\delta\ell), -1\right) \in V_{k+1} \cap V_{k+1}^{\perp}.$$

Since $J'' > 0$, $V_{k+1} \cap V_{k+1}^{\perp} = \{0\}$, this is impossible.
Since $\dim T_{\widehat{x}_k} \pi \mathcal{H}(\Sigma_k) = n-1$ and $h_k(\widehat{x}_k) \notin T_{\widehat{x}_k} \pi \mathcal{H}(\Sigma_k)$, $T_{\widehat{x}_k} \pi \mathcal{H}(\Sigma_k)$ and $h_k(\widehat{x}_k)$ generate the whole tangent space $T_{\widehat{x}_k} M$ hence there exists $C_k \in \mathbb{R}$, $\delta x \in T_{x_0} L_0''$ such that

$$\begin{aligned} h_{k+1}(\widehat{x}_k) = C_k h_k(\widehat{x}_k) + \Big[& \widehat{S}_{\widehat{\tau}_k *} \delta x + \omega_k(i^{-1}\delta\ell) h_1(\widehat{x}_k) \\ & + \sum_{i=1}^{k-1} \omega_i(i^{-1}\delta\ell) \exp(\widehat{\theta}_k h_k)_* \circ \cdots \circ \exp(\widehat{\theta}_{i+1} h_{i+1})_* h_i(\widehat{x}_i) \Big] . \end{aligned} \tag{36}$$

Equation (36) is equivalent to

$$g_{k+1}(\widehat{x}_k) = C_k g_k(\widehat{x}_k) + \pi_* \mathcal{G}_k''(i^{-1}\delta\ell)$$

i.e. $\left(\delta x, \omega_1(i^{-1}\delta\ell), \ldots, \omega_{k-1}(i^{-1}\delta\ell), \omega_k(i^{-1}\delta\ell) + C_k, -1\right) \in V_{k+1} \cap V_k^{\perp}$.
Since $J''|_{V_{k+1} \cap V_k^{\perp}} > 0$, from equation (29) we get

$$-C_k G_{k+1}'' \left(\overrightarrow{G}_k''\right) > 0$$

i.e. $C_k > 0$. □

6.2. The map $\pi \circ \mathcal{H}$

THEOREM 2. *(1) $\pi \mathcal{H}(\Sigma_k)$ is an hyper-surface in M and the convex hull of $\{h_k(\widehat{x}_k),\ h_{k+1}(\widehat{x}_k)\}$ does not intersect its tangent hyper-plane $T_{\widehat{x}_k} \pi \mathcal{H}(\Sigma_k)$,*
(2) $\pi \mathcal{H}(\mathcal{O})$ is a neighborhood $\mathcal{U}$ of $\widehat{\Xi}$, $\pi \mathcal{H}\colon \mathcal{O} \to \mathcal{U}$ is one-to-one.

Proof. (1) comes directly from Propositions 10 and 11, possibly restricting $\mathcal{O}$. To prove *(2)* we first point out that Propositions 10 and 11 imply (see [1] for analogous techniques) that

$$\mathrm{D}(\pi \mathcal{H})(t, \widehat{\ell}_0) \text{ is onto} \quad \forall\, t \in [-\delta, \widehat{T} + \delta]. \tag{37}$$

This leads to $\pi \mathcal{H}\colon \Omega_i \to \pi \mathcal{H}(\Omega_i)$ to be a diffeomorphism, for a possibly smaller $\mathcal{O}$. Possibly restricting again $\mathcal{O}$ we may assume also

1. $\pi \mathcal{H}(\Omega_i) \cap \pi \mathcal{H}(\Omega_{i+1}) = \pi \mathcal{H}(\Sigma_i)$, by (37) at $t = \widehat{\tau}_i$ and Assumption 5;
2. $\pi \mathcal{H}(\Omega_i) \cap \pi \mathcal{H}\left(\Omega_j\right) = \emptyset$ for $|j - i| \geq 2$, by Assumption 5.

Therefore $\pi \mathcal{H}\colon \mathcal{O} \to \mathcal{U}$ is one-to-one. □

We shall denote $\pi \circ \mathcal{H}$ by ψ.

REMARK 6. Theorem 2 states that through ψ we can cover a neighborhood of the reference trajectory $\widehat{\xi}$ on M with a field of extremals.

6.3. Reduction to a finite-dimensional problem

We shall use the exactness of the 1-form $\boldsymbol{s} = p\,\mathrm{d}q$ on $\mathcal{H}(\Omega) \subset T^*M$.
Let $\xi\colon [0, T] \to M$ be a trajectory of the control system whose range is in the neighborhood $\mathcal{U}$ of $\widehat{\Xi}$. Call $x_0 = \xi(0)$ and $x_T = \xi(T)$. Apply ψ^{-1} to ξ and $\widehat{\xi}$ and close the path.

$$
\begin{aligned}
0 = \oint \mathcal{H}^* p\,\mathrm{d}q &= \int_{\psi^{-1}(\xi)} \mathcal{H}^* p\,\mathrm{d}q + \int_{\psi^{-1}(x_T)\psi^{-1}(\widehat{x}_{\widehat{T}})} \mathcal{H}^* p\,\mathrm{d}q \\
&\quad - \int_{\psi^{-1}(\widehat{\xi})} \mathcal{H}^* p\,\mathrm{d}q + \int_{\psi^{-1}(\widehat{x}_0)\psi^{-1}(x_0)} \mathcal{H}^* p\,\mathrm{d}q \,.
\end{aligned}
$$

The first integral is less than or equal to $\displaystyle\int_0^T p_0 f^0(\xi(s), u(s))\,\mathrm{d}s$, while the third one is $\displaystyle\int_0^{\widehat{T}} p_0 \widehat{f}^0(\widehat{\xi}(s))\,\mathrm{d}s$, hence

$$
\begin{aligned}
\int_{\psi^{-1}(\widehat{x}_0)\psi^{-1}(x_0)} \mathcal{H}^* p\,\mathrm{d}q \quad &\geq \quad \int_{\psi^{-1}(\widehat{x}_{\widehat{T}})\psi^{-1}(x_T)} \mathcal{H}^* p\,\mathrm{d}q \\
&+ \; p_0 \left[\int_0^{\widehat{T}} \widehat{f}^0(\widehat{\xi}(s))\,\mathrm{d}s \; - \; \int_0^T f^0(\xi(s, x_0), u(s))\,\mathrm{d}s \right] .
\end{aligned} \tag{38}
$$

By definition of the function ψ

$$
\psi^{-1}(\widehat{x}_0) = \left(0, \widehat{\ell}_0\right) , \qquad \psi^{-1}(\widehat{x}_{\widehat{T}}) = \left(\widehat{T}, \widehat{\ell}_0\right) .
$$

Let us define

$$
\begin{aligned}
&(t_0, \ell_0) \equiv \psi^{-1}(x_0) , \; q_0 \equiv \pi \ell_0 , \qquad (t_T, \ell_T) \equiv \psi^{-1}(x_T) , \; q_T \equiv \pi \ell_T , \\
&t_0 \in (-\delta, \widehat{\tau}_1) , \qquad t_T \in (\widehat{\tau}_{r-1}, \widehat{T} + \delta) .
\end{aligned}
$$

If x_0 is in a small neighborhood of $\widehat{x}_0$, then $x_0 = \exp t_0 h_1(q_0)$, so if we consider the difference $\alpha(x_0) - \alpha(\widehat{x}_0)$ of the costs on the initial point, we have:

$$
\begin{aligned}
\alpha(x_0) - \alpha(\widehat{x}_0) =& \alpha(x_0) - \alpha(q_0) + \alpha(q_0) - \alpha(\widehat{x}_0) \\
=& \left[\alpha\left(\exp(t_0 h_1) q_0\right) - \alpha(q_0) - p_0 \int_0^{t_0} f_1^0(\xi(r, q_0))\, \mathrm{d}r \right] \\
&+ \left[\alpha(q_0) - \alpha(\widehat{x}_0) + p_0 \int_0^{t_0} f_1^0(\xi(r, q_0))\, \mathrm{d}r \right] \\
=& t_0 \left[L_{h_1}\alpha(q_0) - p_0 f_1^0(q_0) \right] + \frac{t_0^2}{2} L_{h_1}\left(L_{h_1}\alpha - p_0 f_1^0 \right)(\bar{q}) \\
&+ \left[\alpha(q_0) - \alpha(\widehat{x}_0) + p_0 \int_0^{t_0} f_1^0(\xi(r, q_0))\, \mathrm{d}r \right]
\end{aligned}
\tag{39}
$$

for a certain $\bar{q}$ in a neighborhood of q_0.
The coefficient of t_0 is null because $q_0 \in \pi\Sigma_0$ while the coefficient of t_0^2 is strictly positive by continuity and the positivity of J'' on V_1. From equation (38) we now get

$$
\begin{aligned}
\alpha(x_0) - \alpha(\widehat{x}_0) \geq& \ \alpha(q_0) - \alpha(\widehat{x}_0) + p_0 \int_0^{t_0} f_1^0(\xi(r, q_0))\, \mathrm{d}r \\
=& \int_{\psi^{-1}(\widehat{x}_0)\psi^{-1}(x_0)} \mathcal{H}^* p\, \mathrm{d}q \\
\geq& \int_{\psi^{-1}(\widehat{x}_{\widehat{T}})\psi^{-1}(x_T)} \mathcal{H}^* p\, \mathrm{d}q + p_0 \left[\int_0^{\widehat{T}} \widehat{f}^0(\widehat{\xi}(s))\, \mathrm{d}s - \int_0^T f^0(\xi(s, x_0), u(s))\, \mathrm{d}s \right].
\end{aligned}
\tag{40}
$$

Let us now evaluate the difference of the costs $C(T, x_0, u)$ and $C(\widehat{T}, \widehat{x}_0, \widehat{u})$:

$$
\begin{aligned}
C(T, x_0, u) - C(\widehat{T}, \widehat{x}_0, \widehat{u}) =& \left(\alpha(x_0) + \beta(x_T) + p_0 \int_0^T f^0(\xi(s, x_0), u(s))\, \mathrm{d}s \right) \\
&- \left(\alpha(\widehat{x}_0) + \beta(\widehat{x}_{\widehat{T}}) + p_0 \int_0^{\widehat{T}} \widehat{f}^0(\widehat{\xi}(s))\, \mathrm{d}s \right) \\
\geq& \int_{\psi^{-1}(\widehat{x}_{\widehat{T}})\psi^{-1}(x_T)} \mathcal{H}^* p\, \mathrm{d}q + \beta(x_T) - \beta(\widehat{x}_{\widehat{T}}) \\
=& \left(\alpha \circ \pi \circ \psi^{-1} + \beta + p_0 \int_0^{t \circ \psi^{-1}(\cdot)} f^0_{\psi^{-1}(\cdot)}(\xi(\pi \circ \psi^{-1}(\cdot), r))\, \mathrm{d}r \right)(x_T) \\
&- \left(\alpha \circ \pi \circ \psi^{-1} + \beta + p_0 \int_0^{t \circ \psi^{-1}(\cdot)} f^0_{\psi^{-1}(\cdot)}(\xi(\pi \circ \psi^{-1}(\cdot), r))\, \mathrm{d}r \right)(\widehat{x}_{\widehat{T}})
\end{aligned}
$$

Let $F\colon N_f \to \mathbb{R}$ be defined as

$$
F\colon x \to \left(\alpha \circ \pi \circ \psi^{-1} + \beta + p_0 \int_0^{t \circ \psi^{-1}(\cdot)} f^0_{\psi^{-1}(\cdot)}(\xi(\pi \circ \psi^{-1}(\cdot), r))\, \mathrm{d}r \right)(x).
\tag{41}
$$

To prove that $\widehat{\xi}$ is a state-local minimum it now suffices to prove that F has a strict local minimum in $\widehat{x}_{\widehat{T}}$.

THEOREM 3. *F has a strict local minimum in $\widehat{x}_{\widehat{T}}$.*

Proof. It suffices to prove that

$$\mathrm{d}F(\widehat{x}_{\widehat{T}}) = 0, \tag{42}$$

$$\mathrm{D}^2 F(\widehat{x}_{\widehat{T}}) > 0\,. \tag{43}$$

Equation (42) is an immediate consequence of the transversality conditions (9) - (10). Since $\mathrm{d}\left(\alpha \circ \pi \circ \psi^{-1}\right) = \mathcal{H} \circ \pi \circ \psi^{-1}$

$$\begin{aligned}\mathrm{D}^2 F(\widehat{x}_{\widehat{T}})[\delta x_T]^2 &= \Big((\mathcal{H} \circ \psi^{-1})_* + \mathrm{D}^2(\beta + p_0 B_0)\Big)\,(\widehat{x}_{\widehat{T}})[\delta x_T]^2 \\ &= \boldsymbol{\sigma}\,\Big((\mathcal{H} \circ \psi^{-1})_* \delta x_T\,, \mathrm{d}(-\beta - p_0 B_0)_* \delta x_T\Big)\end{aligned}$$

where B_0 is the integral appearing in formula (41). Since ψ_*^{-1} is a diffeomorpism of $T_{x_{\widehat{T}}} M$ and $\mathbb{R} \times T_{\widehat{\ell}_0} \Sigma_0$, let $(\delta s, \delta\ell) \equiv \psi_*^{-1} \delta x_T$. We get

$$\begin{aligned}\mathrm{D}^2 F(\widehat{x}_{\widehat{T}})[\delta x_T]^2 &= \boldsymbol{\sigma}\,(\mathcal{H}_*(\delta s, \delta\ell), \mathrm{d}(-\beta - p_0 B_0)_* \psi_*(\delta s, \delta\ell)) \\ &= \boldsymbol{\sigma}\,(\mathcal{H}_*(\delta s, \delta\ell), \mathrm{d}(-\beta - p_0 B_0)_* \pi_* \mathcal{H}_*(\delta s, \delta\ell))\,.\end{aligned} \tag{44}$$

Apply $\widehat{\mathcal{H}}_{\widehat{T}*} \circ i$ to each argument of $\boldsymbol{\sigma}$ in inequality (31) of proposition 7. With $\delta\ell \equiv i(\delta p, \delta x)$ we obtain

$$\begin{aligned}0 &< \boldsymbol{\sigma}\Bigg(\widehat{\mathcal{H}}_{\widehat{T}*} i\Bigg((\delta p, \delta x) + \sum_{i=1}^{r-1} \omega_i(\delta p, \delta x) \overrightarrow{G}''_i + \delta t\, \overrightarrow{G}''_r\Bigg), \\ &\qquad \widehat{\mathcal{H}}_{\widehat{T}*} i\Bigg(\delta p_{\widehat{T}}, \delta x + \sum_{i=1}^{r-1} \omega_i(\delta p, \delta x) g_i(\widehat{x}_0) + \delta t\, g_r(\widehat{x}_0)\Bigg)\Bigg) \\ &= \boldsymbol{\sigma}\,(\mathcal{H}_*(\delta t + \langle \mathrm{d}\tau_{r-1}\,,\ \delta\ell\rangle), \mathrm{d}\alpha_* \delta x), \\ &\qquad - \delta p_{\widehat{T}} + \mathrm{d}(-\beta - p_0 B_0)_* \pi_* \mathcal{H}_*(\delta t + \langle \mathrm{d}\tau_{r-1}\,,\ \delta\ell\rangle))\end{aligned}$$

for any $(\delta p, \delta x) \in L''_0$ and any $\delta t \in \mathbb{R}$ such that $\pi_* \left(\mathcal{G}''_{r-1}(\delta p, \delta x) + \delta t\, \overrightarrow{G}''_r\right) \neq 0$ and is in $\pi_* L''_{\widehat{T}}$ and for any $(\delta p_{\widehat{T}}, \delta x_{\widehat{T}}) \in L''_{\widehat{T}}$ such that
$\delta x_{\widehat{T}} = \pi_* \left(\mathcal{G}''_{r-1}(\delta p, \delta x) + \delta t\, \overrightarrow{G}''_r\right)$. This is exactly the expression in equation (44) with $\delta s = \delta t + \langle \mathrm{d}\tau_{r-1}\,,\ \delta\ell\rangle$. □

Proof of Theorem 1. To conclude the proof of Theorem 1 we have to prove that $\widehat{\xi}$ is a strict minimizer and that in the abnormal case $p_0 = 0$ it is isolated.
Assume that (T, ξ, u) is another triplet which achieves the minimum or that $p_0 = 0$. Since $\widehat{x}_{\widehat{T}}$ is a strict minimum for F, we must have $\xi(T) = \widehat{x}_{\widehat{T}}$ and equality must hold

in equations (38) and (40). Let us denote $\psi^{-1}(\xi(s)) = (t(s), \ell(s))$, $s \in [0, T]$ and let $\mu = (\lambda, \xi) \colon s \to \mathcal{H}_{t(s)}(\ell(s))$. Equality in equation (38) implies

$$\langle \lambda(s)\,,\ \dot{\xi}(s) \rangle - p_0 f^0(\xi(s), u(s)) = H(\lambda(s), \xi(s)) \quad \text{a.e.} \tag{45}$$

and by Assumption 2 we obtain $\dot{\xi}(s) = h_i(\xi(s))$ if $\xi(s)$ in in the interior of $\psi(\Omega_i)$. Since $\xi(T) = \widehat{x}_{\widehat{T}}$ we have $\xi(s) = \widehat{\xi}(s - T + \widehat{T})$. By (45) and Assumption 3 the limit points of $\dot{\xi}(s)$ for $s \to \widehat{\tau}_{r-1} - T + \widehat{T}$ must belong to the convex hull of $\{h_{r-1}(\xi(\widehat{\tau}_{r-1} - T + \widehat{T})), h_r(\xi(\widehat{\tau}_{r-1} - T + \widehat{T}))\}$ therefore Theorem 2 implies that ξ crosses $\psi(\Sigma_{r-1})$ and it enters the interior of $\psi(\Omega_{r-1})$. By induction we get $\xi(s) = \widehat{\xi}(s - T + \widehat{T})$, whenever $s \geq \max\{0, T - \widehat{T}\}$. This also means that $\xi(0) = \exp(T - \widehat{T})h_1(\widehat{x}_0)$ so that in equation (39) we have $q_0 = \widehat{x}_0$ and $t_0 = T - \widehat{T}$, so since equality must hold in (40) we must have $t_0 = 0$ in (39). This completes the proof. □

References

[1] AGRACHEV A.A., STEFANI G. AND ZEZZA P. *Strong optimality for a bang-bang trajectory*, SIAM J. Control Optimization **41** (4) (2002), 991–1014.

[2] POGGIOLINI L. AND STEFANI G. *State-local optimality of a bang-bang trajectory: an Hamiltonian approach*, submitted.

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

[4] STEFANI G. AND ZEZZA P., *Time-optimality of a bang-bang trajectory with maple*, in: "2nd IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control", 2003.

AMS Subject Classification: 49K15, 49K30.

Laura POGGIOLINI, Dipartimento di Matematica Applicata "G. Sansone", Università di Firenze, via Santa Marta 3, I-50139 Firenze, ITALY
e-mail: `laura.poggiolini@unifi.it`