STRONG LOCAL OPTIMALITY FOR A BANG-BANG TRAJECTORY IN A MAYER 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 [5, 13] and [16]. We require both the *strict bang-bang Legendre condition*, and the second order conditions for the finite dimensional problem obtained by moving the switching times of the reference trajectory.

Key words. bang-bang controls, second order sufficient conditions, Hamiltonian methods

AMS subject classifications. 49K15, 49J15, 93C10

1. Introduction.

We consider a Mayer problem where the control functions are bounded and enter linearly in the dynamics.

$$\text{minimize} \quad C(\xi, u) := c_0(\xi(0)) + c_f(\xi(T)) \tag{1.1a}$$

$$\text{subject to} \quad \dot{\xi}(t) = f_0(\xi(t)) + \sum_{s=1}^{m} u_s f_s(\xi(t)) \tag{1.1b}$$

$$\xi(0) \in N_0\,, \quad \xi(T) \in N_f \tag{1.1c}$$

$$u = (u_1, \ldots, u_m) \in L^\infty([0, T], [-1, 1]^m). \tag{1.1d}$$

Here $T > 0$ is given, the state space is a n-dimensional manifold M, N_0 and N_f are smooth sub-manifolds of M. The vector fields $f_0, f_1, \ldots, f_m$ and the functions c_0, c_f are C^2 on M, N_0 and N_f, respectively.

We aim at giving second order sufficient conditions for a *reference bang-bang extremal 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 admissible trajectories, regardless of any distance of the associated controls. Therefore, optimality is with respect to neighboring trajectories, independently of the values of the associated controls. In particular, if the extremal is abnormal, we prove that $\widehat{\xi}$ is isolated among admissible trajectories.

We 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 component $\widehat{u}_i$ of the reference control $\widehat{u}$ 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 component changes value at $\hat{t}_k$, while it is called *multiple* if at least two control components change value.

Second order conditions for the optimality of a bang-bang extremal with simple switches only are given in [5, 10, 13, 16] and references therein, while in [18] sufficient conditions are given in the case of the minimum time problem for L^1-local optimality - an intermediate condition between strong and local optimality - of a bang-bang

*The first author was supported by PRIN 2006019927_002 - Controllabilità stabilizzabilità e ottimizzazione di sistemi dinamici non lineari, the second author was supported by PRIN 200789J59E_002 - Comportamento qualitativo delle soluzioni delle Equazioni Differenziali Ordinarie (con eventuale ritardo)

†Dipartimento di Matematica Applicata, Università degli Studi di Firenze, via di Santa Marta, 3, I-50139 Firenze (laura.poggiolini@math.unifi.it and marco.spadini@math.unifi.it)

extremal having both simple and multiple switches with the extra assumption that the Lie brackets of the switching vector fields are annihilated by the adjoint covector.

All the above cited papers require regularity assumptions on the switches (see the subsequent Assumptions 1, 2 and 3 which are the natural strengthening of necessary conditions) and the positivity of a suitable second variation.

Here we consider the problem of strong local optimality, when at most one double switch occurs, but there are finitely many simple ones and no commutativity assumptions on the involved vector fields. More precisely we extend the conditions in [5, 13, 16] by requiring the sufficient second order conditions for the finite dimensional sub-problems that are obtained by allowing the switching times to move. The addition of a double switch is not a trivial extension of the known single-switch cases. In fact, as explained in Section 2.2, any perturbation of the double switching time of $\hat{u}$ creates generically two simple switches, that is a new *bang* arc is generated. On the contrary, small perturbations of a single switch do not change the structure of the reference control, i.e. while in the case of simple switches the only variables are the switching times, each time a double switch occurs one has to consider the two possible combinations of the switching controls. This fact gives rise to a non-smooth flow, whose invertibility is investigated via some topological methods described in the Appendix, or via Clarke's Inverse Function Theorem (see [6, Thm 7.1.1.]) in some particular degenerate case.

We believe that the techniques employed here could be extended to the more general case when there are more than one double switch. However, such an extension may not be straightforward as the technical and notational complexities grow quickly with the number of double switches.

Preliminary results were given in [17], where the authors analyze a study case and in [14] that deals with a Bolza problem in the so-called *non-degenerate case*. Also, stability analysis under parameter perturbations for this kind of bang-bang extremals was studied in [7]. In this paper we focus on the geometric construction, so that in order to keep the analytic machinery at a minimum we rely on the technical paper [15] which contains all the computations.

The paper is organized as follows: Section 2.1 introduces the notation and the regularity hypotheses that are assumed through the paper. Although we are going to use mostly the Hamiltonian formulation, here the regularity assumptions are stated also in terms of the switching functions which are more widely known. In Section 2.2, where our main result Theorem 2.1 is stated, we introduce a finite dimensional sub-problem of (1.1) and its "second variations" (indeed this sub-problem is $C^{1,1}$ but not C^2 so that the classical "second variation" is not well defined). The essence of the paper will be to show that the sufficient conditions for the optimality of $(\widehat{\xi}, \widehat{u})$ for this sub-problem are actually sufficient also for the optimality of the reference pair $(\hat{\xi}, \hat{u})$ in problem (1.1). Here we also briefly describe the Hamiltonian methods the proof is based upon. Section 3 contains the maximized Hamiltonian of the control system and its flow. In Section 4, we write the "second variations" of the finite-dimensional sub-problem and study their sign on appropriate spaces. Section 5 is the heart of the paper; there we prove that the projection onto a neighborhood of the graph of $\hat{\xi}$ in $\mathbb{R} \times M$ of the maximized flow defined in Section 3 is invertible (which is necessary for our Hamiltonian methods to work). Section 6 contains the conclusion of the proof of Theorem 2.1. In the Appendix we treat from an abstract viewpoint the problem, raised in Section 5, of local invertibility of a piecewise C^1 function.

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. The regularity Assumptions 2 and 3 are natural, since we look for sufficient conditions. In fact Pontryagin Maximum Principle yields the necessity of the same inequalities but in weak form.

2.1. Notation and regularity. We are given an admissible reference couple $(\widehat{\xi}, \widehat{u})$ satisfying Pontryagin Maximum Principle (PMP) with adjoint covector $\widehat{\lambda}$ and the reference control $\widehat{u}$ is bang-bang with switching times $\widehat{t}_1, \ldots, \widehat{t}_r$ such that only two kinds of switchings appear:

(i) $\widehat{t}_i$ is a *simple switching time* i.e. only one of the control components $\widehat{u}_1, \ldots, \widehat{u}_m$ switches at time $\widehat{t}_i$;

(ii) $\widehat{t}_i$ is a *double switching time* i.e. exactly two of the control components $\widehat{u}_1, \ldots, \widehat{u}_m$ switch at time $\widehat{t}_i$.

Assume that there is just one double switching time, which we denote by $\hat{\tau}$. Without loss of generality we may assume that the control components switching at time $\hat{\tau}$ are $\widehat{u}_1$ and $\widehat{u}_2$ and that they both switch from the value -1 to $+1$, i.e.

$$\lim_{t \to \hat{\tau}-} \widehat{u}_\nu = -1 \qquad \lim_{t \to \hat{\tau}+} \widehat{u}_\nu = 1 \qquad \nu = 1, 2.$$

In the interval $(0, \hat{\tau})$, J_0 simple switches occur, and J_1 simple switches occur in the interval $(\hat{\tau}, T)$. We denote the simple switching times occurring before the double one by $\hat{\theta}_{0j}$, $j = 1, \ldots, J_0$, and by $\hat{\theta}_{1j}$, $j = 1, \ldots, J_1$ the simple switching times occurring afterward. In order to simplify the notation, we also define $\hat{\theta}_{00} := 0$, $\hat{\theta}_{0,J_0+1} := \hat{\theta}_{10} := \hat{\tau}$, $\hat{\theta}_{1,J_1+1} := T$, i.e. we have

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

We use some basic tools and notation from differential geometry. For any sub-manifold N of M, and any $x \in N$, T_xN and T_x^*N denote the tangent space to N at x and the cotangent space to N at x, respectively, while T^*N denotes the cotangent bundle. For any $w \in T_x^*M$ and any $\delta x \in T_xM$, $\langle w \,,\ \delta x \rangle$ denotes the duality product between a form and a tangent vector; $\pi\colon T^*M \to M$ denotes the canonical projection from the tangent bundle onto the base manifold M. In coordinates $\ell := (p, x)$:

$$\pi\colon \ell = (p, x) \in T^*M \mapsto x \in M.$$

Throughout the paper, for any vector field $f\colon x \in M \mapsto f(x) \in T_xM$, we denote the associated Hamiltonian obtained by lifting f to T^*M by the corresponding capital letter, i.e.

$$F\colon \ell \in T^*M \mapsto \langle \ell \,,\ f(\pi\ell) \rangle \in \mathbb{R},$$

and $\overrightarrow{F}$ denotes the Hamiltonian vector field associated to F. In particular for any $s = 0, 1, \ldots, m$, $F_s(\ell) := \langle \ell \,,\ f_s(\pi\ell) \rangle$ is the Hamiltonian associated to the drift ($s = 0$) and to the controlled vector fields ($s = 1, \ldots, m$) of system (1.1b).

If $f, g\colon M \in TM$, are differentiable vector fields, we denote their Lie bracket as $[f, g]$:

$$[f, g](x) := \mathrm{D}g(x)\, f(x) - \mathrm{D}f(x)\, g(x)$$

The canonical symplectic two-form between $\overrightarrow{F}$ and $\overrightarrow{G}$ at a point ℓ is denoted as $\boldsymbol{\sigma}\left(\overrightarrow{F},\overrightarrow{G}\right)(\ell)$. In coordinates $\ell := (p,x)$:

$$\boldsymbol{\sigma}\left(\overrightarrow{F},\overrightarrow{G}\right)(\ell) := -\langle p\,\mathrm{D}g(x)\,,\ f(x)\rangle + \langle p\,\mathrm{D}f(x)\,,\ g(x)\rangle.$$

For any m-tuple $u = (u_1,\ldots,u_m) \in \mathbb{R}^m$ let us denote the control-dependent Hamiltonian by

$$h_u\colon \ell \in T^*M \mapsto \langle \ell\,,\ f_0(\pi\ell) + \sum_{s=1}^{m} u_s f_s(\pi\ell)\rangle \in \mathbb{R}.$$

Let $\widehat{f}_t$ and $\widehat{F}_t$ be the reference vector field and the reference Hamiltonian, respectively:

$$\widehat{f}_t(x) := f_0(x) + \sum_{s=1}^{m} \widehat{u}_s(t) f_s(x)\,, \quad \widehat{F}_t(\ell) := \langle \ell\,,\ \widehat{f}_t(\pi\ell)\rangle = h_{\widehat{u}(t)}(\ell).$$

Also, let $\widehat{x}_0 := \widehat{\xi}(0)$, $\widehat{x}_d := \widehat{\xi}(\hat{\tau})$ and $\widehat{x}_f := \widehat{\xi}(T)$; the reference flow, that is the flow associated to $\widehat{f}_t$, is defined on the whole interval $[0,T]$ at least in a neighborhood of $\widehat{x}_0$. We denote it as $\widehat{S}\colon (t,x) \mapsto \widehat{S}_t(x)$. Finally, let H be the maximized Hamiltonian associated to the control system:

$$H(\ell) := \max\left\{h_u(\ell)\colon u \in [-1,1]^m\right\}$$

Thus, in our situation PMP reads as follows:
There exist $p_0 \in \{0,1\}$ and $\widehat{\lambda}\colon [0,T] \to T^*M$, absolutely continuous, such that

$$(p_0, \widehat{\lambda}(0)) \neq (0,0) \tag{2.1}$$

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

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

$$\widehat{\lambda}(0)|_{T_{\widehat{x}_0}N_0} = p_0\,\mathrm{d}c_0(\widehat{x}_0), \qquad \widehat{\lambda}(T)|_{T_{\widehat{x}_f}N_f} = -p_0\,\mathrm{d}c_f(\widehat{x}_f) \tag{2.2}$$

$$\widehat{F}_t(\widehat{\lambda}(t)) = H(\widehat{\lambda}(t)) \qquad \text{a.e. } t \in [0,T]. \tag{2.3}$$

We denote $\widehat{\ell}_0 := \widehat{\lambda}(0)$ and $\widehat{\ell}_f := \widehat{\lambda}(T)$.

In terms of the switching functions

$$\sigma_s\colon t \in [0,T] \mapsto F_s \circ \widehat{\lambda}(t) = \langle \lambda(t)\,,\ f_s(\widehat{\xi}(t)\rangle \in \mathbb{R},\ s = 1,\ldots,m,$$

maximality condition (2.3) means $\widehat{u}_s(t)\sigma_s(t) \geq 0$ for any $t \in [0,T]$ and any $s = 1,\ldots,m$. We assume the following regularity condition holds:

Assumption 1 (Regularity). *Let $s \in \{1,\ldots,m\}$. If t is not a switching time for the control component $\widehat{u}_s$, then*

$$\widehat{u}_s(t)\sigma_s(t) = \widehat{u}_s(t)F_s(\widehat{\lambda}(t)) = \widehat{u}_s(t)\langle\widehat{\lambda}(t)\,,\ f_s(\widehat{\xi}(t))\rangle > 0. \tag{2.4}$$

For $j = 0,\ldots,J_i$, $i = 0,1,$, let $k_{ij} := \widehat{f}_t|_{(\widehat{\theta}_{ij},\widehat{\theta}_{i,j+1})}$, be the restrictions of $\widehat{f}_t$ to each of the time intervals where the reference control $\widehat{u}$ is constant and let $K_{ij}(\ell) := \langle \ell\,,\ k_{ij}(\pi\ell)\rangle$ be the associated Hamiltonian. From maximality condition (2.3)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(K_{ij} - K_{i,j-1}\right) \circ \widehat{\lambda}(t)\Big|_{t=\hat{\theta}_{ij}} \geq 0$$

for any $i = 0, 1, \quad j = 1, \dots, J_i$, i.e. if $\hat{u}_{s(ij)}$ is the control component switching at time $\hat{\theta}_{ij}$ and $\Delta_{ij} \in \{-2, 2\}$ is its jump, then

$$\frac{\mathrm{d}}{\mathrm{d}t} \Delta_{ij} \sigma_{s(ij)}(t) \Big|_{t=\hat{\theta}_{ij}} \geq 0$$

We assume that the strong inequality holds at each simple switching time $\hat{\theta}_{ij}$:

Assumption 2.

$$\frac{\mathrm{d}}{\mathrm{d}t} \Delta_{ij} \sigma_{s(ij)}(t) \Big|_{t=\hat{\theta}_{ij}} = \frac{\mathrm{d}}{\mathrm{d}t} \left(K_{ij} - K_{i,j-1} \right) \circ \widehat{\lambda}(t) \Big|_{t=\hat{\theta}_{ij}} > 0 \quad i = 0, 1, \ j = 1, \dots, J_i. \tag{2.5}$$

Assumption 2 is known as the Strong bang-bang Legendre condition for simple switching times.

In geometric terms Assumption 2 means that at time $t = \hat{\theta}_{ij}$ the trajectory $t \mapsto \widehat{\lambda}(t)$ crosses transversely the hyper-surface of T^*M defined by $K_{ij} = K_{i,j-1}$, i.e. by the zero level set of $F_{s(ij)}$, arriving with transverse velocity $\overrightarrow{K}_{i,j-1}(\widehat{\lambda}(\hat{\theta}_{ij}))$ and leaving with transverse velocity $\overrightarrow{K}_{ij}(\widehat{\lambda}(\hat{\theta}_{ij}))$.

As already said we can assume that the double switching time involves the first two components, $\hat{u}_1$ and $\hat{u}_2$ of the reference control $\hat{u}$ which both switch from -1 to $+1$, so that

$$k_{10} = k_{0J_0} + 2f_1 + 2f_2.$$

Define the new vector fields

$$k_\nu := k_{0J_0} + 2f_\nu \, , \quad \nu = 1, 2,$$

with associated Hamiltonians $K_\nu(\ell) := \langle \ell \, , \ k_\nu(\pi\ell) \rangle$. Then, from maximality condition (2.3) we get

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} 2\,\sigma_\nu(t)\Big|_{t=\hat{\tau}-} = \frac{\mathrm{d}}{\mathrm{d}t} 2\,F_\nu \circ \widehat{\lambda}(t)\Big|_{t=\hat{\tau}-} = \frac{\mathrm{d}}{\mathrm{d}t}\left(K_\nu - K_{0J_0}\right) \circ \widehat{\lambda}(t)\Big|_{t=\hat{\tau}-} \geq 0, \\
&\frac{\mathrm{d}}{\mathrm{d}t} 2\,\sigma_\nu(t)\Big|_{t=\hat{\tau}+} = \frac{\mathrm{d}}{\mathrm{d}t} 2\,F_\nu \circ \widehat{\lambda}(t)\Big|_{t=\hat{\tau}+} = \frac{\mathrm{d}}{\mathrm{d}t}\left(K_{10} - K_\nu\right) \circ \widehat{\lambda}(t)\Big|_{t=\hat{\tau}+} \geq 0,
\end{aligned} \qquad \nu = 1, 2.$$

We assume that the strict inequalities hold:

Assumption 3.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(K_\nu - K_{0J_0}\right) \circ \widehat{\lambda}(t)\Big|_{t=\hat{\tau}-} > 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(K_{10} - K_\nu\right) \circ \widehat{\lambda}(t)\Big|_{t=\hat{\tau}+} > 0, \quad \nu = 1, 2. \tag{2.6}$$

Assumption 3 means that at time $\hat{\tau}$ the flow arrives at the hyper-surfaces $F_1 = 0$ and $F_2 = 0$ with transverse velocity $\overrightarrow{K}_{0J_0}(\widehat{\lambda}(\hat{\tau}))$ and leaves with velocity $\overrightarrow{K}_{10}(\widehat{\lambda}(\hat{\tau}))$ which is again transverse to both the hyper-surfaces. We shall call Assumption 3 the Strong bang-bang Legendre condition for double switching times.

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 one. Moving a simple switching time $\hat{\theta}_{ij}$ to time

$\theta_{ij} := \hat{\theta}_{ij} + \delta_{ij}$ amounts to using the values $\widehat{u}|_{(\hat{\theta}_{i,j-1},\hat{\theta}_{ij})}$ and $\widehat{u}|_{(\hat{\theta}_{ij},\hat{\theta}_{i,j+1})}$ of the reference control in the time intervals $\big(\hat{\theta}_{i,j-1}, \theta_{ij}\big)$ and $\big(\theta_{i\,j}, \hat{\theta}_{i,j+1}\big)$, respectively. On the other hand, when we move the double switching time $\hat{\tau}$ we change the switching time of two different components of the reference control and we must allow for each of them to change its switching time independently of the other. This means that between the values of $\widehat{u}|_{(\hat{\theta}_{0J_0},\hat{\tau})}$ and $\widehat{u}|_{(\hat{\tau},\hat{\theta}_{01})}$ we introduce a value of the control which is not assumed by the reference one - at least in a neighborhood of $\hat{\tau}$ - and which may assume two different values according to which component switches first. Let $\tau_\nu := \hat{\tau} + \varepsilon_\nu$, $\nu = 1, 2$. We move the switching time of the first control component $\widehat{u}_1$ from $\hat{\tau}$ to $\tau_1 := \hat{\tau} + \varepsilon_1$, and the switching time of $\widehat{u}_2$ from $\hat{\tau}$ to $\tau_2 := \hat{\tau} + \varepsilon_2$.

Inspired by [5], let us introduce C^2 functions $\alpha \colon M \to \mathbb{R}$ and $\beta \colon M \to \mathbb{R}$ such that $\alpha|_{N_0} = p_0 c_0$, $\mathrm{d}\alpha(\widehat{x}_0) = \widehat{\ell}_0$, $\beta|_{N_f} = p_0 c_f$ and $\mathrm{d}\beta(\widehat{x}_f) = -\widehat{\ell}_f$.

Define $\theta_{ij} := \hat{\theta}_{ij} + \delta_{ij}$, $j = 1, \ldots, J_i$, $i = 0, 1$; $\theta_{0,J_0+1} := \min\{\tau_\nu,\ \nu = 1, 2\}$, $\theta_{10} := \max\{\tau_\nu,\ \nu = 1, 2\}$, $\theta_{00} := 0$ and $\theta_{1,J_1+1} := T$. We have a finite-dimensional sub-problem (FP) given by

$$\text{minimize} \quad \alpha(\xi(0)) + \beta(\xi(T)) \tag{FPa}$$

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

$$\text{and} \quad \xi(0) \in N_0, \quad \xi(T) \in N_f. \tag{FPc}$$

$$\text{where} \quad \theta_{00} = 0, \qquad \theta_{1,J_1+1} = T \tag{FPd}$$

$$\theta_{ij} = \hat{\theta}_{ij} + \delta_{ij}, \quad i = 0, 1, \quad j = 1, \ldots, J_i, \tag{FPe}$$

$$\theta_{0,J_0+1} := \hat{\tau} + \min\{\varepsilon_1,\ \varepsilon_2\}, \quad \theta_{10} := \hat{\tau} + \max\{\varepsilon_1,\ \varepsilon_2\} \tag{FPf}$$

$$\text{and} \quad \begin{cases} \nu = 1 & \text{if } \varepsilon_1 \le \varepsilon_2, \\ \nu = 2 & \text{if } \varepsilon_2 \le \varepsilon_1. \end{cases} \tag{FPg}$$

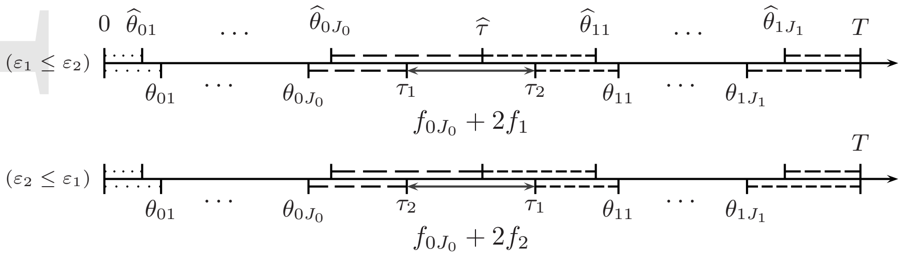


Fig. 2.1. *The different sequences of vector fields in the finite-dimensional sub-problem.*

We denote the solution, evaluated at time t, of (FPb) emanating from a point $x \in M$ at time 0, as $S_t(x, \delta, \varepsilon)$. Observe that $S_t(x, 0, 0) = \widehat{S}_t(x)$, and that the reference control is achieved along $\varepsilon_1 = \varepsilon_2$, i.e. the reference flow is attained by (FP) on a point of non-differentiability of the functions

$$\theta_{0,J_0+1} := \hat{\tau} + \min\{\varepsilon_1,\ \varepsilon_2\}, \qquad \theta_{10} := \hat{\tau} + \max\{\varepsilon_1,\ \varepsilon_2\}.$$

We are going to prove (see Remark 1 in Section 4) that despite this lack of differentiability, (FP) is C^1 (indeed $C^{1,1}$) at $\delta_{ij} = \varepsilon_1 = \varepsilon_2 = 0$.

We can thus consider, on the kernel of the first variation of (FP), its second variation, piece-wisely defined as the second variation of the restrictions of (FP) to the half-spaces $\{(\delta, \varepsilon) \colon \varepsilon_1 \leq \varepsilon_2\}$ and $\{(\delta, \varepsilon) \colon \varepsilon_2 \leq \varepsilon_1\}$. Because of the structure of (FP), this second variation is coercive if and only if both restrictions are positive–definite quadratic forms, see Remark 2 in Section 4. In particular any of their convex combinations is positive-definite on the kernel of the first variation, i.e. Clarke's generalized Hessian at $(x, \delta, \varepsilon) = (\widehat{x}_0, 0, 0)$ is positive-definite on that kernel.

In Section 4 we give explicit formulas both for the first and for the second variations. We ask for such second variations to be positive definite and prove the following theorem:

THEOREM 2.1. *Let $(\widehat{\xi}, \widehat{u})$ be a bang–bang regular extremal (in the sense of Assumption 1) for problem* (1.1) *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 strong Legendre conditions, Assumptions 2 and 3, hold. Assume also that the second variation of problem (FP) is positive definite on the kernel of the first variation. Then $(\widehat{\xi}, \widehat{u})$ is a strict strong local optimizer for problem (1.1)*. If the extremal is abnormal ($p_0 = 0$), then $\widehat{\xi}$ is an isolated admissible trajectory.*

The proof will be carried out by means of Hamiltonian methods. For a general introduction to such methods see e.g. [1, 2, 3], below we illustrate such methods for our particular problem.

- In Section 3 we prove that the maximized Hamiltonian of the control system H is well defined and Lipschitz continuous on the whole cotangent bundle T^*M. Its Hamiltonian vector field $\overrightarrow{H}$ is piecewise smooth in a neighborhood of the range of $\widehat{\lambda}$ and its classical flow, denoted by

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

is well defined in a neighborhood of $[0, T] \times \{\widehat{\ell}_0\}$. We also show that $\widehat{\lambda}$ is a trajectory of $\overrightarrow{H}$, i.e. $\hat{\lambda}(t) = \mathcal{H}_t(\widehat{\ell}_0)$.

- In Sections 4-5 we prove that there exist a C^2 function α such that $\alpha|_{N_0} = p_0 c_0$, $\mathrm{d}\alpha(x_0) = \widehat{\ell}_0$ and enjoying the following property: the map

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

is one–to–one onto a neighborhood of the graph of $\widehat{\xi}$, where $\Lambda := \{\mathrm{d}\alpha(x) \colon x \in \mathcal{O}(x_0)\}$. Indeed the proof of this invertibility is the main core of the paper and its main novelty.

- Under the above conditions the one–form $\omega := \mathcal{H}^*(p\,\mathrm{d}q - H\,\mathrm{d}t)$ is exact on $[0, T] \times \Lambda$, hence there exists a C^1 function

$$\chi \colon (t, \ell) \in [0, T] \times \Lambda \mapsto \chi_t(\ell) \in \mathbb{R}$$

such that $\mathrm{d}\chi = \omega$. Also it may be shown (see, e.g. [5]) that $\mathrm{d}(\chi_t \circ (\pi\mathcal{H}_t)^{-1}) = \mathcal{H}_t \circ (\pi\mathcal{H}_t)^{-1}$ for any $t \in [0, T]$. Moreover we may assume $\chi_0 = \alpha \circ \pi$

Observe that $(t, \widehat{\xi}(t)) = (\mathrm{id} \times \pi\mathcal{H})(t, \widehat{\ell}_0)$ and let us show how this construction leads to the reduction. Define

$$\mathcal{V} := (\mathrm{id} \times \pi\mathcal{H})([0, T] \times \Lambda), \qquad \psi := (\mathrm{id} \times \pi\mathcal{H})^{-1} \colon \mathcal{V} \to [0, T] \times \Lambda$$

and let (ξ, u) be an admissible pair (i.e. a pair satisfying (1.1b)–(1.1c)–(1.1d)) such that the graph of ξ is in $\mathcal{V}$. We can obtain a closed path Γ in $\mathcal{V}$ with a concatenation of the following paths:

(i) $\Xi\colon t \in [0,T] \mapsto (t, \xi(t)) \in \mathcal{V}$,

(ii) $\Phi_T\colon s \in [0,1] \mapsto (T, \phi_T(s)) \in \mathcal{V}$, where $\phi_T\colon s \in [0,1] \mapsto \phi_T(s) \in M$ is such that $\phi_T(0) = \xi(T)$, $\phi_T(1) = \widehat{x}_f$,

(iii) $\widehat{\Xi}\colon t \in [0,T] \mapsto (t, \widehat{\xi}(t)) \in \mathcal{V}$, ran backward in time,

(iv) $\Phi_0\colon s \in [0,1] \mapsto (0, \phi_0(s)) \in \mathcal{V}$, where $\phi_0\colon s \in [0,1] \mapsto \phi_0(s) \in M$ is such that $\phi_0(0) = \widehat{x}_0$, $\phi_0(1) = \xi(0)$.

Since the one-form ω is exact we get

$$0 = \oint_\Gamma \omega = \int_{\psi(\Xi)} \omega + \int_{\psi(\Phi_T)} \omega - \int_{\psi(\widehat{\Xi})} \omega + \int_{\psi(\Phi_0)} \omega.$$

From the definition of ω and the maximality properties of H we get

$$\int_{\psi(\widehat{\Xi})} \omega = 0, \qquad \int_{\psi(\Xi)} \omega \leq 0 \tag{2.8}$$

so that

$$\int_{\psi(\Phi_T)} \omega + \int_{\psi(\Phi_0)} \omega \geq 0. \tag{2.9}$$

Since

$$\int_{\psi(\Phi_T)} \omega = \int_{(\pi\mathcal{H}_T)^{-1}\circ\Phi_T} \mathrm{d}(\chi_T \circ (\pi\mathcal{H}_T)^{-1}) = \chi_T \circ (\pi\mathcal{H}_T)^{-1}(\widehat{x}_f) - \chi_T \circ (\pi\mathcal{H}_T)^{-1}(\xi(T)),$$

$$\int_{\psi(\Phi_0)} \omega = \int_0^1 \langle \mathrm{d}\alpha(\phi_0(s))\,,\ \dot{\phi}_0(s)\rangle \,\mathrm{d}s = \alpha(\xi(0)) - \alpha(\widehat{x}_0),$$

inequality (2.9) yields

$$\alpha(\xi(0)) - \alpha(\widehat{x}_0) + \chi_T \circ (\pi\mathcal{H}_T)^{-1}(\widehat{x}_f) - \chi_T \circ (\pi\mathcal{H}_T)^{-1}(\xi(T)) \geq 0. \tag{2.10}$$

Thus

$$\begin{aligned}\alpha(\xi(0)) + \beta(\xi(T)) - \alpha(\widehat{x}_0) - \beta(\widehat{x}_f)& \\ \geq \left(\chi_T \circ (\pi\mathcal{H}_T)^{-1} + \beta\right)(\xi(T)) - \left(\chi_T \circ (\pi\mathcal{H}_T)^{-1} + \beta\right)(\widehat{x}_f)&\end{aligned} \tag{2.11}$$

that is: we only have to prove the local minimality at $\widehat{x}_f$ of the function

$$\mathcal{F}\colon x \in N_f \cap \mathcal{O}(\widehat{x}_f) \mapsto \left(\chi_T \circ (\pi\mathcal{H}_T)^{-1} + \beta\right)(x) \in \mathbb{R}.$$

where $\mathcal{O}(\widehat{x}_f)$ is a small enough neighborhood of $\widehat{x}_f$.
In proving both the invertibility of $\mathrm{id} \times \pi\mathcal{H}$ and the local minimality of $\widehat{x}_f$ for $\mathcal{F}$ we will analyze the positivity of the second variations of problem (FP).

3. The maximized flow. We are now going to prove the properties of the maximized Hamiltonian H and of the flow – given by classical solutions – of the associated Hamiltonian vector field $\overrightarrow{H}$. Such flow will turn out to be Lipschitz continuous and piecewise–C^1. In such construction we use only the regularity Assumptions 1–2–3 and not the positivity of the second variations of problems (FP). We proceed as follows: in *Step 1* we consider the simple switches occurring before the double one. We explain the procedure in details for the first simple switch. The others are treated iterating such procedure as in [5]; in *Step 2* we decouple the double switch obtaining two simple switches and that give rise to as many flows. Finally in *Step 3* we consider the simple switches that occur after the double one. For each of the flows originating from the double switch we apply the same procedure of Step 1.

Step 1: Regularity Assumption 1 implies that locally around $\widehat{\ell}_0$, the maximized Hamiltonian is K_{00} and that $\widehat{\lambda}(t)$, i.e. the flow of $\overrightarrow{K}_{00}$ evaluated in $\widehat{\ell}_0$, intersects the set $\{\ell \in T^*M\colon K_{01}(\ell) = K_{00}(\ell)\}$ at time $\hat{\theta}_{01}$. Assumption 2 yields that such intersection is transverse. This suggests to define $\theta_{01}(\ell)$ as the time when the flow of $\overrightarrow{K}_{00}$, emanating from ℓ, intersects such set and to switch to the flow of $\overrightarrow{K}_{01}$ afterwards. To be more precise, we apply the implicit function theorem to the map

$$\Phi_{01}(t,\ell) := (K_{01} - K_{00}) \circ \exp t\overrightarrow{K}_{00}(\ell)$$

in a neighborhood of $(t,\ell) := (\hat{\theta}_{01}, \widehat{\ell}_0)$ in $[0,T] \times T^*M$, so that $H(\ell) = K_{00}(\ell)$ for any $t \in [0, \theta_{01}(\ell)]$. We then iterate this procedure and obtain the switching surfaces $\{(\theta_{0j}(\ell), \ell)\colon \ell \in \mathcal{O}(\widehat{\ell}_0)\}$, $j = 0, \ldots, J_0$ where:

1. $\theta_{00}(\ell) := 0$, $\phi_{00}(\ell) := \ell$;
2. $\theta_{0j}(\ell)$ is the unique solution to

$$(K_{0j} - K_{0,j-1}) \circ \exp \theta_{0j}(\ell)\overrightarrow{K}_{0,j-1}\left(\phi_{0,j-1}(\ell)\right) = 0$$

defined by the implicit function theorem in a neighborhood of $(t,\ell) = (\hat{\theta}_{0j}, \widehat{\ell}_0)$ and $\phi_{0j}(\ell)$ is defined by

$$\phi_{0j}(\ell) := \exp\big(-\theta_{0j}(\ell)\overrightarrow{K}_{0j}\big) \circ \exp \theta_{0j}(\ell)\overrightarrow{K}_{0,j-1}\left(\phi_{0,j-1}(\ell)\right).$$

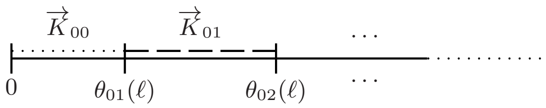


Fig. 3.1. *Construction of the maximized flow.*

Step 2: Let us now show how to decouple the double switching time in order to define the maximized Hamiltonian $H(\ell)$ in a neighborhood of $(\hat{\tau}, \widehat{\lambda}(\hat{\tau}))$. In this we depart from [5] in that we introduce the vector fields k_1, k_2 in the sequence of values assumed by the reference vector field. We do this in four stages:

1. for $\nu = 1,\ 2$ let $\tau_\nu(\ell)$ be the unique solution to

$$(K_\nu - K_{0J_0}) \circ \exp \tau_\nu(\ell)\overrightarrow{K}_{0J_0}(\phi_{0J_0}(\ell)) = 0$$

defined by the implicit function theorem in a neighborhood of $(\hat{\tau}, \widehat{\ell}_0)$;

2. choose

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

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

$$\phi^\nu_{0,J_0+1}(\ell) := \exp\big(-\tau_\nu(\ell)\overrightarrow{K}_\nu\big) \circ \exp \tau_\nu(\ell)\overrightarrow{K}_{0J_0}\left(\phi_{0J_0}(\ell)\right); \tag{3.1}$$

3. for $\nu = 1, 2$ let $\theta^\nu_{10}(\ell)$ be the unique solution to

$$(K_{10} - K_\nu) \circ \exp \theta_{10}(\ell)\overrightarrow{K}_\nu\left(\phi^\nu_{0,J_0+1}(\ell)\right) = 0 \tag{3.2}$$

defined by the implicit function theorem in a neighborhood of $(\hat{\tau}, \widehat{\ell}_0)$ and define

$$\phi^\nu_{10}(\ell) := \exp\big(-\theta^\nu_{10}(\ell)\overrightarrow{K}_{10}\big) \circ \exp \theta^\nu_{10}(\ell)\overrightarrow{K}_\nu\left(\phi^\nu_{0,J_0+1}(\ell)\right);$$

4. choose

$$\theta_{10}(\ell) = \begin{cases} \theta^1_{10}(\ell) & \text{if } \tau_1(\ell) \le \tau_2(\ell), \\ \theta^2_{10}(\ell) & \text{if } \tau_2(\ell) < \tau_1(\ell). \end{cases}$$

Notice that if $\tau_1(\ell) = \tau_2(\ell)$, then $\theta^1_{10}(\ell) = \theta^2_{10}(\ell) = \tau_1(\ell) = \tau_2(\ell)$ so that $\theta_{10}(\cdot)$ is continuous. To be more precise, the function $\theta_{10}(\cdot)$ is Lipschitz continuous on its domain and is actually C^1 except possibly on the set $\{\ell \in T^*M \colon \tau_1(\ell) = \tau_2(\ell)\}$.

Step 3: Finally we define analogous quantities for the simple switching times that follow the double one. For each $j = 1, \ldots, J_1$ we proceed in two stages:

1. for $\nu = 1, 2$ let $\theta^\nu_{1j}(\ell)$ be the unique solution to

$$(K_{1j} - K_{1,j-1}) \circ \exp \theta^\nu_{1j}(\ell) \overrightarrow{K}_{1,j-1} \left(\phi^\nu_{1,j-1}(\ell) \right) = 0$$

defined by the implicit function theorem in a neighborhood of $(\hat{\theta}^\nu_{1j}, \widehat{\ell}_0)$ and define

$$\phi^\nu_{1j}(\ell) := \exp \big(- \theta^\nu_{1j}(\ell) \overrightarrow{K}_{1j} \big) \circ \exp \theta^\nu_{1j}(\ell) \overrightarrow{K}_{1,j-1} \left(\phi^\nu_{i,j-1}(\ell) \right);$$

2. choose

$$\theta_{1j}(\ell) = \begin{cases} \theta^1_{1j}(\ell) & \text{if } \tau_1(\ell) \le \tau_2(\ell) \\ \theta^2_{1j}(\ell) & \text{if } \tau_2(\ell) < \tau_1(\ell) \end{cases}$$

We conclude the procedure by setting $\theta_{1,J_1+1}(\ell) = \theta^1_{1,J_1+1}(\ell) = \theta^2_{1,J_1+1}(\ell) := T$. Thus we get that the flow of the maximized Hamiltonian coincides with the flow of the Hamiltonian $H \colon (t, \ell) \in [0, T] \times T^*M \mapsto H_t(\ell) \in T^*M$

$$H_t(\ell) := \begin{cases} K_{0j}(\ell) & t \in (\theta_{0j}(\ell), \theta_{0,j+1}(\ell)], \quad j = 0, \ldots, J_0 \\ K_\nu(\ell) & t \in (\theta_{0,J_0+1}(\ell), \theta_{10}(\ell)], \quad \text{if } \theta_{0,J_0+1}(\ell) = \tau_\nu(\ell) \\ K_{1j}(\ell) & t \in (\theta_{1j}(\ell), \theta_{1,j+1}(\ell)], \quad j = 0, \ldots, J_1. \end{cases} \tag{3.3}$$

4. The second variation. To choose an appropriate horizontal Lagrangian manifold Λ we write the second variations of sub–problem (FP) and exploit their positivity. To write an invariant second variation, as introduced in [4], we write the pull–back $\zeta_t(x, \delta, \varepsilon)$ of the flows S_t along the reference flow $\widehat{S}_t$. Define the pullbacks of the vector fields k_{ij} and h_ν

$$g_{ij}(x) := \widehat{S}^{-1}_{\hat{\theta}_{ij}\,*} k_{ij} \circ \widehat{S}_{\hat{\theta}_{ij}}(x), \quad h_\nu(x) := \widehat{S}^{-1}_{\hat{\tau}\,*} k_\nu \circ \widehat{S}_{\hat{\tau}}(x)$$

and let $\delta_{0,J_0+1} := \min\{\varepsilon_1, \varepsilon_2\}$, $\delta_{10} := \max\{\varepsilon_1, \varepsilon_2\}$. At time $t = T$ we have

$$\begin{aligned} \zeta_T(x, \delta, \varepsilon) = \widehat{S}^{-1}_T \circ S_T(x, \delta, \varepsilon) = \exp\left(-\delta_{1J_1}\right) g_{1J_1} \circ \ldots \circ \exp\left(\delta_{11} - \delta_{10}\right) g_{10} \circ \\ \circ \exp\left(\delta_{10} - \delta_{0,J_0+1}\right) h_\nu \circ \exp\left(\delta_{0,J_0+1} - \delta_{0J_0}\right) g_{0J_0} \circ \ldots \circ \exp \delta_{01} g_{00}(x) \end{aligned}$$

where $\nu = 1$ if $\varepsilon_1 \le \varepsilon_2$, $\nu = 2$ otherwise. In order to analyze the influence of the double switch on the flow we need to introduce some further notation: let $\widetilde{f}_1$ and $\widetilde{f}_2$ be the pull–backs of f_1 and f_2 from time $\hat{\tau}$ to time $t = 0$, i.e., $\widetilde{f}_\nu := \widehat{S}^{-1}_{\hat{\tau}\,*} f_\nu \circ \widehat{S}_{\hat{\tau}}$, $\nu = 1, 2$, so that $h_\nu = g_{0J_0} + 2\widetilde{f}_\nu$, $\nu = 1, 2$, and $g_{10} = g_{0J_0} + 2\widetilde{f}_1 + 2\widetilde{f}_2$.

First, to better understand the situation, assume that no simple switch occurs ($J_0 = J_1 = 0$). In this case the linearized flow at time T has the following form:

$$L(\delta x, \delta, \varepsilon) = \delta x + (\delta_{11} - \delta_{01}) g_{01}(x) + 2(\delta_{11} - \varepsilon_1)\widetilde{f}_1(x) + 2(\delta_{11} - \varepsilon_2)\widetilde{f}_2(x),$$

which shows that the flow is C^1.

We now proceed to the more general situation. Define

$$a_{00} := \delta_{01}, \qquad a_{0j} := \delta_{0,j+1} - \delta_{0j} \quad j = 1, \dots, J_0, \qquad b := \delta_{10} - \delta_{0,J_0+1};$$
$$a_{1j} := \delta_{1,j+1} - \delta_{1j} \quad j = 0, \dots, J_1 - 1, \qquad a_{1J_1} := -\delta_{1J_1}.$$

Then $b + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} = 0$ and, with a slight abuse of notation, we may write

$$\begin{aligned}\zeta_T(x,a,b) = & \exp a_{1J_1} g_{1J_1} \circ \ldots \circ \exp a_{11} g_{11} \circ \exp a_{10} g_{10} \\ & \circ \exp b h_\nu \circ \exp a_{0J_0} g_{0J_0} \circ \ldots \circ \exp a_{01} g_{01} \circ \exp a_{00} g_{00}(x),\end{aligned}$$

where $\nu = 1$ if $\varepsilon_1 \le \varepsilon_2$, $\nu = 2$ otherwise. Henceforward we denote by a the (J_0+J_1+2)-tuple $(a_{00}, \dots, a_{0J_0}, a_{10}, \dots, a_{1J_1})$.

The reference flow is the one associated to $(a,b) = (0,0)$ and the first order approximation of ζ_T at a point $(x,0,0)$ is given by

$$\begin{aligned}L(\delta x, a, b) = & \delta x + b h_\nu(x) + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} g_{ij}(x) = \delta x + \sum_{j=0}^{J_0-1} a_{0j} g_{0j}(x) + \\ & + (\delta_{0,J_0+1} - \delta_{0J_0})\, g_{0J_0}(x) + (\delta_{10} - \delta_{0,J_0+1}) h_\nu(x) + (\delta_{11} - \delta_{10}) g_{10}(x) + \sum_{j=1}^{J_1} a_{1j} g_{1j}(x)\end{aligned}$$

where $\nu = 1$ if $\varepsilon_1 \le \varepsilon_2$, $\nu = 2$ otherwise. Thus

$$\begin{aligned}L(\delta x, a, b) = & \delta x + \sum_{j=0}^{J_0-1} a_{0j} g_{0j}(x) + (\delta_{11} - \delta_{0J_0}) g_{0J_0}(x) + 2(\delta_{11} - \varepsilon_1) \widetilde{f}_1(x) + \\ & + 2(\delta_{11} - \varepsilon_2) \widetilde{f}_2(x) + \sum_{j=1}^{J_1} a_{1j} g_{1j}(x).\end{aligned} \tag{4.1}$$

Remark 1. *Equation* (4.1) *shows that in* $L(\delta x, a, b)$ *we have the same first order expansion, whatever the sign of* $\varepsilon_2 - \varepsilon_1$*. This proves that the finite–dimensional problem (FP) is* C^1*.*

Let $\widehat{\beta} := \beta \circ \widehat{S}_T$ and $\widehat{\gamma} := \alpha + \widehat{\beta}$. Then the cost (FPa) can be written as

$$J(x,a,b) = \alpha(x) + \widehat{\beta} \circ \zeta_T(x,a,b)$$

and, by PMP, $\mathrm{d}\widehat{\gamma}(\widehat{x}_0) = 0$. The first variation of J at $(x,a,b) = (\hat{x}_0, 0, 0)$ is given by

$$J'(\delta x, a, b) = \Big(b h_\nu + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} g_{ij} \Big) \cdot \widehat{\beta}(\widehat{x}_0)$$

which, by (4.1), does not depend on ν, i.e. it does not depend on the sign of $\varepsilon_2 - \varepsilon_1$. On the other hand, the second order expansion of $\zeta_T^\nu(x,\cdot,\cdot)$ at $(a,b) = (0,0)$ is

$$\exp\Bigg(b h_\nu + \sum_{i=0}^{1} \sum_{j=0}^{J_i} a_{ij} g_{ij} + \frac{1}{2} \Bigg\{ \sum_{j=0}^{J_0} a_{0j} \Big[g_{0j}, \sum_{s=j+1}^{J_0} a_{0s} g_{0s} + b h_\nu + \sum_{j=0}^{J_1} a_{1j} g_{1j} \Big] +$$

$$+ b\Big[h_\nu, \sum_{j=0}^{J_1} a_{1j} g_{1j}\Big] + \sum_{j=0}^{J_1} a_{1j}\Big[g_{1j}, \sum_{s=j+1}^{J_1} a_{1s} g_{1s}\Big]\Big\}\Big)(x).$$

where $\nu = 1$ if $\varepsilon_1 \leq \varepsilon_2$, $\nu = 2$ otherwise. Using this expansion and proceeding as in [5] we get, for all $(\delta x, a, b) \in \ker J'$,

$$J''_\nu[(\delta x, a, b)]^2 = \frac{1}{2}\Big\{ \mathrm{d}^2\widehat{\gamma}(\widehat{x}_0)[\delta x]^2 + 2\,\delta x \cdot \Big(\sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij}\, g_{ij} + b h_\nu \Big) \cdot \widehat{\beta}(\widehat{x}_0) +$$
$$+ \Big(\sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij}\, g_{ij} + b h_\nu \Big)^2 \cdot \widehat{\beta}(\widehat{x}_0) + \sum_{j=0}^{J_0}\sum_{i=0}^{j-1} a_{0i} a_{0j} [g_{0i}, g_{0j}] \cdot \widehat{\beta}(\widehat{x}_0) +$$
$$+ b \sum_{i=0}^{J_0} a_{0i}[g_{0i}, h_\nu] \cdot \widehat{\beta}(\widehat{x}_0) + \sum_{j=0}^{J_1} a_{1j} \Big(\sum_{i=0}^{J_0} a_{0i}[g_{0i}, g_{1j}] + b[h_\nu, g_{1j}] +$$
$$+ \sum_{i=0}^{j-1} a_{1i}[g_{1i}, g_{1j}] \Big) \cdot \widehat{\beta}(\widehat{x}_0) \Big\}$$

where, again, $\nu = 1$ if $\varepsilon_1 \leq \varepsilon_2$, $\nu = 2$ otherwise.

Remark 2. *The previous formula clearly shows that $J''_1 = J''_2$ on $\{(\delta x, a, b) \colon b = 0\}$, i.e. on $\{(\delta x, \delta, \varepsilon) \colon \varepsilon_1 = \varepsilon_2\}$. The second variation is J''_1 if $\varepsilon_1 \leq \varepsilon_2$, J''_2 otherwise. Its coercivity means that both J''_1 and J''_2 are coercive quadratic forms.*

Remark 3. *Isolating the addenda where a_{0J_0}, b, a_{10} appear, as in (4.1), one can easily see that $J''_1 = J''_2$ if and only if $[\widetilde{f}_1, \widetilde{f}_2] \cdot \widehat{\beta}(\widehat{x}_0) = 0$, i.e. if and only if $\langle \widehat{\lambda}(\hat{\tau})\,,\ [f_1, f_2](\hat{x}_d) \rangle = 0$. In other words: problem (FP) is twice differentiable at $(x, \delta, \varepsilon) = (\hat{x}_0, 0, 0)$ if and only if $\langle \widehat{\lambda}(\hat{\tau})\,,\ [f_1, f_2](\hat{x}_d) \rangle = 0$.*

By assumption, for each $\nu = 1, 2$, J''_ν is positive definite on

$$\mathcal{N}_0 := \Big\{ (\delta x, a, b) \in T_{\widehat{x}_0} N_0 \times \mathbb{R}^{J_0 + J_1 + 2} \times \mathbb{R} \colon \\ b + \sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij} = 0, \quad L(\delta x, a, b) \in T_{\widehat{x}_f} N_f \Big\}.$$

Again following the procedure of [5] we may modify α by adding a suitable second–order penalty at $\widehat{x}_0$ (see e.g. [8], Theorem 13.2) so that we may assume that each second variation J''_ν is positive definite on

$$\mathcal{N} := \Big\{ (\delta x, a, b) \in T_{\widehat{x}_0} M \times \mathbb{R}^{J_0 + J_1 + 2} \times \mathbb{R} \colon \\ b + \sum_{i=0}^{1}\sum_{j=0}^{J_i} a_{ij} = 0, \quad L(\delta x, a, b) \in T_{\widehat{x}_f} N_f \Big\},$$

i.e. we can remove the constraint on the initial point of admissible trajectories.

We choose Λ as the Lagrangian manifold defined by such α, that is

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

and we study the positivity of J''_ν as follows: consider

$$V := \Big\{ (\delta x, a, b) \in \mathcal{N} \colon L(\delta x, a, b) = 0 \Big\}$$

and the sequence $V_{01} \subset \ldots \subset V_{0J_0} \subset V_{10} \subset \ldots \subset V_{1J_1} = V$ of sub–spaces of V, defined as

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

Then J''_ν is positive definite on $\mathcal{N}$ if and only if it is positive definite on each $V_{ij} \cap V_{i,j-1}^{\perp_{J''_\nu}}$, $V_{10} \cap V_{0J_0}^{\perp_{J''_\nu}}$ and $\mathcal{N} \cap V^{\perp_{J''_\nu}}$, and notice that

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

for any $j = 2, \ldots, J_0$, $k = 0, \ldots, J_1$ and $\nu = 1, 2$.

As in [5] one can prove a characterization, in terms of the maximized flow, of the intersections above. We state here such characterization without proofs which can be found in [5]. Recall that the G_{ij}'s and the H_ν's denote the Hamiltonian obtained by lifting the vector fields g_{ij}'s h_ν's.

LEMMA 4.1. *Let $j = 1, \ldots, J_0$ and $\delta e = (\delta x, a, b) \in V_{0j}$. Assume J''_ν is coercive on $V_{0,j-1}$. Then $\delta e \in V_{0j} \cap V_{0,j-1}^{\perp_{J''_\nu}}$ if and only if*

$$a_{0s} = \langle \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_* \delta x\rangle \quad \forall s = 0, \ldots, j-2. \tag{4.2}$$

In this case

$$J''_\nu[\delta e]^2 = -a_{0j}\,\boldsymbol{\sigma}\Big(\mathrm{d}\alpha_*\delta x + \sum_{s=0}^{j-1} a_{0s}\overrightarrow{G}_{0s}(\widehat{\ell}_0), (\overrightarrow{G}_{0j} - \overrightarrow{G}_{0,j-1})(\widehat{\ell}_0)\Big). \tag{4.3}$$

LEMMA 4.2. *Let $\nu = 1, 2$ and $\delta e = (\delta x, a, b) \in V_{10}$. Assume J''_ν is coercive on V_{0J_0}. Then $\delta e \in V_{10} \cap V_{0J_0}^{\perp_{J''_\nu}}$ if and only if*

$$a_{0s} = \langle \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_* \delta x\rangle \quad \forall s = 0, \ldots, J_0 - 1. \tag{4.4}$$

In this case

$$\begin{aligned} J''_\nu[\delta e]^2 = &- b\,\boldsymbol{\sigma}\Big(\mathrm{d}\alpha_*\delta x + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}_{0s}(\widehat{\ell}_0), (\overrightarrow{H}_\nu - \overrightarrow{G}_{0,J_0})(\widehat{\ell}_0)\Big) - \\ &- a_{10}\,\boldsymbol{\sigma}\Big(\mathrm{d}\alpha_*\delta x + \sum_{s=0}^{J_0} a_{0s}\overrightarrow{G}_{0s}(\widehat{\ell}_0) + b\overrightarrow{H}_\nu(\widehat{\ell}_0), (\overrightarrow{G}_{10} - \overrightarrow{H}_\nu)(\widehat{\ell}_0)\Big). \end{aligned} \tag{4.5}$$

LEMMA 4.3. *Let $\nu = 1, 2$, $j = 1, \ldots, J_1$ and $\delta e = (\delta x, a, b) \in V_{1j}$. Assume J''_ν is coercive on $V_{1,j-1}$. Then $\delta e \in V_{1j} \cap V_{1,j-1}^{\perp_{J''_\nu}}$ if and only if*

$$\begin{aligned} a_{0s} &= \langle \mathrm{d}(\theta_{0,s+1} - \theta_{0s})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_*\delta x\rangle \quad \forall s = 0, \ldots, J_0 \\ b &= \langle \mathrm{d}(\theta_{10} - \theta_{0,J_0+1})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_*\delta x\rangle \\ a_{1s} &= \langle \mathrm{d}(\theta_{1,s+1} - \theta_{1s})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_*\delta x\rangle \quad \forall s = 0, \ldots, j-2. \end{aligned}$$

In this case $J''_\nu[\delta e]^2$ *is given by*

$$-a_{1j}\,\boldsymbol{\sigma}\left(\mathrm{d}\alpha_*\delta x+\sum_{s=0}^{J_0}a_{0s}\overrightarrow{G}_{0s}(\widehat{\ell}_0)+b\overrightarrow{H}_\nu(\widehat{\ell}_0)+\sum_{i=0}^{j-1}a_{1i}\overrightarrow{G}_{1i}(\widehat{\ell}_0),(\overrightarrow{G}_{1j}-\overrightarrow{G}_{1,j-1})(\widehat{\ell}_0)\right).$$

Lemma 4.4. *Let* $\nu=1,2$ *and* $\delta e=(\delta x,a,b)\in\mathcal{N}$. *Assume* J''_ν *is coercive on* V_{1J_1}. *Then* $\delta e\in\mathcal{N}\cap V_{1J_1}^{\perp_{J''_\nu}}$ *if and only if* $\delta e\in\mathcal{N}$ *and*

$$\begin{aligned}
a_{0s}&=\langle \mathrm{d}(\theta_{0,s+1}-\theta_{0s})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_*\delta x\rangle\quad\forall s=0,\dots,J_0\\
b&=\langle \mathrm{d}(\theta_{10}-\theta_{0,J_0+1})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_*\delta x\rangle\\
a_{1s}&=\langle \mathrm{d}(\theta_{1,s+1}-\theta_{1s})\,(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_*\delta x\rangle\quad\forall s=0,\dots,J_1-1.
\end{aligned}$$

In this case

$$\begin{aligned}
J''_\nu[\delta e]^2=-\boldsymbol{\sigma}\Big(\,\mathrm{d}(-\widehat{\beta})_*\big(\delta x+\sum_{i=0}^{1}\sum_{s=0}^{J_i}a_{is}g_{is}(\widehat{x}_0)+bh_\nu(\widehat{x}_0)\big)\,,&\\
\mathrm{d}\alpha_*\delta x+\sum_{i=0}^{1}\sum_{s=0}^{J_i}a_{is}\overrightarrow{G}_{is}(\widehat{\ell}_0)+b\overrightarrow{H}_\nu(\widehat{\ell}_0)\Big).&
\end{aligned}$$

5. The invertibility of the flow. We now prove that the map

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

is one-to-one onto a neighborhood of the graph of $\widehat{\xi}$. Since the time interval $[0,T]$ is compact and by the properties of flows, it suffices to show that $\pi\mathcal{H}_{\hat\tau}$ and $\pi\mathcal{H}_{\hat\theta_{ij}}$, for $i=1,2$ and $j=1,\dots,J_i$, are one-to-one onto a neighborhood of $\widehat{\xi}(\hat\tau)$ and $\widehat{\xi}(\hat\theta_{ij})$ in M, respectively. The proof of the invertibility at the simple switching times $\hat\theta_{0j}$, $j=1,\dots,J_0$ may be carried out either as in [5] or by means of Clarke's Inverse Function Theorem (see [15], Lemma 6.1). Here we skip this proof but give some details on the invertibility at the double switching time and at the simple switching times $\hat\theta_{1j}$, $j=1,\dots,J_1$ that follow it. This proof will be performed, depending on the dimension of the kernel of $\mathrm{d}(\tau_1-\tau_2)|_{T_{\widehat{\ell}_0}\Lambda}$, by means of Clarke's Inverse Function Theorem or using topological methods (see Theorem 7.6). The invertibility at the simple switching times $\widehat{\theta}_{0j}$ yields the invertibility of

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

about $[0,\hat\tau-\varepsilon]\times\{\widehat{\ell}_0\}$.

We now show that such procedure can be carried out also on $[\hat\tau-\varepsilon,T]\times\{\widehat{\ell}_0\}$, so that $\mathrm{id}\times\pi\mathcal{H}$ will turn out to be locally invertible from a neighborhood $[0,T]\times\mathcal{O}\subset[0,T]\times\Lambda$ of $[0,T]\times\{\widehat{\ell}_0\}$ 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 $\pi\mathcal{H}_{\hat\tau}$ at $\widehat{\ell}_0$.

In a neighborhood of $\widehat{\ell}_0$, $\pi\mathcal{H}_{\hat\tau}$ has the following piecewise representation:

1. if $\min\{\tau_1(\ell),\ \tau_2(\ell)\}\geq\hat\tau$, then $\quad\pi\mathcal{H}_{\hat\tau}(\ell)=\exp\hat\tau\overrightarrow{K}_{0J_0}\circ\phi_{0J_0}(\ell)$,
2. if $\min\{\tau_1(\ell),\tau_2(\ell)\}=\tau_1(\ell)\leq\hat\tau\leq\theta_{10}(\ell)$, then

$$\pi\mathcal{H}_{\hat\tau}(\ell)=\exp(\hat\tau-\tau_1(\ell))\overrightarrow{K}_1\circ\exp\tau_1(\ell)\overrightarrow{K}_{0J_0}\circ\phi_{0J_0}(\ell),$$

3. if $\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_2(\ell) \le \hat\tau \le \theta_{10}(\ell)$, then

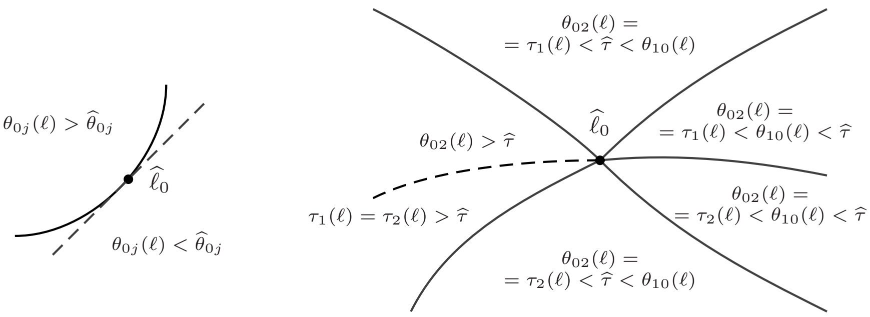
$$\pi\mathcal{H}_{\hat\tau}(\ell) = \exp(\hat\tau - \tau_2(\ell))\overrightarrow{K}_2 \circ \exp\tau_2(\ell)\overrightarrow{K}_{0J_0} \circ \phi_{0J_0}(\ell),$$

4. if $\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_1(\ell) \le \theta_{10}(\ell) \le \hat\tau$, then

$$\pi\mathcal{H}_{\hat\tau}(\ell) = \exp(\hat\tau - \theta_{10}(\ell))\overrightarrow{K}_{10} \circ \exp(\theta_{10}(\ell) - \tau_1(\ell))\overrightarrow{K}_1 \circ \exp\tau_1(\ell)\overrightarrow{K}_{0J_0} \circ \phi_{0J_0}(\ell),$$

5. if $\min\{\tau_1(\ell), \tau_2(\ell)\} = \tau_2(\ell) \le \theta_{10}(\ell) \le \hat\tau$, then

$$\pi\mathcal{H}_{\hat\tau}(\ell) = \exp(\hat\tau - \theta_{10}(\ell))\overrightarrow{K}_{10} \circ \exp(\theta_{10}(\ell) - \tau_2(\ell))\overrightarrow{K}_2 \circ \exp\tau_2(\ell)\overrightarrow{K}_{0J_0} \circ \phi_{0J_0}(\ell).$$



(a) $t = \widehat\theta_{0j}$ (b) $t = \widehat\tau$

FIG. 5.1. *Local behaviour of $\mathcal{H}_t$ near $\widehat\ell_0$ at a simple switching time and at the double one.*

Let us denote by L^0, L^{11}, L^{21}, L^{12}, L^{22} the linearization of the five expressions for $\pi\mathcal{H}_{\hat\tau}$, and let M^0, M^{11}, M^{21}, M^{12}, M^{22} be the polyhedral cones where they respectively hold.

LEMMA 5.1. *The piecewise linearized maps L^0, L^{11}, L^{21}, L^{12}, L^{22} have the same orientation in the following sense: given any basis of $T_{\widehat\ell_0}\Lambda_0$ and any basis of $T_{\widehat\xi(\hat\tau)}M$, the determinants of the matrices associated to the linear maps L^0, $L^{\nu j}$, $\nu, j = 1, 2$, in such bases, have the same sign.*

Proof. The assertion follows from Lemma 7.1 in the Appendix if one shows that the following claims hold:

Claim 1. If $\langle \mathrm{d}\tau_\nu(\widehat\ell_0)\,,\ \delta\ell_2\rangle < 0 < \langle \mathrm{d}\tau_\nu(\widehat\ell_0)\,,\ \delta\ell_1\rangle$ then $L^0_{\widehat\tau}(\delta\ell_1) \ne L^{\nu 1}_{\widehat\tau}(\delta\ell_2)$, i.e.

$$\exp(\hat\tau k_{0J_0})_*\pi_*\phi_{0J_0\,*}(\delta\ell_1) \ne \exp(\hat\tau k_{0J_0})_*\pi_*\phi_{0J_0\,*}(\delta\ell_2) - \langle \mathrm{d}\tau_\nu(\widehat\ell_0)\,,\ \delta\ell_2\rangle(k_\nu - k_{0J_0})(\widehat x_{\hat\tau}).$$

Claim 2. If $\langle \mathrm{d}\theta^\nu_{01}(\widehat\ell_0)\,,\ \delta\ell_2\rangle < 0 < \langle \mathrm{d}\theta^\nu_{01}(\widehat\ell_0)\,,\ \delta\ell_1\rangle$ then $L^{\nu 1}_{\widehat\tau}(\delta\ell_1) \ne L^{\nu 2}_{\widehat\tau}(\delta\ell_2)$, i.e.

$$\begin{aligned}
&\exp(\hat\tau k_{0J_0})_*\pi_*\phi_{0J_0\,*}(\delta\ell_1) - \langle \mathrm{d}\tau_\nu(\widehat\ell_0)\,,\ \delta\ell_1\rangle(k_\nu - k_{0J_0})(\widehat x_{\hat\tau}) \ne \\
&\quad \ne \exp(\hat\tau k_{0J_0})_*\pi_*\phi_{0J_0\,*}(\delta\ell_2) - \langle \mathrm{d}\tau_\nu(\widehat\ell_0)\,,\ \delta\ell_2\rangle(k_\nu - k_{0J_0})(\widehat x_{\hat\tau}) - \\
&\qquad - \langle \mathrm{d}\theta^\nu_{10}(\widehat\ell_0)\,,\ \delta\ell_2\rangle(k_{10} - k_\nu)(\widehat x_{\hat\tau})
\end{aligned}$$

These claims can be proved by a contradiction argument, using the explicit expressions for the piecewise linearized map $(\pi\mathcal{H}_{\hat\tau})_*$ and the fact that the second variation on

$V_{10} \cap V_{0J_0}^{\perp_{J''_\nu}}$, $\nu = 1, 2$, is positive definite (see the proof of Lemma 6.2 in [15] for details). $\square$

We can now complete the proof of the local invertibility of $\pi\mathcal{H}_{\hat{\tau}}$. As previously said the proof depends on the dimension of the kernel of the map $\mathrm{d}(\tau_1 - \tau_2)(\widehat{\ell}_0)\Big|_{T_{\widehat{\ell}_0}\Lambda}$.

Differentiating (3.1)–(3.2) one easily gets the following formulas

$$\begin{aligned} \langle \mathrm{d}\theta^1_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle &= \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle - \langle \mathrm{d}(\tau_1 - \tau_2)(\widehat{\ell}_0)\,,\,\delta\ell\rangle \frac{\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_2\right)(\widehat{\ell}_0)}{\boldsymbol{\sigma}\left(\overrightarrow{H}_1, \overrightarrow{G}_{10}\right)(\widehat{\ell}_0)}, \\ \langle \mathrm{d}\theta^2_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle &= \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle - \langle \mathrm{d}(\tau_2 - \tau_1)(\widehat{\ell}_0)\,,\,\delta\ell\rangle \frac{\boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_1\right)(\widehat{\ell}_0)}{\boldsymbol{\sigma}\left(\overrightarrow{H}_2, \overrightarrow{G}_{10}\right)(\widehat{\ell}_0)}. \end{aligned} \tag{5.1}$$

which are crucial to the proof. In particular notice that $\langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle = \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle$ implies $\langle \mathrm{d}\theta^\nu_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle = \langle \mathrm{d}\tau_\nu(\widehat{\ell}_0)\,,\,\delta\ell\rangle$, $\nu = 1, 2$.

Case 1. Consider the generic case when $\mathrm{d}(\tau_1 - \tau_2)(\widehat{\ell}_0)\Big|_{T_{\widehat{\ell}_0}\Lambda} \not\equiv 0$. We need to express the boundaries between the adjacent sectors M^0, $M^{\nu j}$.

(i) The boundary between M^0 and M^{11} is given by

$$\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda\colon 0 = \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle \le \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle\};$$

(ii) The boundary between M^0 and M^{21} is given by

$$\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda\colon 0 = \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle \le \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle\};$$

(iii) The boundary between M^{11} and M^{12} is given by

$$\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda\colon \langle \mathrm{d}\theta^1_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle = 0,\ \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle \le \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle\};$$

(iv) The boundary between M^{21} and M^{22} is given by

$$\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda\colon \langle \mathrm{d}\theta^2_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle = 0, \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle \le \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle\};$$

(v) The boundary between M^{12} and M^{22} is given by

$$\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda\colon \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle = \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle \le 0\}.$$

According to Theorem 7.6 in the Appendix, 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 maintain the orientation and that there exists a point $\overline{\delta y}$ whose preimage according $(\pi\mathcal{H}_{\hat{\tau}})_*$ is a singleton that belongs to at most two of the above defined sectors.

Notice that the continuity of $\pi\mathcal{H}_{\hat{\tau}}$ follows from the very definition of the maximized flow. Discontinuities of $(\pi\mathcal{H}_{\hat{\tau}})_*$ may occur only at the boundaries described above, but a direct computation shows that this is not the case. Let us now prove the existence of a $\overline{\delta y}$ with the required properties.

For "symmetry" reasons it is convenient to look for the vector $\overline{\delta y}$ among those which belong to the image of the set $\{\delta\ell \in T_{\widehat{\ell}_0}\Lambda\colon 0 < \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle = \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle\}$.

Let $\overline{\delta\ell} \in T_{\widehat{\ell}_0}\Lambda$ such that $0 < \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\overline{\delta\ell}\rangle = \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\overline{\delta\ell}\rangle$ and let $\overline{\delta y} := L^0\overline{\delta\ell}$. Clearly $\overline{\delta y}$ has at most one preimage per each of the above polyhedral cones. Let us prove that actually its preimage is the singleton $\{\overline{\delta\ell}\}$. In fact we show that for $\nu, j = 1, 2$, there is no $\delta\ell \in M^{\nu j}$ such that $L^{\nu j}(\delta\ell) = \overline{\delta y}$.

$j = 1$. Fix $\nu \in \{1, 2\}$ and assume, by contradiction, that there exists $\delta\ell \in M^{\nu 1}$ such that $L^{\nu 1}\delta\ell = \overline{\delta y}$. The contradiction is shown exactly as in the proof of Claim 1 in Lemma 5.1, by using the explicit expression for the piecewise linearized map $(\pi\mathcal{H}_{\hat{\tau}})_*$ and the fact that the second variation on $V_{10} \cap V_{0J_0}^{\perp_{J''_\nu}}$ is positive definite (see Section 5 of [15] for detailed computations of the second variation).

$j = 2$. Fix $\nu \in \{1, 2\}$ and assume, by contradiction, that there exists $\delta\ell \in M^{\nu 2}$ such that $L^{\nu 2}\delta\ell = \overline{\delta y}$. Let $\overline{\delta x} := \pi_*\overline{\delta\ell}$, and $\delta x := \pi_*\delta\ell$. Taking the pull-back along the reference flow at time $\widehat{\tau}$, $L^{\nu 2}\delta\ell = \overline{\delta y}$ is equivalent to assuming that

$$\overline{\delta x} - \sum_{s=1}^{J_0} \langle \mathrm{d}\theta_{0s}(\widehat{\ell}_0)\,,\,\overline{\delta\ell}\rangle (g_{0s} - g_{0,s-1})(\widehat{x}_0) = \delta x - \sum_{s=1}^{J_0} \langle \mathrm{d}\theta_{0s}(\widehat{\ell}_0)\,,\,\delta\ell\rangle (g_{0,s} - g_{0,s-1})(\widehat{x}_0) -$$
$$- \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle (h_\nu - g_{0J_0})(\widehat{x}_0) - \langle \mathrm{d}\theta^\nu_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle (g_{10} - h_\nu)(\widehat{x}_0).$$

Let $\delta e := (\overline{\delta x} - \delta x, a, b)$, where,

$$a_{0s} := \begin{cases} \langle \mathrm{d}(\theta_{0,s+1} - \theta_{0s})(\widehat{\ell}_0)\,,\,\overline{\delta\ell} - \delta\ell\rangle & s = 0, \ldots, J_0 - 1, \\ \langle \mathrm{d}\theta_{0J_0}(\widehat{\ell}_0)\,,\,\overline{\delta\ell} - \delta\ell\rangle - \langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle & s = J_0, \end{cases}$$

$$b := -\langle \mathrm{d}(\theta^\nu_{10} - \tau_\nu)(\widehat{\ell}_0)\,,\,\delta\ell\rangle, \qquad a_{1s} := \begin{cases} \langle \mathrm{d}\theta^\nu_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle & s = 0, \\ a_{1s} = 0 & s = 1, \ldots, J_1. \end{cases}$$

Then $\delta e \in V_{10} \cap V_{0J_0}^{\perp_{J''_\nu}}$ and Lemma 4.2 applies, so that

$$0 < J''_\nu[\delta e]^2 = \langle \mathrm{d}(\theta^\nu_{10} - \tau_\nu)(\widehat{\ell}_0)\,,\,\delta\ell\rangle \langle \mathrm{d}\tau_\nu(\widehat{\ell}_0)\,,\,\overline{\delta\ell}\rangle \boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_\nu\right)(\widehat{\ell}_0) -$$
$$- \langle \mathrm{d}\theta^\nu_{10}(\widehat{\ell}_0)\,,\,\delta\ell\rangle \Big(\langle \mathrm{d}\theta^\nu_{10}(\widehat{\ell}_0)\,,\,\overline{\delta\ell}\rangle \boldsymbol{\sigma}\left(\overrightarrow{H}_\nu, \overrightarrow{G}_{10}\right)(\widehat{\ell}_0) +$$
$$+ \langle \mathrm{d}\tau_\nu(\widehat{\ell}_0)\,,\,\overline{\delta\ell}\rangle \boldsymbol{\sigma}\left(\overrightarrow{G}_{0J_0}, \overrightarrow{H}_{3-\nu}\right)(\widehat{\ell}_0) \Big),$$

a contradiction, since all the addenda are negative.

By Theorem 7.6 this proves the invertibility of $\pi\mathcal{H}_{\hat{\tau}}$, hence $\mathrm{id} \times \pi\mathcal{H}$ is one–to–one in a neighborhood of $[0, \hat{\theta}_{10} - \varepsilon] \times \{\widehat{\ell}_0\}$.

Case 2. Assume now that the non generic case $T_{\widehat{\ell}_0}\Lambda \subset \ker \mathrm{d}(\tau_1 - \tau_2)(\widehat{\ell}_0)$ holds. The generalized Jacobian $\partial(\pi\mathcal{H}_{\hat{\tau}})(\widehat{\ell}_0)$ (in the sense of Clarke, see [6]) of $\pi\mathcal{H}_{\hat{\tau}} \colon \Lambda \to M$ at $\widehat{\ell}_0$ is the closed convex hull of the linear maps L^0, $L^{\nu j}$, $\nu, j = 1, 2$. We distinguish between two sub–cases:

Case 2.1. $\langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle = \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle = 0$ for any $\delta\ell \in T_{\widehat{\ell}_0}\Lambda$

In this case, by (5.1) we have $\mathrm{d}\theta^1_{10}(\widehat{\ell}_0)|_{T_{\widehat{\ell}_0}\Lambda} \equiv \mathrm{d}\theta^2_{10}(\widehat{\ell}_0)|_{T_{\widehat{\ell}_0}\Lambda} \equiv 0$, hence the linear maps $L^{\nu j}$ for $\nu, j = 1, 2$, coincide with the map L^0, so that $\pi\mathcal{H}_{\hat{\tau}}$ is differentiable at $\widehat{\ell}_0$. The invertibility of L^0 and Clarke's invertibility theorem yield the claim.

Case 2.2. $\langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\,\delta\ell\rangle = \langle \mathrm{d}\tau_2(\widehat{\ell}_0)\,,\,\delta\ell\rangle$ for any $\delta\ell \in T_{\widehat{\ell}_0}\Lambda$ but $\ker(\mathrm{d}\tau_1(\widehat{\ell}_0)|_{T_{\widehat{\ell}_0}\Lambda}) \neq T_{\widehat{\ell}_0}\Lambda$. In this case, by (5.1) we have $\mathrm{d}\theta^1_{10}(\widehat{\ell}_0)|_{T_{\widehat{\ell}_0}\Lambda} \equiv \mathrm{d}\theta^2_{10}(\widehat{\ell}_0)|_{T_{\widehat{\ell}_0}\Lambda} \equiv \mathrm{d}\tau_1(\widehat{\ell}_0)|_{T_{\widehat{\ell}_0}\Lambda}$ so that $L^{12} \equiv L^{22}$.

Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of $T_{\widehat{x}_0}M$ such that $\langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_* v_1\rangle = 1$ and $\langle \mathrm{d}\tau_1(\widehat{\ell}_0)\,,\ \mathrm{d}\alpha_* v_i\rangle = 0$ for $i = 2, \ldots, n$. We show that $\partial(\pi\mathcal{H}_{\hat{\tau}})(\widehat{\ell}_0)$ consists of invertible matrices by showing that

$$(L^0)^{-1}\big(t_0L^0 + t_1L^{11} + t_2L^{21} + t_3L^{12} + t_4L^{22}\big) \circ \mathrm{d}\alpha_* \tag{5.2}$$

is invertible for any $t_0, \ldots, t_4 \geq 0$ such that $\sum_{i=0}^4 t_i = 1$.

Let c_i^ν, $\nu = 1, 2$, $i = 1, \ldots, n$ be such that $(h_\nu - g_{0J_0})(\widehat{x}_0) = \sum_{i=1}^n c_i^\nu v_i$. For each $\nu = 1, 2$ we have

$$\begin{aligned}
&(L^0)^{-1}L^{\nu 1}\,\mathrm{d}\alpha_* v_1 = v_1 - (h_\nu - g_{0J_0})(\widehat{x}_0) = (1 - c_1^\nu)v_1 - \sum_{k=2}^n c_k^\nu v_k\,,\\
&(L^0)^{-1}L^{\nu 2}\,\mathrm{d}\alpha_* v_1 = v_1 - (h_\nu - g_{0J_0})(\widehat{x}_0) - (g_{10} - h_\nu)(\widehat{x}_0) =\\
&\qquad\qquad = (1 - c_1^1 - c_1^2)v_1 - \sum_{k=2}^n (c_k^1 + c_k^2)v_k,\\
&(L^0)^{-1}L^{\nu j}\,\mathrm{d}\alpha_* v_i = v_i \quad \text{ for } i = 2, \ldots, n \text{ and } \nu,\ j = 1, 2.
\end{aligned}$$

Thus the determinant of the matrix in (5.2) is given by $t_0 + t_1 \det(L^0)^{-1}L^{11}\,\mathrm{d}\alpha_* + t_2 \det(L^0)^{-1}L^{21}\,\mathrm{d}\alpha_* + (t_3 + t_4)\det(L^0)^{-1}L^{12}\,\mathrm{d}\alpha_*$ which cannot be zero since all the addenda are positive as it follows from Lemmas 5.1 and 7.1. This concludes the proof of the invertibility of $\pi\mathcal{H}_{\hat{\tau}}$ whereas the invertibility of $\pi\mathcal{H}_{\hat{\theta}_{1j}}$, $j = 1, \ldots, J_1$ follows the same lines. Therefore the proof is omitted here (although it can be found in [15], Section 6).

6. Proof of Theorem 2.1. Let

$$\mathrm{id} \times \pi\mathcal{H} \colon [0, T] \times \mathcal{O} \to \mathcal{V} = [0, T] \times \mathcal{U}$$

be one–to–one and let $\xi \colon [0, T] \to M$ be an admissible trajectory whose graph is in $\mathcal{V}$. Let us recall that applying the Hamiltonian methods, as explained in Section 2.2 we get

$$C(\xi, u) - C(\widehat{\xi}, \widehat{u}) \geq \mathcal{F}(\xi(T)) - \mathcal{F}(\widehat{x}_f).$$

where $\mathcal{F} := \theta_T \circ (\pi\mathcal{H}_T)^{-1} + \beta$, thus, to complete the proof of Theorem 2.1 it suffices to show that $\mathcal{F}$ has a local minimum at $\widehat{x}_f$. For the sake of simplicity put $\psi_T := (\pi\mathcal{H}_T)^{-1}$.

Theorem 6.1. *$\mathcal{F}$ has a strict local minimum at $\widehat{x}_f$.*

Proof. It suffices to prove that

$$\mathrm{d}\,\mathcal{F}(\widehat{x}_f) = 0 \text{ and } \mathrm{D}^2\mathcal{F}(\widehat{x}_f) > 0\,. \tag{6.1}$$

The first equality in (6.1) is an immediate consequence of the definition of $\mathcal{F}$ and of PMP. In fact, since $\mathrm{d}(\theta_T \circ \psi_T) = \mathcal{H}_T \circ \psi_T$, we have $\mathrm{d}\,\mathcal{F}(\widehat{x}_f) = \mathcal{H}_T(\widehat{\ell}_0) + \mathrm{d}\beta(\widehat{x}_f) = 0$. Moreover

$$\begin{aligned}
\mathrm{D}^2\mathcal{F}(\widehat{x}_f)[\delta x_f]^2 &= \big((\mathcal{H}_T \circ \psi_T)_* + \mathrm{D}^2\beta\big)\,(\widehat{x}_f)[\delta x_f]^2 =\\
&= \boldsymbol{\sigma}\,\left((\mathcal{H}_T \circ \psi_T)_*\delta x_f, \mathrm{d}(-\beta)_*\delta x_f\right).
\end{aligned} \tag{6.2}$$

From Lemma 4.4 we get

$$\begin{aligned}
\boldsymbol{\sigma}\Big(\ &\mathrm{d}(-\widehat{\beta})_*\big(\delta x + \sum_{i=0}^1\sum_{s=0}^{J_i} a_{is}g_{is}(\widehat{x}_0) + bh_\nu(\widehat{x}_0)\big)\,,\\
&\qquad \mathrm{d}\alpha_*\delta x + \sum_{i=0}^1\sum_{s=0}^{J_i} a_{is}\overrightarrow{G}_{is}(\widehat{\ell}_0) + b\overrightarrow{H}_\nu(\widehat{\ell}_0)\Big) < 0.
\end{aligned} \tag{6.3}$$

Applying $\widehat{\mathcal{H}}_{T*}$ to both arguments and using the anti–symmetry property of $\boldsymbol{\sigma}$ we get

$$\boldsymbol{\sigma}\left(\mathcal{H}_{T*}\,\mathrm{d}\alpha_*\delta x, \mathrm{d}(-\beta)_*((\pi\mathcal{H}_T)_*\,\mathrm{d}\alpha_*\delta x)\right) > 0$$

which is exactly (6.2) choosing $\delta x := \pi_*\psi_{T*}\delta x_f$. □

To conclude the proof of Theorem 2.1 we have to show that $\widehat{\xi}$ is a strict minimizer. Assume $C(\xi, u) = C(\widehat{\xi}, \widehat{u})$. Since $\widehat{x}_f$ is a strict minimizer for F, then $\xi(T) = \widehat{x}_f$ and equality must hold in (2.8):

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

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. The control u takes the value $\widehat{u}_{|(\hat{\theta}_{1J_1},T)}$ 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 $s = \hat{\theta}_{1,J_1}$. At such time, again by regularity assumptions, u must switch to $\widehat{u}|_{(\hat{\theta}_{1,J_1-1},\hat{\theta}_{1,J_1})}$, so that $\xi(s) = \widehat{\xi}(s)$ also for s in a left neighborhood of $\hat{\theta}_{1,J_1}$. Proceeding backward in time, with an induction argument we finally get $(\xi(s), u(s)) = (\widehat{\xi}(s), \widehat{u}(s))$ for any $s \in [0, T]$.

In the abnormal case the cost is zero, thus the existence of a strict local minimizer implies that the trajectory is isolated among admissible ones.

7. Appendix: Invertibility of piecewise C^1 maps. This Section is devoted to piecewise linear maps and to piecewise C^1 maps. Our aim is to prove a sufficient condition, in terms of the "piecewise linearization", of piecewise C^1 maps.

Some linear algebra preliminaries are needed. The straightforward proof of the following fact can be found in [15], Lemma 7.1:

LEMMA 7.1. *Let A and B be linear automorphisms of $\mathbb{R}^n$. Assume that for some $v \in (\mathbb{R}^n)^* \setminus \{0\}$, A and B coincide on the hyperplane $\{x \in \mathbb{R}^n \colon \langle v\,,\ x\rangle = 0\}$. Then, the map $\mathcal{L}_{AB}$ defined by $x \mapsto Ax$ if $\langle v\,,\ x\rangle \geq 0$, and by $x \mapsto Bx$ if $\langle v\,,\ x\rangle \leq 0$, is a homeomorphism if and only if* $\det(A) \cdot \det(B) > 0$.

Let $G : \mathbb{R}^n \to \mathbb{R}^n$ be continuous and such that there exists a decomposition $S_1, \ldots, S_k$ of $\mathbb{R}^n$ in closed polyhedral cones (intersection of half spaces, hence convex) with nonempty interior and common vertex in the origin and such that $\partial S_i \cap \partial S_j = S_i \cap S_j$ for $i \neq j$, and $G(x) = L_i x$ for all $x \in S_i$, where $L_1, \ldots, L_k$ are linear maps such that $L_i x = L_j x$ for any $x \in S_i \cap S_j$, and $\det L_i \neq 0$, $\forall i = 1, \ldots, k$. We call any such map *continuous piecewise linear*.

Observe that any continuous piecewise linear map G as above is differentiable in $\mathbb{R}^n \setminus \cup_{i=1}^k \partial S_i$. Moreover 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 in [11], Chapter 5, is still valid if the assumption on the compactness of the manifolds is replaced with the assumption that G is proper,). Also, $\deg(G, \mathbb{R}^n, p)$ is constant with respect to p so we simply denote it by $\deg(G)$. We also assume that $\det L_i > 0$ for any $i = 1, \ldots, k$.

LEMMA 7.2. *If G is as above, then $\deg(G) > 0$. In particular, if there exists $q \neq 0$ such that its preimage $G^{-1}(q)$ is a singleton that belongs to at most two of the convex polyhedral cones S_i, 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 $A := G(S_1) \setminus \cup_{i=1}^k G\big(\partial S_i\big)$ is non–empty.

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

$$\deg(G) = \sum_{y \in G^{-1}(x)} \operatorname{sign}\left(\det\left(\mathrm{d}G(y)\right)\right) = \#G^{-1}(x) > 0 \tag{7.1}$$

since $G^{-1}(x) \neq \emptyset$. The second part of the assertion follows taking $x = q$ in (7.1).

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 $i \neq j$. By assumption $p \neq 0$ does not belong to any cone ∂S_s for $s \notin \{i, j\}$, 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 7.1; by the assumption on the signs of the determinants of L_i and L_j, $\mathcal{L}_{L_i L_j}$ 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 $1 = \deg(\mathcal{L}_{L_i L_j}) = \deg(\mathcal{L}_{L_i L_j}, V, p) = \deg(G, V, p) = \deg(G)$, as claimed. □

7.1. Piecewise differentiable functions.

LEMMA 7.3. *Let A and B be linear endomorphisms of $\mathbb{R}^n$. Assume that for some $v \in \mathbb{R}^n \setminus \{0\}$, A and B coincide on the hyperplane $\{x \in \mathbb{R}^n : \langle x\,,\ v\rangle = 0\}$. Then*

$$\det(tA + (1-t)B) = t \det A + (1-t)\det B \quad \forall t \in \mathbb{R}.$$

Proof. Without loss of generality we can assume that $|v| = 1$ and choose vectors $w_2, \ldots, w_n \in \mathbb{R}^n \setminus \{0\}$ such that $v, w_2, \ldots, w_n$ is an orthonormal basis of $\mathbb{R}^n$. In this basis, for $t \in [0, 1]$ we can represent the operator $tA + (1-t)B$ in matrix form:

$$\begin{pmatrix} ta_{11} + (1-t)b_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ta_{n1} + (1-t)b_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} = \begin{pmatrix} ta_{11} + (1-t)b_{11} & b_{12} & \ldots & b_{1n} \\ \vdots & \vdots & & \vdots \\ ta_{n1} + (1-t)b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix}$$

Thus, if A_{i1} and B_{i1} represent the $(i1)$-th cofactor of A and B respectively, then $A_{i1} = B_{i1}$ for $i = 1, \ldots, n$ therefore

$$\det\big(tA + (1-t)B\big) = \sum_{i=1}^{n} (-1)^{i+1}\big(ta_{i1} + (1-t)b_{i1}\big) \det A_{i1} = t \det A + (1-t)\det B$$

as claimed. □

Lemmas 7.1 and 7.3 imply the following fact:

LEMMA 7.4. *Let A and B be linear automorphisms of $\mathbb{R}^n$. Assume that for some $v \in \mathbb{R}^n \setminus \{0\}$, A and B coincide on the hyperplane $\{x \in \mathbb{R}^n \colon \langle x\,,\ v\rangle = 0\}$. Assume that 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. Then,* $\det(A) \cdot \det\big(tA + (1-t)B\big) > 0$ *for any $t \in [0, 1]$.*

Let $\sigma_1, \ldots, \sigma_r$ be a family of C^1–regular pairwise transverse hyper-surfaces in $\mathbb{R}^n$ with $x_0 \in \cap_{i=1}^r \sigma_i$ 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 $U_1, \ldots, U_k$.

Let $f : \overline{U} \to \mathbb{R}^n$ be a continuous map such that there exist $f_1, \ldots, f_k \in C^1(\overline{U})$ with the property that

$$f(x) = f_i(x),\ x \in \overline{U}_i, \text{ and } f_i(x) = f_j(x) \text{ for any } x \in \overline{U}_i \cap \overline{U}_j. \tag{7.2}$$

Notice that such a function is $PC^1(\overline{U})$ (see e.g. [9] for a definition) and Lipschitz continuous in U.

Let $S_1, \ldots, S_k$ be the tangent cones (in the sense of Boulingand) at x_0 to the sets $U_1, \ldots, U_k$, (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

$$F(x) = \mathrm{d}f_i(x_0)x \qquad x \in S_i. \tag{7.3}$$

so that F is a continuous piecewise linear map (compare [9]).

One can see that f is Bouligand differentiable and that its B-derivative is the map F (compare [9, 12]). 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 7.5. *Let f and F be as in* (7.2)–(7.3)*, and assume* $\det \mathrm{d}f_i(x_0) > 0$ *for* $i = 1, \ldots, k$. *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,* $\deg\big(f, B(x_0, \rho), y_0\big) = \deg(F)$.

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 $m := \inf\{|F(v)| : \ |v| = 1\} = \min_{i=1,\ldots,k} \|df_i\| > 0$. 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. □

THEOREM 7.6. *Let f and F be as in* (7.2)–(7.3) *and assume* $\det \mathrm{d}f_i(x_0) > 0$. *Assume also that there exists $p \in \mathbb{R}^n$ whose preimage belongs to at most two of the convex polyhedral cones S_i and such that $F^{-1}(p)$ is a singleton. Then f is a Lipschitzian homeomorphism in a sufficiently small neighborhood of x_0.*

Proof. From Lemmas 7.2–7.5, it follows that $\deg(f, B(x_0, \rho), y_0) = 1$ for sufficiently small $\rho > 0$. By Theorem 4 in [12], we obtain the assertion. □

REFERENCES

[1] Andrei A. Agrachev and Revaz V. Gamkrelidze. Symplectic geometry for optimal control. In Hector 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. An invariant second variation in optimal control. *Internat. J. Control*, 71(5):689–715, 1998.

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

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

[7] Ursula Felgenhauer, Laura Poggiolini, and Gianna Stefani. Optimality and stability result for bang–bang optimal controls with simple and double switch behaviour. *Control and Cybernetics*, 4, 2009.

[8] Magnus R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to calculus of variations. *Pacific J. Math*, 1:525–581, 1951.

[9] 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.

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

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

[12] 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.

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

[14] Laura Poggiolini and Marco Spadini. *Sufficient optimality conditions for a bang-bang trajectory in a Bolza Problem*, chapter In Mathematical Control Theory and Finance, Sarychev, A.; Shiryaev, A.; Guerra, M.; Grossinho, M.d.R. (Eds.), pages 337–358. Springer, 2008.

[15] Laura Poggiolini and Marco Spadini. Bang–bang trajectories with a double switching time: sufficient strong local optimality conditions. arXiv:1010.1149v1[math.OC], October 2010.
[16] Laura Poggiolini and Gianna Stefani. State-local optimality of a bang-bang trajectory: a Hamiltonian approach. *Systems & Control Letters*, 53:269–279, 2004.
[17] Laura Poggiolini and Gianna Stefani. Sufficient optimality conditions for a bang-bang trajectory. In *45th IEEE Conference on Decision and Control*, pages on cd–rom. IEEE, 2006.
[18] Andrei V. Sarychev. First and second order sufficient optimality conditions for bang-bang controls. *SIAM J. Control Optimization*, 1(35):315–340, 1997.