STRONG OPTIMALITY FOR A BANG-BANG TRAJECTORY[*]

ANDREI A. AGRACHEV[†], GIANNA STEFANI[‡], AND PIERLUIGI ZEZZA[§]

Abstract. In this paper we give sufficient conditions for a bang-bang regular extremal to be a strong local optimum for a control problem in the Mayer form; strong means that we consider the C^0 topology in the state space. The controls appear linearly and take values in a polyhedron, and the state space and the end point constraints are finite-dimensional smooth manifolds. In the case of bang-bang extremals, the kernel of the first variation of the problem is trivial, and hence the usual second variation, which is defined on the kernel of the first one, does not give any information. We consider the finite-dimensional subproblem generated by perturbing the switching times, and we prove that the sufficient second order optimality conditions for this finite-dimensional subproblem yield local strong optimality. We give an explicit algorithm to check the positivity of the second variation which is based on the properties of the Hamiltonian fields.

Key words. optimal control, bang-bang controls, sufficient optimality condition, strong local optima

AMS subject classifications. Primary, 49K15; Secondary, 49K30, 58E25

PII. S036301290138866X

1. Introduction. This paper is part of a general research program whose aim is to further extend the use of Hamiltonian methods in the study of optimal control problems. We believe that these methods can play a relevant role in control theory because they allow a general approach to sufficient conditions for strong local optimality, as we wish to show here.

The Hamiltonian approach to strong optimality consists of constructing a field of state extremals covering a neighborhood of a given trajectory which has to be tested. This field of extremals is obtained by projecting on the state manifold M the flow $\mathcal{H}_t$ of the maximized Hamiltonian emanating from the Lagrangian submanifold of the initial transversality conditions. If this projection admits a Lipschitz continuous local inverse, then we can estimate the variation of the cost function at a neighboring trajectory by a function ψ which depends only on the final point, and it is hence independent of the control differential equation; in this way we reduce the problem to a finite-dimensional one. The existence of a Lipschitz continuous local inverse is guaranteed by the surjectivity of the projection on M of the tangent map to the flow $\mathcal{H}_t$. This construction corresponds to the classical one of a nonselfintersecting family of state extremals. This is enough to obtain optimality if the final point is fixed since the submanifold of the final end points reduces to a singleton; otherwise we need some further optimality condition on the function ψ.

We use the relations existing between a suitable second variation and the symplectic properties of the Hamiltonian flow to show that when this second variation is positive definite then the projection on the state manifold M of the tangent map to

[†]Steklov Math. Inst., Gubkina ul.8, 117966 Moscow, Russia, and SISSA, Via Beirut 4, 34014 Trieste, Italy (agrachev@sissa.it).

[‡]Dipartimento di Matematica Applicata, G. Sansone, Via S. Marta 3, 50139 Firenze, Italy (stefani@dma.unifi.it).

[§]Dipartimento di Matematica per le Decisioni, Via C. Lombroso 6/17, 50134 Firenze, Italy (pzezza@unifi.it, http://www.dmd.unifi.it/zezza).

the flow $\mathcal{H}_t$ is surjective; moreover the positivity of this second variation leads also to the sufficient optimality conditions for the function ψ.

To make this general approach possible we need an intrinsic formulation of the second variation as an accessory linear-quadratic minimization problem on the tangent space; this will allow us to exploit one of the crucial ideas underlying the Hamiltonian approach: the tangent map of the flow of the maximized Hamiltonian is the linear Hamiltonian flow of an associated linear-quadratic problem, i.e., the flow of the Jacobi system.

Another important issue is that, when the initial point is not free, it is not possible to cover a neighborhood of the initial point by the projection of the Hamiltonian flow. In the calculus of variations this problem has been solved by perturbing the initial time, but this method does not always work in optimal control because the projection could be singular for a time interval of positive length; this is always the case for bang-bang controls if there is a constraint on the initial point. We propose a different approach: when the second variation is positive we add a penalty term, which allows us to reduce the original problem to another one without constraints on the initial point.

Some of these issues have already been addressed. In [ASZ98b] we stated sufficient conditions for strong local optimality for an optimal control problem in $\mathbb{R}^n$ with unbounded controls, while in [ASZ98a] we gave an intrinsic expression of the accessory problem and studied the relations between the Hamiltonian flow and the index of the second variation. The geometric properties of the field of extremals necessary for proving sufficient conditions for strong optimality were studied in [ASZ99].

In this paper we study a control problem in the Mayer form where the controls appear linearly and take values in a polyhedron, the state space and the end point constraints are finite-dimensional smooth submanifolds, and we give sufficient conditions for a bang-bang extremal to be a strong local minimizer.

In the bang-bang case we have to face some new problems. Since the maximized Hamiltonian is not smooth at the switching points we need to give conditions (see Assumptions 2.1, 2.2, 2.3) which assure us that its flow is defined and piecewise smooth around the reference adjoint covector. Moreover, in the case of bang-bang extremals, the kernel of the first variation of the problem is trivial, and hence the usual second variation, which is defined on the kernel of the first one, does not give any information. We solve this problem by considering the finite-dimensional subproblem generated by perturbing the switching times. The usual (finite-dimensional) second order optimality conditions for this problem give an appropriate second variation. Indeed we prove that the positivity of this second variation yields that the Hamiltonian flow has the properties we have described so that we can prove strong local optimality for the reference trajectory. The set of admissible variations on which we test the second variation can be very small, its dimension can be less than the state space dimension, and when it is zero, we directly have optimality.

By introducing an analogue of the strict Legendre condition, Assumption 2.3, we can eliminate the control from the extremality conditions for the second variation, and its extremals are then described by a discrete version of the Jacobi system, (2.9); the flow of this system describes the tangent subspaces to the flow of the maximized Hamiltonian at the points of nonsmoothness. Since the optimality can be lost only at these points, then the positivity of the second variation can be checked by an algorithm (see Lemma 2.8) which is based on the properties of the discrete flow of the bang-bang Jacobi system. For analogous conditions in the case of unbounded controls, see [ASZ98b].

The literature on second order sufficient conditions for the optimality of a bang-bang trajectory is scarce; we refer to [PS00] and the references therein for results based on the existence of a regular synthesis, and to [Sar92] and [Sar97], where the author studies local minima in the L^1 norm on the control in the time-optimal case. For a general description of the classical study of strong local optimality in the one-dimensional calculus of variations, see [GH96a, GH96b].

2. Statement of the results. Let X_i, $i = 1, 2, \ldots, m$, be distinct C^∞ vector fields defined on the C^∞ finite-dimensional manifold M and let $\Delta = co\,\{e_1,\, e_2,\, \ldots,\, e_m\}$ be the unitary simplex in $\mathbb{R}^m$.

We are interested in the optimal control problem

$$\text{Minimize} \quad J(\xi) := c_0(\xi(0)) + c_T(\xi(T))$$

subject to

(2.1) $$\dot{\xi}(t) = \sum_{i=1}^m u_i(t)\, X_i(\xi(t)), \quad u(t) \in \Delta$$

(2.2) $$\xi(0) \in N_0\,, \ \xi(T) \in N_T,$$

where the time interval $[0, T]$ is fixed, N_0, N_T are given C^∞ submanifolds of M, and c_0, c_T are real-valued smooth functions. We will give sufficient conditions for a trajectory to be a strong local optimum, where strong means that we consider the C^0 topology in the state space.

As a candidate optimal solution we are given a *bang-bang Pontryagin extremal* $(\hat{\xi}, \hat{u})$, that is, an absolutely continuous solution $\hat{\xi} : [0, T] \to M$ of system (2.1)–(2.2) with corresponding control $\hat{u}$ satisfying the Pontryagin maximum principle (PMP); moreover there is a partition of $[0, T]$

$$0 = t_0 < t_1 < t_2 < \cdots < t_r < t_{r+1} = T$$

such that

$$\hat{u}(t) = e_{j_i}, \quad t \in (t_{i-1}, t_i),$$

for some $j_i \in \{1, 2, \ldots, m\}$. Therefore $\hat{\xi}$ is a solution of

(2.3) $$\dot{\xi}(t) = X_{j_i}(\xi(t)), \quad t \in [t_{i-1}, t_i],$$

in each subinterval. The values t_i for $i = 1, 2, \ldots, r$ will be called *switching times*, and we set

$$x_0 := \hat{\xi}(0), \quad x_T := \hat{\xi}(T)$$

to simplify notation. Corresponding to the reference extremal we define the time-dependent vector field

$$\hat{h} : [0, T] \times M \to TM \quad \text{as} \quad \hat{h}_{|(t_{i-1}, t_i)} := X_{j_i},$$

and we set $h_i := X_{j_i}$. Therefore the reference trajectory is a solution of the differential equation

(2.4) $$\dot{\xi}(t) = \hat{h}_t(\xi(t)), \quad t \in [0, T].$$

By lifting the vector field $\hat{h}_t$ to the cotangent bundle, we define the time-dependent Hamiltonian

$$\widehat{H} : [0,T] \times T^*M \to \mathbb{R}, \quad (t,\ell) \mapsto \langle \ell, \hat{h}_t(\pi\ell) \rangle,$$

where $\pi : T^*M \to M$ is the canonical projection; let us set $H_i := \widehat{H}_{|(t_{i-1},t_i)}$.

For our problem the maximized Hamiltonian

$$H : T^*M \to \mathbb{R}, \quad \ell \mapsto \max_{u \in \Delta} \left\langle \ell, \sum_{i=1}^m u_i \, X_i(\pi\ell) \right\rangle$$

is well defined and Lipschitz.

Recall that any piecewise smooth Hamiltonian $H_t : T^*M \to \mathbb{R}$ defines a Hamiltonian vector field $\vec{H}_t$ whose flow will be denoted by $\mathcal{H}_t$. Moreover for any time-independent vector field Y we denote its flow by $(t,x) \mapsto \exp t\,Y(x)$; see [Arn80].

We can express the PMP by saying that there exist $p_0 \in \{0,1\}$ and a lift $\hat{\lambda}$ of $\hat{\xi}$ to the cotangent bundle, which is a solution of

$$\dot{\lambda}(t) = \vec{\hat{H}}_t(\lambda(t)),$$

(2.5) $$\lambda(0) = p_0 \, dc_0(x_0) \quad \text{on } T_{x_0} N_0,$$

(2.6) $$\lambda(T) = -p_0 \, dc_T(x_T) \quad \text{on } T_{x_T} N_T$$

such that

$$|p_0| + \|\hat{\lambda}\| \neq 0,$$
$$\widehat{H}_t(\hat{\lambda}(t)) = H(\hat{\lambda}(t)).$$

Let us now introduce our first assumption.

ASSUMPTION 2.1 (bang-bang regular extremal). *The maximum*

$$\max_{u \in \Delta} \left\langle \hat{\lambda}(t), \sum_{i=1}^m u_i \, X_i(\hat{\xi}(t)) \right\rangle$$

is attained at a vertex of Δ *for all* $t \in [0,T]$, $t \neq t_1, t_2, \ldots, t_r$.

Assumption 2.1 means that on each subinterval (t_{i-1}, t_i) there is a unique index j_i such that

$$H(\hat{\lambda}(t)) = \langle \hat{\lambda}(t), \, X_{j_i}(\hat{\xi}(t)) \rangle.$$

The smooth functions $p_0 \, c_0$ and $p_0 \, c_T$ are defined on N_0 and N_T, respectively, but they can be extended to the whole manifold M in such a way that the transversality conditions (2.5) and (2.6) hold on the whole tangent space. We denote by $\alpha, \beta : M \to \mathbb{R}$ two functions such that

$$\alpha = p_0 \, c_0 \quad \text{on } N_0, \quad \beta = p_0 \, c_T \quad \text{on } N_T,$$

(2.7) $$\hat{\lambda}(0) = d\,\alpha(x_0) \quad \text{on } T_{x_0}M, \quad \hat{\lambda}(T) = -d\,\beta(x_T) \quad \text{on } T_{x_T}M.$$

Consider the two Lagrangian submanifolds

$$\Lambda_0 := \left\{ d\,\alpha(x) + (T_x N_0)^\perp \mid x \in N_0 \right\},$$
$$\Lambda_T := \left\{ -d\,\beta(x) + (T_x N_T)^\perp \mid x \in N_T \right\};$$

the transversality conditions (2.5) and (2.6) of the PMP can be equivalently stated by saying that

$$\lambda(0) \in \Lambda_0, \quad \lambda(T) \in \Lambda_T.$$

In the normal case ($p_0 = 1$) α, β are cost functions equivalent to the original ones, while in the abnormal case ($p_0 = 0$) they are extensions of the zero function. When $p_0 = 0$ all the costs disappear, and indeed we will study a problem with a zero cost; therefore, proving that $\hat{\xi}$ is a *strict* strong minimizer will imply that it is isolated with respect to the C^0 topology among the admissible trajectories. In the case of sub-Riemannian metrics isolated trajectories are called rigid geodesics.

The sufficient conditions will be derived by studying the following optimal control problem, which is equivalent to the original one:

(P)   Minimize $J(\xi) := \alpha(\xi(0)) + \beta(\xi(T))$

subject to (2.1) and (2.2).

The points

$$\ell_i := \hat{\lambda}(t_i), \quad i = 0, 1, \dots, r+1,$$

will be called the *switching points* of the adjoint covector $\hat{\lambda}$. From the PMP we can deduce the following relations, which represent necessary optimality conditions:

$$H_i(\ell_i) = H_{i+1}(\ell_i), \quad i = 1, 2, \dots, r,$$
$$\langle d(H_{i+1} - H_i), \vec{H}_{i+1} \rangle(\ell_i) \geq 0, \quad i = 1, 2, \dots, r.$$

To state sufficient conditions for $\hat{\xi}$ to be a strong local minimizer we need to strengthen these two conditions, and hence we assume the following.

ASSUMPTION 2.2 (simple switching points). *The maximum*

$$\max_{u \in \Delta} \left\langle \hat{\lambda}(t), \sum_{i=1}^{m} u_i X_i(\hat{\xi}(t)) \right\rangle$$

is attained along a one-dimensional edge of Δ for $t = t_1, t_2, \dots, t_r$.

ASSUMPTION 2.3 (strict bang-bang Legendre condition).

$$\langle d(H_{i+1} - H_i), \vec{H}_{i+1} \rangle(\ell_i) > 0, \quad i = 1, 2, \dots, r.$$

Remark 2.4. The PMP implies that the switching point ℓ_i belongs to the level set $H_{i+1} - H_i = 0$ for $i = 1, 2, \dots, r$. The strict bang-bang Legendre condition yields that, near the point ℓ_i, this level set is a hypersurface which will be called the *switching surface*.

Our assumptions are strictly related to the properties of the flow of the maximized Hamiltonian H, and they guarantee that the Hamiltonian flow is piecewise smooth; see Corollary 4.2. In particular, Assumptions 2.1 yields that locally around the reference extremal we can switch from one vector field to another only on the switching surfaces, while Assumption 2.2 yields that, on the switching surfaces, we can choose only between two vector fields, and the last one, Assumption 2.3, yields that we are forced to switch. Let $\boldsymbol{s}$ be the Liouville one form in T^*M and denote by $\boldsymbol{\sigma} = d\boldsymbol{s}$ the

canonical symplectic two form on T^*M; see [Arn80] for the definitions. Taking into account the basic properties of $\boldsymbol{\sigma}$, Assumption 2.3 can be equivalently written as

$$\langle d(H_{i+1}-H_i), \vec{H}_{i+1}\rangle(\ell_i)$$
$$= \boldsymbol{\sigma}\left(\vec{H}_i, \vec{H}_{i+1}\right)(\ell_i)$$
$$= \{H_i, H_{i+1}\}(\ell_i)$$
$$= \langle \ell_i, [h_i, h_{i+1}](\hat{\xi}(t_i))\rangle \geq 0,$$

where $\{\ ,\ \}$ and $[\ ,\ ]$ denote the Poisson and Lie brackets, respectively.

2.1. A finite-dimensional subproblem. We are going to choose an appropriate r-dimensional family of variations corresponding to bang-bang trajectories; they are generated by perturbing the switching times. The optimality with respect to these variations will be not only necessary but also sufficient (under the previously stated assumptions) to prove that the reference trajectory is a strong local minimizer.

For a given $a > 0$ such that $\min_{i=1,\ldots,r+1}(t_i - t_{i-1}) > 2a$, let $\varepsilon \in B(0,a) \subset \mathbb{R}^r$, set $\varepsilon_0 = \varepsilon_{r+1} = 0$, and consider the time-dependent vector field

$$(\varepsilon, x) \mapsto h_t(\varepsilon, x) = h_i(x) \quad \text{if} \quad t \in (t_{i-1} + \varepsilon_{i-1}, t_i + \varepsilon_i).$$

This new vector field is obtained from the reference one by moving the switching time t_i by ε_i.

Remark 2.5. A small ε corresponds to a control variation which is small in the L^1 norm but not in the L^∞ norm.

Denote the flow of $\dot{\xi}(t) = h_t(\varepsilon, \xi(t))$ by

$$S_t : M \times B(0, a) \to M$$

and consider the following finite-dimensional subproblem of problem (P):

(sub-P)   Minimize $\gamma(x, \varepsilon) := \alpha(x) + \beta(S_T(x, \varepsilon))$

subject to

$$x \in N_0, \quad S_T(x, \varepsilon) \in N_T.$$

Note that $\widehat{S}_t := S_t(\cdot, 0)$ is the flow of $\hat{h}_t$, $S_t(x_0, 0) = \hat{\xi}(t)$ and $(x_0, 0)$ is the candidate optimal solution for the subproblem.

By using the relations (2.7) and the extremality properties of the reference trajectory, it is easy to prove that $(x_0, 0)$ is a critical point for γ, that is,

$$d\gamma(x_0, 0) = 0,$$

and hence

$$J'' := \frac{1}{2} D^2 \gamma(x_0, 0)$$

is a well-defined quadratic form on $T_{x_0} M \times \mathbb{R}^r$, which gives the second order approximation of γ. The second variation of (sub-P) is the restriction of J'' to the linearization of the constraints; namely, if we set

(2.8) $\mathcal{N} = \left\{(\delta x, \varepsilon) \in T_{x_0} N_0 \times \mathbb{R}^r : S_{T*}(\delta x, \varepsilon) \in T_{x_T} N_T\right\},$

then the second variation of (sub-P) is $J''_{|\mathcal{N}}$, and it will be called the *second variation at the switching points*. Let us remark that $J''_{|\mathcal{N}} \geq 0$ is a necessary optimality condition when the subproblem is normal.

The main result of the paper states that under the regularity conditions on the maximized Hamiltonian previously stated, the positivity of the second variation at the switching points is sufficient to prove that the reference trajectory $(\hat{\xi}, \hat{u})$ is a strict local minimizer for the original problem in the C^0 topology on the state (strong minimizer).

From an intuitive point of view the idea underlying this result can be summarized by saying that the flow of the maximized Hamiltonian projects onto the trajectories of the finite-dimensional subproblem, which then generates a field of extremals that can be used to prove sufficiency.

THEOREM 2.6. *Assume that the given bang-bang Pontryagin extremal $\hat{\xi}$ is regular and has simple switching points and that the strict bang-bang Legendre condition is satisfied. If the second variation at the switching points is positive definite, then $\hat{\xi}$ is a strict strong local minimizer, i.e., a strict local minimizer in the C^0 topology. In the abnormal case $\hat{\xi}$ is an isolated admissible solution of the constrained control system.*

Remark 2.7. If $\mathcal{N}$ reduces to $\{0\}$, then the second variation at the switching points is positive definite, and hence we obtain a first order sufficient condition.

2.2. The bang-bang Jacobi system. We can check the positivity of the second variation at the switching points in a complete Hamiltonian form. Let

$$\Pi := T_{\pi\,\ell} M \hookrightarrow T_\ell T^* M$$

be the *vertical* subspace and define

$$L_0 := T_{\ell_0} \Lambda_0, \quad L_T := T_{\ell_T} \Lambda_T.$$

The regularity assumptions on the maximized Hamiltonian yield that $\mathcal{H}_t$ is smooth everywhere except at the switching times where it is left and right smooth; see Corollary 4.2. The positivity of the second variation at the switching points can be checked through the properties of the tangent subspaces to $\mathcal{H}_{t_k}(\Lambda_0)$ from the left and from the right and by their relative positions with respect to Π. Thanks to the strict bang-bang Legendre condition, these Lagrangian subspaces can be described through the flow of the following discrete version of the Jacobi system:

$$(2.9) \quad \begin{cases} \delta\ell_k^- = \left(\exp(t_{k+1} - t_k)\,\vec{H}_{k+1}\right)_* \delta\ell_{k-1}^+, \\[2mm] \delta\ell_k^+ = \delta\ell_k^- + \dfrac{\boldsymbol{\sigma}(\delta\ell_k^-, (\vec{H}_k - \vec{H}_{k+1})(\ell_k))}{\boldsymbol{\sigma}(\vec{H}_k, \vec{H}_{k+1})(\ell_k)}\,(\vec{H}_k - \vec{H}_{k+1})(\ell_k). \end{cases}$$

Denote the flow of $\delta\ell_k^-$ and $\delta\ell_k^+$ by Δ_k^-, Δ_k^+ and set

$$L_k^- := \Delta_k^- L_0, \quad L_k^+ := \Delta_k^+ L_0.$$

In section 4 we prove that L_k^+, L_k^- are the left and right tangent subspaces to $\mathcal{H}_{t_k}(\Lambda_0)$; see Remark 4.6.

LEMMA 2.8. *The positivity of the second variation at the switching points can be checked through the following algorithm.*
STEP 1: Set $k = 1$.
STEP 2: If $k \leq r$, then go to STEP 3
        else go to STEP 4.
STEP 3: If $(\vec{H}_k - \vec{H}_{k+1})(\ell_k) \notin L_k^- + \Pi$ or $(\vec{H}_k - \vec{H}_{k+1})(\ell_k) \in L_k^-$, then set $k = k + 1$ and go to STEP 2.
        else
           if $L_k^+ \cap \Pi \subseteq L_k^- \cap \Pi$ and for every $\delta\ell^+ \in L_k^+$, $\delta\ell^- \in L_k^-$ such that $\pi_* \delta\ell^+ = \pi_* \delta\ell^-$ we have that $\boldsymbol{\sigma}(\delta\ell^-, \delta\ell^+) \geq 0$, then set $k = k + 1$ and go to STEP 2.
           else $J''_{|\mathcal{N}}$ is not positive definite, STOP.
STEP 4: If for every $\delta\ell \in L_{r+1}^-$, $\delta\ell_T \in L_T$ such that $\pi_* \delta\ell = \pi_* \delta\ell_T \neq 0$ we have that $\boldsymbol{\sigma}(\delta\ell, \delta\ell_T) > 0$, then $J''_{|\mathcal{N}}$ is positive definite, END.
        else $J''_{|\mathcal{N}}$ is not positive definite, STOP.

Remark 2.9. Let us explain the meaning of each step of this algorithm.
1. The algorithm first checks the positivity of the second variation associated to the corresponding problem with fixed final point (STEP 3).
2. Each iteration of the algorithm is associated to a new variation obtained by perturbing the corresponding switching time; this procedure generates an increasing family of variations.
3. STEP 3 deserves some comment: if $(\vec{H}_k - \vec{H}_{k+1})(\ell_k) \notin L_k^- + \Pi$, then there is no new variation, and hence there is no condition to check; if $(\vec{H}_k - \vec{H}_{k+1})(\ell_k) \in L_k^-$, then the flow $\mathcal{H}_t$ is differentiable also at t_k and the properties of the second variation remain unchanged.
4. STEP 4 checks the positivity conditions related to the presence of a nontrivial final cost, and hence when the final point is fixed, STEP 4 is void and the algorithm becomes the following.
    STEP 1: $k = 1$.
    STEP 2: If $k \leq r$, then go to STEP 3
            else $J''_{|\mathcal{N}}$ is positive definite, END.
    STEP 3: If $(\vec{H}_k - \vec{H}_{k+1})(\ell_k) \notin L_k^- + \Pi$ or $(\vec{H}_k - \vec{H}_{k+1})(\ell_k) \in L_k^-$, then set $k = k + 1$ and go to STEP 2.
            else
                if $L_k^+ \cap \Pi \subseteq L_k^- \cap \Pi$ and for every $\delta\ell^+ \in L_k^+$, $\delta\ell^- \in L_k^-$ such that $\pi_* \delta\ell^+ = \pi_* \delta\ell^-$ we have $\boldsymbol{\sigma}(\delta\ell^-, \delta\ell^+) \geq 0$, then set $k = k + 1$ and go to STEP 2.
                else $J''_{|\mathcal{N}}$ is not positive definite, STOP.

Since the maximized Hamiltonian is a piecewise lift of a vector field on M, then the vertical directions remain vertical under the action of the flow, and at the switching points the dimension of the projection increases at most by one. To obtain a flow which projects locally onto M we will reduce the problem to an equivalent one with free initial point; for this reason we describe explicitly the algorithm in this special case.

COROLLARY 2.10. *If the initial point is free, we have that $L_0 \cap \Pi = \{0\}$, and hence the algorithm becomes the following.*
STEP 1: Set $k = 1$.
STEP 2: If $k \leq r$, then go to STEP 3
        else go to STEP 4.

```
STEP 3: If (H⃗ₖ − H⃗ₖ₊₁)(ℓₖ) ∈ Lₖ⁻ , then set k = k + 1 and go to STEP 2.
        else
            if π∗Lₖ⁺ = T_ξ̂(tₖ)M and for every δℓ⁺ ∈ Lₖ⁺, δℓ⁻ ∈ Lₖ⁻ such that
        π∗δℓ⁺ = π∗δℓ⁻ we have that σ(δℓ⁻,δℓ⁺) ≥ 0, then set k = k + 1 and
        go to STEP 2.
            else J''|N is not positive definite, STOP.
STEP 4: If for every δℓ ∈ L⁻_{r+1}, δℓ_T ∈ L_T such that π∗δℓ = π∗δℓ_T ≠ 0 we
        have that σ(δℓ,δℓ_T) > 0, then J''|N is positive definite, END.
        else J''|N is not positive definite, STOP.
```

Remark 2.11. In STEP 3 we check the fixed final point problem, and the algorithm stops when we find a direction on which the quadratic form is negative or zero. For this reason we call the corresponding switching time t_k the *conjugate point*; a conjugate point can occur only at a switching time.

2.3. The Bolza problem. We deal with an optimal control problem in the Mayer form only for simplicity; all the results can be stated for a problem in the Bolza form when the cost function includes an integral term, that is,

$$\text{Minimize} \quad J(\xi) := c_0(\xi(0)) + c_T(\xi(T)) + \int_0^T \sum_{i=1}^m u_i(t)\, X_i^0(\xi(t))\, dt$$

subject to (2.1) and (2.2), where X_i^0, $i = 1, 2, \ldots, m$, are C^∞ functions defined on M.

The same proofs can be carried out using as reference and maximized Hamiltonian those defined as

$$\widehat{H} : \ell \mapsto \langle \ell, \hat{h}(\pi\ell)\rangle - p_0 \sum_{i=1}^m \hat{u}_i(t)\, X_i^0(\pi\ell),$$

$$H : \ell \mapsto \max_{u \in \Delta} \left(\left\langle \ell, \sum_{i=1}^m u_i\, X_i(\pi\ell) \right\rangle - p_0 \sum_{i=1}^m u_i\, X_i^0(\pi\ell) \right).$$

3. The second variation at the switching points. This section is necessarily technical, but it contains the main ideas and the technical lemmas needed to carry out this kind of approach.

To study the relations existing between the second variation at the switching points and the properties of the Hamiltonian flow, let us reduce (sub-P) to a single-input affine problem with piecewise constant control maps having the t_i's as switching times. This reduction can be achieved by the following time reparametrization:

$$\dot{\varphi}(\tau) = 1 + \nu(\tau), \quad \nu \in (-1, 1),$$
$$\varphi(0) = 0, \quad \varphi(T) = T,$$

where ν is piecewise constant, i.e., $\nu(\tau) \equiv \nu_i$, $\tau \in [t_{i-1}, t_i)$.

Any solution of this boundary value problem is an increasing isomorphism of the interval $[0, T]$ onto itself. If we set $\varepsilon_i := \varphi(t_i) - t_i$, $i = 1, 2, \ldots, r$, then we have that $S_{\varphi(\tau)}(x, \varepsilon)$ is the solution of the differential equation

$$(3.1) \qquad\qquad \dot{\zeta}(\tau) = [1 + \nu(\tau)]\, \hat{h}_\tau(\zeta(\tau)), \quad \zeta(0) = x.$$

If we set $u_i := \nu_i \left(t_i - t_{i-1}\right)$, then from our construction it follows that $\varepsilon_i = \sum_{j=1}^{i} u_j$ and $\sum_{j=1}^{r+1} u_j = 0$ as it follows from the boundary condition $\varphi(T) = T$. Therefore we can take as control space the r-dimensional vector space

$$U := \left\{ (u_1, u_2, \ldots, u_{r+1}) \in \mathbb{R}^{r+1} \; \middle| \; \sum_{j=1}^{r+1} u_j = 0 \right\}.$$

For a given $u \in U$ we denote by ν_u the corresponding control map. Since there is a one-to-one correspondence between ε and u we still denote by $\mathcal{N}$ the subset of $T_{x_0} N_0 \times U$ corresponding to $\mathcal{N} \subseteq T_{x_0} N_0 \times \mathbb{R}^r$, which is defined in (2.8). We can now study the second variation of the problem of minimizing $\alpha(\zeta(0)) + \beta(\zeta(T))$ subject to (3.1) with the boundary conditions $\zeta(0) \in N_0$ and $\zeta(T) \in N_T$.

Following the same approach used in [ASZ98a] we can define the second variation as a linear quadratic problem on $T_{x_0} M$ by the using pull-back system defined through the time-dependent vector field

$$(3.2) \qquad \hat{g}_t := \widehat{S}_{t*}^{-1} \, \hat{h}_t \circ \widehat{S}_t.$$

$\hat{g}_t$ is piecewise constant with the same switching times as $\hat{h}_t$, and we set $g_i := \hat{g}_{|(t_{i-1}, t_i)}$. Consider the pull-back control system

$$(3.3) \qquad \dot{\eta}(t) = \nu(t) \, \hat{g}_t(\eta(t))$$

and the associated linearized equation at $\eta(t) \equiv x_0$,

$$(3.4) \qquad \dot{\delta\eta}(t) = \nu(t) \, \hat{g}_t(x_0).$$

If we also pull back the costs by setting

$$\hat{\beta} = \beta \circ \widehat{S}_T, \quad \hat{\gamma} = \alpha + \hat{\beta},$$

then, reasoning as in [ASZ98a], the second variation at the switching points can be equivalently written as the restriction to $\mathcal{N}$ of the linear-quadratic form

$$J''[\delta e]^2 = \frac{1}{2} D^2 \hat{\gamma}(x_0)[\delta x]^2 + \int_0^T \nu_u(s) \, \langle Q_s, \delta\eta_s(\delta e) \rangle \, ds,$$

where $\delta e := (\delta x, u) \in T_{x_0} M \times U$ and

$$\langle Q_t, \delta x \rangle = \langle D \langle D \, \hat{\beta}, \hat{g}_t \rangle(x_0), \, \delta x \rangle.$$

The Hamiltonian associated with this linear-quadratic problem is

$$(\omega, \delta x, u) \mapsto G_t''(\omega, \delta x) \nu_u(t),$$

where G_t'' is the following piecewise constant linear Hamiltonian:

$$G_t'' : T_{x_0}^* M \times T_{x_0} M \to \mathbb{R}, \quad (\omega, \delta x) \mapsto \langle \omega, \hat{g}_t(x_0) \rangle + \langle Q_t, \delta x \rangle.$$

With notation analogous to previous ones, we set $G_i'' := G''_{|(t_{i-1}, t_i)}$ and define the Lagrangian subspace of the initial and final transversality conditions as

$$L_0'' := \left\{ \left(-D^2 \hat{\gamma}(x_0)(\delta x, \cdot) + \omega, \, \delta x\right) \; \middle| \; \delta x \in T_{x_0} N_0, \; \omega \in \left(T_{x_0} N_0\right)^{\perp} \right\},$$

$$L_T'' := \left(T_{x_0} \widehat{S}_T^{-1}(N_T)\right)^{\perp} \times T_{x_0} \widehat{S}_T^{-1}(N_T).$$

We want to express the value of the form J'' in Hamiltonian notation. Let $\delta e = (\delta x, u)$ and $\delta f = (\delta y, v)$ belong to $T_{x_0} M \times U$ and let $\delta \ell \in L_0''$ be such that

$$\pi_* \delta \ell = \delta x.$$

If we denote by $\mathcal{G}_t''(\delta \ell, u) := (\omega_t(\delta \ell, u), \delta \eta_t(\delta e))$ the solution of the Hamiltonian system

$$(3.5) \qquad \dot{\lambda}(t) = \vec{G}_t''(\lambda(t)) \, \nu_u(t), \quad \lambda(0) = \delta \ell,$$

then we obtain, as in the proof of Lemma 4 in [ASZ98a, p. 700],

$$\begin{aligned}
J''(\delta e, \delta f) &= D^2 \hat{\gamma}(x_0)(\delta x, \delta y) + \int_0^T \langle Q_s, \nu_u(s) \delta \eta_s(\delta f) + \nu_v(s) \delta \eta_s(\delta e) \rangle \, ds \\
(3.6) \qquad &= D^2 \hat{\gamma}(x_0)(\delta x, \delta y) - \langle \omega_T(\delta \ell, u), \delta \eta_T(\delta f) \rangle + \langle \omega_0(\delta \ell, u), \delta y \rangle \\
&\quad + \int_0^T G_t'' \Big(\mathcal{G}_t''(\delta \ell, u) \Big) \nu_v(t) \, dt.
\end{aligned}$$

The positivity of the second variation at the switching points will be checked in two steps. We first consider the problem with fixed final point and check the positivity of the corresponding second variation, that is, J'' restricted to

$$V := \{ (\delta x, u) \in T_{x_0} N_0 \times U \mid \delta \eta(\delta x, \nu_u, T) = 0 \} \subseteq \mathcal{N}.$$

Afterwards we check the positivity of J'' on $\mathcal{N} \cap V^{\perp_{J''}}$, where $\perp_{J''}$ means orthogonality with respect to J''.

To study the signature of the second variation on V we take an increasing sequence of subspaces $V_k \subset V$ obtained by considering as admissible controls those u for which ν_u is zero from t_{k+1} on; i.e., we will study the second variation on each

$$V_k := \{ (\delta x, u) \in V \mid u_j = 0 \text{ for } j \geq k + 2 \}.$$

The extremals of J'' on V are essential in the study of its signature, and they are those δe belonging to $V \cap V^{\perp_{J''}}$. For this reason we characterize the J''-orthogonality in the following integral version of the Jacobi system.

LEMMA 3.1. *For $k \in \{1, 2, \ldots, r\}$, $\delta e = (\delta x, u) \in \mathcal{N} \cap V_k^{\perp_{J''}}$ if and only if there exists $\delta \ell \in L_0''$ such that*

$$(3.7) \qquad \pi_* \delta \ell = \delta x, \quad \pi_* \mathcal{G}_T''(\delta \ell, u) \in \pi_* L_T'',$$

$$(3.8) \qquad \int_0^T G_t'' \Big(\mathcal{G}_t''(\delta \ell, u) \Big) \nu_v(t) \, dt = 0 \quad \forall v: \ v_j = 0, \ j \geq k + 2.$$

Proof. $\delta e \in \mathcal{N} \cap V_k^{\perp_{J''}}$ if and only if there exist $\bar{\omega}_0 \in (T_{x_0} N_0)^{\perp}$ and $\bar{\omega}_T \in \Pi$ such that

$$J''(\delta e, \delta f) = \langle \bar{\omega}_0, \, \delta y \rangle + \langle \bar{\omega}_T, \, \delta \eta_T(\delta f) \rangle$$

for all $\delta f = (\delta y, v) \in T_{x_0} M \times U$ such that $v_j = 0, \ j \geq k + 2$.
If we choose

$$\delta \ell = (-D^2 \hat{\gamma}_0(\delta x, \cdot) + \bar{\omega}_0, \, \delta x),$$

then $\delta\ell \in L''_0$ and (3.7) is satisfied. For $v = 0$, from (3.6) we obtain

$$-\omega_T(\delta\ell, u) = \bar{\omega}_T$$

and hence (3.8).

To prove the converse, let us remark that (3.7) yields that $\delta e \in \mathbb{N}$; moreover, using (3.6) for $\delta f \in V_k$, from (3.7) and (3.8), it follows that $J''[\delta e, \delta f] = 0$. □

COROLLARY 3.2. *Let* $\delta e = (\delta x, u) \in \mathbb{N} \cap V^{\perp_{J''}}$ *and let* $\delta\ell \in L''_0$ *be the one given in Lemma 3.1. If* $\delta\ell_1 \in L''_T$ *is such that* $\pi_*\delta\ell_1 = \pi_*\mathcal{G}''_T(\delta\ell, u)$, *then*

$$(3.9) \qquad J''\,[\delta e]^2 = \boldsymbol{\sigma}\Big(\delta\ell_1, \mathcal{G}''_T(\delta\ell, u)\Big).$$

Let $\delta e = (\delta x, u) \in V_k \cap V^{\perp_{J''}}_{k-1}$ *and let* $\delta\ell \in L''_0$ *be the one given in Lemma 3.1. Then*

$$(3.10) \qquad J''[\delta e]^2 = \boldsymbol{\sigma}\Big(\mathcal{G}''_{t_k}(\delta\ell, u), u_{k+1}(\vec{G}''_{k+1} - \vec{G}''_k)\Big).$$

Proof. Equality (3.9) is an easy consequence of Lemma 3.1 and (3.6). Integrating by parts and using the symplectic properties of the Hamiltonian flow, again from (3.6), it follows that

$$
\begin{aligned}
J''[\delta e]^2 &= \int_{t_{k-1}}^{t_k} G''_k\Big(\mathcal{G}''_t(\delta\ell, u)\Big)\,\frac{-u_{k+1}}{t_k - t_{k-1}}\,dt + \int_{t_k}^{t_{k+1}} G''_{k+1}\Big(\mathcal{G}''_t(\delta\ell, u)\Big)\,\nu_u(t)\,dt \\
&= -u_{k+1}G''_k\Big(\mathcal{G}''_{t_k}(\delta\ell, u)\Big) + u_{k+1}G''_{k+1}\Big(\mathcal{G}''_{t_{k+1}}(\delta\ell, u)\Big) \\
&= u_{k+1}\Big(G''_{k+1} - G''_k\Big)\Big(\mathcal{G}''_{t_k}(\delta\ell, u)\Big).
\end{aligned}
$$

Equality (3.10) now follows thanks to the symplectic properties of the Hamiltonian flow. □

Let us remark that (3.7) characterizes those δe in $\mathbb{N}$, while (3.8) characterizes those in $V_k^{\perp_{J''}}$. In particular the extremals of the second variation are described by those $\delta\ell \in L''_0$, $u \in U$ such that $\mathcal{G}''_T(\delta\ell, u) \in L''_T$, and

$$(3.11) \qquad \int_0^T G''_t\Big(\mathcal{G}''_t(\delta\ell, u)\Big)\,\nu_v(t)\,dt = 0 \quad \forall\, v \in U.$$

The relations between the second variation and the Hamiltonian of the original problem can be better understood by using the following map:

$$\boxed{\imath : T^*_{x_0}M \times T_{x_0}M \to T_{\ell_0}T^*M, \quad (\omega, \delta x) \mapsto -\omega + d(-\hat{\beta})_*\delta x}$$

It is easy to check that the map $\imath$ is an antisymplectic isomorphism

$$(3.12) \qquad \boldsymbol{\sigma}\,(\imath\,\delta\ell_1, \imath\,\delta\ell_2) = -\boldsymbol{\sigma}\,(\delta\ell_1, \delta\ell_2)$$

and that it is an isomorphism between L''_0 and L_0 which acts as

$$\imath\,\delta\ell = d\alpha_* \pi_* \delta\ell.$$

The map $\imath$ connects the Hamiltonians associated with the second variation with the original ones through the following relation:

$$(3.13) \qquad \imath\,\vec{G}''_k = \widehat{\mathcal{H}}^{-1}_{t_{k*}}\vec{H}_k(\ell_k) = \widehat{\mathcal{H}}^{-1}_{t_{k-1*}}\vec{H}_k(\ell_{k-1}).$$

Equation (3.13) can be proved starting from the equality

$$\vec{G}_k'' = \left(-D\langle D\,\hat{\beta}, g_k\rangle(x_0),\ g_k(x_0) \right)$$

and applying the map $\imath$ to obtain, in coordinates,

$$\imath\,\vec{G}_k'' = \left(d\hat{\beta}(x_0)\,Dg_k(x_0), g_k(x_0) \right);$$

finally since $d\hat{\beta}(x_0) = -d\alpha(x_0)$, then (3.13) follows.

Thanks to the above properties of the map $\imath$ we can restate the *strict bang-bang Legendre condition* as

(3.14)
$$\boldsymbol{\sigma}\left(\vec{G}_k'',\ \vec{G}_{k+1}'' \right) < 0, \quad k = 1, 2, \ldots, r.$$

The strict bang-bang Legendre condition allows us to solve recursively equation (3.11) with respect to the control, and hence we are able to define a discrete version of the Jacobi system by substituting this control back into (3.5). The resulting system is defined below, and its construction is described in the subsequent Lemma 3.4.

DEFINITION 3.3. *Suppose that the strict bang-bang Legendre condition is satisfied and consider the discrete dynamical system on $\mathbb{R} \times T^*(T_{x_0} M)$,*

$$
\begin{cases}
\mathrm{w}_k = \dfrac{\boldsymbol{\sigma}\left(\delta\ell_{k-1}, \vec{G}_k'' - \vec{G}_{k+1}'' \right)}{\boldsymbol{\sigma}\left(\vec{G}_k'', \vec{G}_{k+1}'' \right)}, \\[4ex]
\delta\ell_k = \delta\ell_{k-1} + (\vec{G}_k'' - \vec{G}_{k+1}'')\,\mathrm{w}_k.
\end{cases}
$$

For $k = 1, 2, \ldots, r$ we define the flows of w_k and $\delta\ell_k$ as the linear functions

$$\omega_k : L_0'' \to \mathbb{R}$$

and the symplectic isomorphisms

$$\mathcal{G}_T^k : L_0'' \to T_{x_0}^* M \times T_{x_0} M.$$

LEMMA 3.4. *Suppose that the strict bang-bang Legendre condition is satisfied and let $(\delta\ell, u) \in L_0'' \times U$; then (3.8) holds if and only if*

$$u_i = \langle (\omega_i - \omega_{i-1}), \delta\ell \rangle, \quad i = 1, 2, \ldots, k,$$
$$\mathcal{G}_{t_i}''(\delta\ell, u) = \mathcal{G}_T^i(\delta\ell) + \langle \omega_i, \delta\ell \rangle\,\vec{G}_{i+1}'', \quad i = 1, 2, \ldots, k.$$

Proof. From the properties of the Hamiltonian flows, by integrating by parts equality (3.8), it follows that

$$\int_0^T G_t''\left(\mathcal{G}_t''(\delta\ell, u) \right) \nu_v(t)\,dt$$

$$= \sum_{i=1}^k v_i \left(G_i''\left(\mathcal{G}_{t_i}''(\delta\ell, u) \right) - G_{k+1}''\left(\mathcal{G}_{t_{k+1}}''(\delta\ell, u) \right) \right) = 0$$

for all $v \in U$ such that $v_j = 0$, $j \geq k + 2$. Hence (3.8) is equivalent to

$$G_1''\left(\mathcal{G}_{t_1}''(\delta\ell, u) \right) = G_2''\left(\mathcal{G}_{t_2}''(\delta\ell, u) \right) = \cdots = G_{k+1}''\left(\mathcal{G}_{t_{k+1}}''(\delta\ell, u) \right).$$

If we compute explicitly

$$\left(G_i'' - G_{i+1}''\right)\left(\mathcal{G}_{t_{i-1}}''(\delta\ell, u) + u_i\,\vec{G}_i''\right) = 0,$$

we obtain, for each $i = 1, 2, \ldots, k$,

$$\boldsymbol{\sigma}\left(\mathcal{G}_{t_{i-1}}''(\delta\ell, u) + u_i\,\vec{G}_i'', \vec{G}_i'' - \vec{G}_{i+1}''\right) = 0.$$

The equivalence now follows by finite induction and from Definition 3.3. $\square$

Remark 3.5. Let us remark that being J''-orthogonal to V_k implies that the values of the control maps $u_1, u_2, \ldots, u_k$ are uniquely determined by the value of $\delta\ell$. Moreover we obtain the flow of the Hamiltonian system of the second variation up to time t_k through the flow of the discrete bang-bang Jacobi system. More precisely, if we define the control $\mathrm{w}^k : L_0'' \to U$ as

(3.15)
$$\begin{cases} \mathrm{w}_i^k := (\omega_i - \omega_{i-1}), & i = 1, 2, \ldots, k, \\ \mathrm{w}_{k+1}^k := -\omega_k, & \\ \mathrm{w}_i^k := 0, & i \geq k + 2, \end{cases}$$

then from Lemma 3.4 it follows that this control is such that $(\delta\ell, \langle \mathrm{w}^k, \delta\ell\rangle)$ satisfies (3.8) and

$$\mathcal{G}_T^k(\delta\ell) = \mathcal{G}_T''(\delta\ell, \langle \mathrm{w}^k, \delta\ell\rangle).$$

A possible way to check the positivity of J'' on V_k is to study the behavior of J'' on $V_k \cap V_{k-1}^{\perp_{J''}}$. Thanks to the properties of the bang-bang Jacobi system the variations belonging to $V_k \cap V_{k-1}^{\perp_{J''}}$ and the values of J'' can be described through the following subspaces:

$$L_k'' := \mathcal{G}_T^k\,L_0''.$$

The results are given in the two following lemmas. Let us notice that the first statement of the next lemma states that the extremals of J'' on V_k are the solutions of the Jacobi system that become vertical at step k and that the third statement characterizes the occurrence of a new variation.

LEMMA 3.6. *Suppose that the strict bang-bang Legendre condition is satisfied. The following statements hold:*

1. $\delta e = (\delta x, u) \in V_k \cap V_k^{\perp_{J''}}$ *if and only if there exists $\delta\ell \in L_0''$ such that*

$$\pi_* \delta\ell = \delta x, \quad u = \langle \mathrm{w}^k, \delta\ell\rangle, \quad \mathcal{G}_T^k(\delta\ell) \in \Pi.$$

2. $\delta e \in V_k \cap V_{k-1}^{\perp_{J''}}$ *if and only if there exists $\delta\ell \in L_0''$ such that*

$$\pi_* \delta\ell = \delta x, \quad u_j = \langle \mathrm{w}_j^{k-1}, \delta\ell\rangle, \quad j = 1, 2, \ldots, k - 1,$$
$$\mathcal{G}_T^{k-1}(\delta\ell) - u_{k+1}\,(\vec{G}_k'' - \vec{G}_{k+1}'') \in \Pi,$$

and in this case we have that

(3.16) $$J''[\delta e]^2 = \boldsymbol{\sigma}\left(\mathcal{G}_T^{k-1}(\delta\ell) - u_{k+1}\vec{G}_k'', -u_{k+1}(\vec{G}_k'' - \vec{G}_{k+1}'')\right).$$

3. If $J''_{|V_{k-1}} > 0$, then $V_k = V_{k-1}$ if and only if $\vec{G}''_k - \vec{G}''_{k+1} \notin L''_{k-1} + \Pi$.

4. If $\delta\ell_{k-1} \in L''_{k-1}$ and $\delta\ell_k \in L''_k \setminus L''_{k-1}$ are such that

$$\pi_* \delta\ell_{k-1} = \pi_* \delta\ell_k,$$

then there exists a nontrivial $\delta e \in V_k \cap V_{k-1}^{\perp_{J''}}$ such that

$$J[\delta e]^2 = \boldsymbol{\sigma}\left(\delta\ell_k,\, \delta\ell_{k-1}\right).$$

Proof. 1. From (3.7) and (3.8) and from the properties of the control (3.15) it follows that

$$u = \langle \mathrm{w}^k,\, \delta\ell\rangle \quad \text{and} \quad \mathcal{G}''_T(\delta\ell, u) = \mathcal{G}^k_T(\delta\ell) \in \Pi.$$

2. Letting $\delta e \in V_k \cap V_{k-1}^{\perp_{J''}}$ be the first part is an immediate consequence of Lemmas 3.1 and 3.4. Once again from the properties of the control (3.15) we have that

$$\begin{aligned}
\mathcal{G}''_{t_{k+1}}(\delta\ell, u) &= \mathcal{G}^{k-1}_T(\delta\ell) + \langle \omega_{k-1},\, \delta\ell\rangle \vec{G}''_k + u_k \vec{G}''_k + u_{k+1} \vec{G}''_{k+1} \\
&= \mathcal{G}^{k-1}_T(\delta\ell) - u_{k+1}(\vec{G}''_k - \vec{G}''_{k+1}).
\end{aligned}$$

From (3.10) it follows that

$$\begin{aligned}
J''[\delta e]^2 &= \boldsymbol{\sigma}\left(\mathcal{G}''_{t_k}(\delta\ell, u),\, -u_{k+1}(\vec{G}''_k - \vec{G}''_{k+1})\right) \\
&= \boldsymbol{\sigma}\left(\mathcal{G}^{k-1}_T(\delta\ell) - u_{k+1}\vec{G}''_k,\, -u_{k+1}(\vec{G}''_k - \vec{G}''_{k+1})\right).
\end{aligned}$$

3. $\vec{G}''_k - \vec{G}''_{k+1} \notin L''_{k-1} + \Pi$ if and only if

$$\mathcal{G}^{k-1}_T(\delta\ell) - u_{k+1}(\vec{G}''_k - \vec{G}''_{k+1}) \in \Pi \Rightarrow u_{k+1} = 0.$$

From statement 2 we have that $J[\delta e]^2 = 0$, and hence the statement follows.

4. By definition there are $\delta\ell_0,\, \delta\ell_1 \in L''_0$ such that

$$\delta\ell_{k-1} = \mathcal{G}^{k-1}_T(\delta\ell_0), \quad \delta\ell_k = \mathcal{G}^k_T(\delta\ell_1).$$

From the assumptions we have that

$$\mathcal{G}^k_T(\delta\ell_1) - \mathcal{G}^{k-1}_T(\delta\ell_0) = \mathcal{G}^{k-1}_T(\delta\ell_1) - \mathcal{G}^{k-1}_T(\delta\ell_0) + \langle \omega_k, \delta\ell_1\rangle (\vec{G}''_k - \vec{G}''_{k+1}) \in \Pi.$$

If we define $\delta\ell := \delta\ell_1 - \delta\ell_0$ and $\delta e := (\pi_* \delta\ell, \langle \mathrm{w}^k,\, \delta\ell_1\rangle - \langle \mathrm{w}^{k-1},\, \delta\ell_0\rangle)$, then we have that $u_{k+1} = -\langle \omega_k, \delta\ell_1\rangle$, and from statement 2 we have that $\delta e \in V_k \cap V_{k-1}^{\perp_{J''}}$, and it is nontrivial because if $u_{k+1} = 0$, then $\delta\ell_k \in L''_{k-1}$. Moreover we have that

$$\begin{aligned}
J''[\delta e]^2 &= \boldsymbol{\sigma}\left(\mathcal{G}^{k-1}_T(\delta\ell_1 - \delta\ell_0) - u_{k+1}\vec{G}''_k,\, -u_{k+1}(\vec{G}''_k - \vec{G}''_{k+1})\right) \\
&= \boldsymbol{\sigma}\left(-\mathcal{G}^{k-1}_T(\delta\ell_0),\, \mathcal{G}^k_T(\delta\ell_1)\right) + \boldsymbol{\sigma}\left(\mathcal{G}^{k-1}_T(\delta\ell_0),\, \mathcal{G}^{k-1}_T(\delta\ell_1)\right) \\
&\quad - u_{k+1}\boldsymbol{\sigma}\left(\mathcal{G}^{k-1}_T(\delta\ell_1) + \langle\omega_k,\, \delta\ell_1\rangle\vec{G}''_k,\, \vec{G}''_k - \vec{G}''_{k+1}\right).
\end{aligned}$$

The final statement now follows because the second addend is zero since both the arguments belong to the same Lagrangian subspace L''_{k-1} and the third one is zero by the properties of ω_k. $\quad\square$

LEMMA 3.7. $J''_{|V} > 0$ *if and only if one of the following statements holds for each* $k = 1, 2, \ldots, r$:
1. $\vec{G}''_k - \vec{G}''_{k+1} \notin L''_{k-1} + \Pi$.
2. $\vec{G}''_k - \vec{G}''_{k+1} \in L''_{k-1}$.
3. $L''_k \cap \Pi \subseteq L''_{k-1} \cap \Pi$ *and for all* $\delta\ell_k \in L''_k$ *and* $\delta\ell_{k-1} \in L''_{k-1}$ *such that* $\pi_* \delta\ell_k = \pi_* \delta\ell_{k-1}$, *we have that* $\boldsymbol{\sigma}(\delta\ell_k, \delta\ell_{k-1}) \geq 0$.

Proof. The idea of the proof is the following: we first show that these conditions together with $J''_{|V_{k-1}} > 0$ imply that

$$J''_{|V_k \cap V^{\perp_{J''}}_{k-1}} > 0$$

and hence $J''_{|V_k} > 0$; since $V_0 = \{0\}$, then the lemma will follow by finite induction on k. Let us show that the induction step is valid for each k if and only if one of the statements of the lemma holds.

Assume that $J''_{|V_{k-1}} > 0$.

- From statement 2 of Lemma 3.6 it follows that $\vec{G}''_k - \vec{G}''_{k+1} \notin L''_{k-1} + \Pi$ is equivalent to $V_k \cap V^{\perp_{J''}}_{k-1} = \{0\}$ and the induction step is trivial.
- If $\vec{G}''_k - \vec{G}''_{k+1} \in L''_{k-1}$, then $L''_k = L''_{k-1}$; moreover we can choose $u_{k+1} = 1$ in part 2 of Lemma 3.6 to show that $V_k \cap V^{\perp_{J''}}_{k-1} \neq \{0\}$ to obtain a nontrivial $\delta e \in V_k \cap V^{\perp_{J''}}_{k-1}$ such that

$$\begin{aligned} J''[\delta e]^2 &= \boldsymbol{\sigma}\left(\mathcal{G}^{k-1}_T(\delta\ell) - \vec{G}''_k, -(\vec{G}''_k - \vec{G}''_{k+1})\right) \\ &= \boldsymbol{\sigma}\left(\vec{G}''_{k+1}, \vec{G}''_k\right). \end{aligned}$$

Equation (3.14) completes the proof.
- If $\vec{G}''_k - \vec{G}''_{k+1} \in \{L''_{k-1} + \Pi\} \setminus L''_{k-1}$, then $\dim V_k \cap V^{\perp_{J''}}_{k-1} = 1$. From the first statement of Lemma 3.6 it follows that the condition $L''_k \cap \Pi \subseteq L''_{k-1} \cap \Pi$ is equivalent to

$$V_k \cap V^{\perp_{J''}}_k = V_{k-1} \cap V^{\perp_{J''}}_{k-1} = \{0\},$$

and hence it will be enough to prove that $J''_{|V_k \cap V^{\perp_{J''}}_{k-1}} \geq 0$.

Under our assumptions there is $\delta\ell \in L''_{k-1}$ such that $\langle \omega_k, \delta\ell \rangle = 1$. If we set

$$\delta\ell_k := \delta\ell + (\vec{G}''_k - \vec{G}''_{k+1}) \in L''_k,$$

then we can find $\delta\ell_{k-1} \in L''_{k-1}$ such that $\pi_* \delta\ell_{k-1} = \pi_* \delta\ell_k$. From the fourth statement of Lemma 3.6 it follows that

$$J[\delta e]^2 = \boldsymbol{\sigma}\left(\delta\ell_k, \delta\ell_{k-1}\right).$$

Since $\dim V_k \cap V^{\perp_{J''}}_{k-1} = 1$ then $J_{|V_k \cap V^{\perp_{J''}}_{k-1}} \geq 0$ if and only if $\boldsymbol{\sigma}\left(\delta\ell_k, \delta\ell_{k-1}\right) \geq 0$.   □

LEMMA 3.8. *Assume that* $J''_{|V} > 0$; *then the quadratic form* $J''_{|N \cap V^{\perp_{J''}}}$ *is positive definite if and only if for every* $\delta\ell \in L''_0$ *and* $\delta\ell_T \in L''_T$ *such that*

$$\pi_* \delta\ell_T = \pi_* \mathcal{G}^r_T(\delta\ell) \neq 0$$

we have

$$\boldsymbol{\sigma}(\delta\ell_T, \mathcal{G}_T^r(\delta\ell)) > 0.$$

Proof. From Lemma 3.1 and the properties of the control w^k (see (3.15)), we have that $\delta e = (\delta x, u) \in \mathcal{N} \cap V^{\perp_{J''}}$ if and only if there is $\delta\ell \in L_0''$ such that

$$\pi_* \delta\ell = \delta x, \quad u = \langle \mathrm{w}^r, \delta\ell \rangle, \quad \pi_* \mathcal{G}_T^r(\delta\ell) \in \pi_* L_T''.$$

If $\pi_* \mathcal{G}_T^r(\delta\ell) = 0$, then $\delta e \in V \cap V^{\perp_{J''}}$ and hence $\delta e = 0$; otherwise we can use equation (3.9). $\square$

3.1. The algorithm. We have essentially already shown that the algorithm can be used to check the positivity of the second variation at the switching points. This can be easily seen since from (3.12), (3.13), it follows that

$$L_k^- = \widehat{\mathcal{H}}_{t_{k*}} \, \imath \, L_{k-1}'' \quad \text{for } k = 1, 2, \ldots, r+1,$$
$$L_k^+ = \widehat{\mathcal{H}}_{t_{k*}} \, \imath \, L_k'' \quad \text{for } k = 1, 2, \ldots, r.$$

Moreover STEP 3 follows from Lemma 3.7, while STEP 4 follows from Lemma 3.8, taking into account (3.12).

4. Proof of the theorem. In order to demonstrate the Hamiltonian method we now give the proof of our main result step by step following the approach described in the introduction.

All the proofs make strong use of the properties (see [Arn80]) of the Poincaré–Cartan form $\boldsymbol{\omega} = \boldsymbol{s} - H \, dt$ on $I \times T^*M$ associated to the Hamiltonian H. Namely,

- $\boldsymbol{\omega}$ evaluated along a lift of a solution of (2.1) is nonpositive and it is zero along $\hat{\lambda}$;
- $\boldsymbol{\omega}$ is exact on the Legendre submanifold generated by the flow of $\vec{H}$ emanating from a Lagrangian submanifold.

4.1. Flow properties. The first step shows that our assumptions guarantee that the flow of the maximized Hamiltonian is locally well defined and piecewise C^∞ and describes the structure of the switching surfaces.

LEMMA 4.1. *There exists a neighborhood $\mathcal{U}$ of ℓ_0 such that we can define recursively for $i = 1, \ldots, r$ the C^∞-maps*

$$\tau_i : \mathcal{U} \to \mathbb{R} \quad and \quad \phi_i : \mathcal{U} \to T^*M$$

in the following way: set

$$\tau_0 := 0, \quad \phi_0 := Id.$$

The τ_i's are implicitly defined by

$$\begin{cases} (H_i - H_{i+1}) \left(\exp \tau_i(\ell) \vec{H}_i(\phi_{i-1}(\ell)) \right) = 0, \\ \tau_i(\ell_i) = t_i, \end{cases}$$

while the ϕ_i's are defined as

$$\phi_i := \ell \mapsto \exp(-\tau_i(\ell) \vec{H}_{i+1}) \circ \exp \tau_i(\ell) \vec{H}_i \, (\phi_{i-1}(\ell)).$$

The neighborhood $\mathcal{U}$ can be chosen such that

$$(4.1) \qquad \sup_{\ell \in \mathcal{U}} \tau_i(\ell) < \inf_{\ell \in \mathcal{U}} \tau_{i+1}(\ell),$$

and the ϕ_i's are C^∞ symplectic diffeomorphisms.

Proof. Thanks to the strict bang-bang Legendre condition we can apply the implicit function theorem to show that the τ_i's are well defined and C^∞. Therefore, by continuity, we can guarantee that (4.1) holds. Let us show by induction that ϕ_i's are symplectic diffeomorphisms. From the definition of the τ_i's we have that

$$(4.2) \qquad \boldsymbol{\sigma}\left(\vec{H}_i, \vec{H}_{i+1}\right) d\,\tau_i(\ell) = \boldsymbol{\sigma}\left((\exp \tau_i(\ell)\vec{H}_i)_* \phi_{i-1\,*}, \vec{H}_i - \vec{H}_{i+1}\right),$$

and from the definition of the ϕ_i's we have that

$$(4.3) \qquad \begin{aligned} \phi_{i*} &= \exp(-\tau_i(\ell)\vec{H}_{i+1})_* \exp(\tau_i(\ell)\vec{H}_i)_* \phi_{i-1\,*} \\ &+ \left\{\exp(-\tau_i(\ell)\vec{H}_{i+1})_* \vec{H}_i\left(\exp(\tau_i(\ell)\vec{H}_i(\phi_{i-1}(\ell)))\right) - \vec{H}_{i+1}(\phi_i(\ell))\right\} d\,\tau_i(\ell). \end{aligned}$$

The result now follows from (4.2) and from the general fact that $\exp(s\,\vec{G})$ is a symplectic diffeomorphism for any Hamiltonian vector field $\vec{G}$.   □

Let $\mathcal{U}$ be the neighborhood of ℓ_0 given in Lemma 4.1. If we set

$$\mathcal{O}_i := \left\{(t,\ell) \mid \ell \in \mathcal{U}, \ \tau_{i-1}(\ell) \le t \le \tau_i(\ell)\right\} \subseteq [0,T] \times T^*M,$$

then the $\mathcal{O}_i$'s are $2\,n+1$-dimensional C^∞ submanifolds with boundary $\partial\mathcal{O}_i = S_{i-1} \cup S_i$, where

$$S_i := \mathcal{O}_i \cap \mathcal{O}_{i+1} = \left\{(\tau_i(\ell), \ell)\,, \quad \ell \in \mathcal{U}\right\}.$$

From Lemma 4.1 we can easily deduce the following.

COROLLARY 4.2. *Under Assumptions* 2.1, 2.2, *and* 2.3 *the Hamiltonian system*

$$\begin{aligned} \dot{\lambda}(t) &= \vec{H}(\lambda(t)), \\ \lambda(0) &= \ell \end{aligned}$$

has a unique solution, which can be represented on $[0,T] \times \mathcal{U}$ by the map $\mathcal{H}: (t,\ell) \mapsto (t, \mathcal{H}_t(\ell))$ given by

$$(4.4) \qquad \mathcal{H}_t(\ell) = \exp t\vec{H}_{i+1}(\phi_i(\ell)), \quad t \in [\tau_i(\ell), \tau_{i+1}(\ell)],$$

where $\tau_{r+1} \equiv T$; moreover the flow $\mathcal{H}$ is C^∞ on each $\mathcal{O}_i$.

Let us remark that every solution of the Hamiltonian system (4.4) has the same number of switches as the reference trajectory $\hat{\xi}$; moreover, from the above equation (4.4), we can deduce that for $t \in [\tau_{i-1}(\ell), \tau_i(\ell)]$ we can write

$$\mathcal{H}_t(\ell) = \exp(t - \tau_{i-1}(\ell))\vec{H}_i \circ \cdots \circ \exp(\tau_2(\ell) - \tau_1(\ell))\vec{H}_2 \circ \exp \tau_1(\ell)\vec{H}_1\,(\ell)$$

and the ϕ_i's can be written as

$$\phi_i(\ell) = \exp(-\tau_i(\ell)\vec{H}_{i+1}) \circ \mathcal{H}_{\tau_i(\ell)}\,(\ell).$$

Remark 4.3. We can interpret Lemma 4.1 as saying that, thanks to the strict bang-bang Legendre condition, we can define, in a tube around the adjoint covector, a time-dependent maximized Hamiltonian as

$$(t, \ell) \mapsto H_i(t, \ell) \quad \text{if} \quad \mathcal{H}^{-1}(t, \ell) \in \mathcal{O}_i.$$

This Hamiltonian switches from one vector field to another when its flow crosses the switching surfaces and hence when changing the vector field results in an energy increase. Assumptions 2.1 and 2.2 ensure that with this choice we obtain the maximized Hamiltonian.

4.2. Hamiltonian methods. For a general introduction to the use of these methods and their application to optimal control, we refer to [AG90, AG97].

Without loss of generality we can assume that $\Lambda_0 \subseteq \mathcal{U}$ and that Λ_0 is a smooth simply connected Lagrangian submanifold; if necessary we take the restriction to a neighborhood of x_0. Define

$$\Omega_i := \left\{ (t, \ell) \in \mathcal{O}_i \mid \ell \in \Lambda_0 \right\}, \quad \Sigma_i := \Omega_i \cap \Omega_{i+1},$$

and

$$\Omega := \bigcup_{i=1}^{r+1} \Omega_i.$$

The Ω_i's are $n + 1$-dimensional C^∞ submanifolds with boundary $\partial \Omega_i = \Sigma_{i-1} \cup \Sigma_i$.

From (4.4) it follows that $\mathcal{H}_t(\Lambda_0)$ is a Lagrangian submanifold, although it might be not C^1 at the switching surfaces. We now investigate the properties of the Cartan form $\boldsymbol{\omega}$ and of the map

$$\pi_t := \pi \circ \mathcal{H}_t : \Lambda_0 \mapsto M.$$

LEMMA 4.4. *The form $\mathcal{H}^* \boldsymbol{\omega}$ is closed on each Ω_i and hence exact on Ω so that it can be written as*

$$\mathcal{H}^* \boldsymbol{\omega} = d\vartheta,$$

where ϑ is a continuous function on $[0, T] \times \Lambda_0$, which is C^∞ on each Ω_i. Moreover ϑ can be chosen such that

$$\vartheta(0, \cdot) := \vartheta_0 = \alpha \circ \pi.$$

If π_t is Lipschitz invertible, then

$$d(\vartheta_t \circ \pi_t^{-1}) = \mathcal{H}_t \circ \pi_t^{-1}.$$

Proof. The proof of the first statement is a standard consequence of the properties of $\boldsymbol{\omega}$ (see [Arn80]).

Let $\gamma : [a, b] \to M$ be a Lipschitz curve; then from the first part of the lemma it follows that

$$\int_\gamma \mathcal{H}_t \circ \pi_t^{-1} = \int_{\mathcal{H}_t \circ \pi_t^{-1} \circ \gamma} \boldsymbol{s} = \int_{\pi_t^{-1} \circ \gamma} \mathcal{H}_t^* \boldsymbol{s} = \vartheta_t \circ \pi_t^{-1} \Big|_{\gamma(a)}^{\gamma(b)},$$

and the statement follows. $\square$

If $\lambda : [0, T] \to T^*M$ is a Lipschitz lift of a solution ξ of equation (2.1) such that $(t, \lambda(t)) \in \Omega$ for $t \in [0, T]$, then

$$(4.5) \qquad \int_\lambda \boldsymbol{\omega} \leq 0 \quad \text{and} \quad \int_{\hat{\lambda}} \boldsymbol{\omega} = 0$$

because $\boldsymbol{\omega}$ is defined by the maximized Hamiltonian H. From this property and by the previous lemma, we obtain that

$$(4.6) \qquad J(\pi\lambda) - J(\hat{\xi}) \geq \vartheta_T\Big(\mathcal{H}_T^{-1}(\lambda(T))\Big) - \vartheta_T(\ell_0) + \beta(\xi(T)) - \beta(x_T).$$

Hence, as we mentioned in the introduction, the variation of the cost is estimated from below by the function $\vartheta_T + \beta \circ \pi$, which depends only on the final point.

Let us remark that if π_T is invertible, then the same estimate can be obtained by the function

$$\chi := \vartheta_T \circ \pi_T^{-1} + \beta.$$

For this function, Lemma 4.4 and the transversality conditions imply that

$$d\chi(x_T) = 0.$$

4.3. An equivalent free initial point problem. To be able to lift to Ω any trajectory in a neighborhood of the reference one, we need that π_t is locally onto for each t and in particular for $t = 0$. This last condition can be fulfilled by constructing an equivalent problem with a free initial point. Let Q be any nonnegative quadratic form on $T_{x_0}M$, whose nullity is $T_{x_0}N_0$. We extend it to $T_{x_0}M \times \mathbb{R}^r$ by setting $Q[\delta x, \epsilon]^2 = Q[\delta x]^2$. If the quadratic form J'' is positive on $\mathcal{N}$, then we can find $\rho > 0$ such that

$$J'' + \frac{1}{2}\rho Q > 0 \quad \text{on} \ \Big\{(\delta x, \varepsilon) \in T_{x_0}M \times \mathbb{R}^r : S_{T*}(\delta x, \varepsilon) \in T_{x_T}N_T\Big\},$$

as can be easily proved by elementary arguments of linear algebra. Let us choose a function α_ρ such that

$$\begin{aligned} \alpha_\rho &= \alpha \quad \text{on} \ N_0, \\ d\alpha_\rho &= d\alpha \quad \text{on} \ T_{x_0}N_0, \\ D^2\alpha_\rho(x_0) &= D^2\alpha(x_0) + \rho Q \end{aligned}$$

and consider the problem

$$(4.7) \qquad \text{Minimize} \quad \alpha_\rho(\xi(0)) + \beta(\xi(T))$$

subject to

$$\dot{\xi}(t) = \sum_{i=1}^m u_i(t)\, X_i(\xi(t)), \quad u \in \Delta,$$

$$\xi(T) \in N_T.$$

Since the reference trajectory satisfies the initial boundary conditions, then proving that it is optimal for this new problem yields its optimality for the original one. Therefore without loss of generality we can assume that the original problem has already free

initial point, i.e., $N_0 \equiv M$ and $\alpha \equiv \alpha_\rho$; in this case the initial Lagrangian submanifold is horizontal and its projection covers a neighborhood of the initial point x_0.

Remark 4.5. This reduction is possible because the new cost on the initial point contains an exact penalty which can be constructed assuming that the second variation is positive definite.

Let us now see which properties of the symplectic map $\mathcal{H}_{t*} : T_{\ell_0}\Lambda_0 \to T_{\hat{\lambda}(t)}M$ lead to the optimality of $\hat{\xi}$. Let us remark that (4.4) yields, for $\delta\ell \in L_0$,

$$(4.8) \qquad \mathcal{H}_{t*}(\delta\ell) = (\exp t\vec{H}_{i+1})_*\phi_{i*}\delta\ell \quad \text{for } t \in (t_i, t_{i+1}),$$

$$(4.9) \qquad \mathcal{H}_{t_i*}(\delta\ell) = \begin{cases} (\exp t_i\vec{H}_i)_*\phi_{i-1*}\delta\ell & \text{for } \langle d\,\tau_i(\ell_0), \delta\ell \rangle \leq 0, \\ (\exp t_i\vec{H}_{i+1})_*\phi_{i*}\delta\ell & \text{for } \langle d\,\tau_i(\ell_0), \delta\ell \rangle \geq 0. \end{cases}$$

Remark 4.6. By (4.2) and (4.3) one can easily see that $L_k^- = (\exp t_k\vec{H}_k)_*\phi_{k-1*}L_0$ and $L_k^+ = (\exp t_k\vec{H}_{k+1})_*\phi_{k*}$; therefore L_k^- and L_k^+ are tangent to $\mathcal{H}_{t_k}(L_0)$ from the left and from the right, respectively. Moreover if $d\,\tau_k(\ell_0)_{|L_0} = 0$, then the flow is differentiable at (t_k, ℓ_0).

LEMMA 4.7. *If the map* $\pi_* \mathcal{H}_{t*} : L_0 \to T_{\hat{\xi}(t)}M$ *is onto for* $t \in [0, T]$, *then there exists a neighborhood* $\mathcal{V} \subseteq \Lambda_0$ *of* ℓ_0 *such that* $[0, T] \times \mathcal{V}$ *is mapped by* $\pi\mathcal{H}$ *onto a neighborhood of* $\hat{\xi}$ *in* $[0, T] \times M$ *and* $\pi\mathcal{H}$ *has a piecewise* C^∞ *local inverse. Without loss of generality we set* $\mathcal{V} = \Lambda_0$.

Proof. Thanks to the invertibility assumption on $\pi_* \mathcal{H}_{t_i*}$ and by possibly taking a smaller neighborhood of ℓ_0, we can apply the inverse function theorem on each submanifold with boundary Ω_i to show that the image under $\pi\mathcal{H}$ of Ω is a neighborhood of $\hat{\xi}$ in $[0, T] \times M$. $\square$

THEOREM 4.8. *The equality*

$$\pi_*\mathcal{H}_{t*}L_0 = T_{\hat{\xi}(t)}M$$

holds for $t \in [0, T]$ *if and only if the following statements hold for* $i = 1, 2, \ldots, r$:
  1. $\pi_*\phi_{i*}L_0 = T_{x_0}M$.
  2. *If* $\delta\ell_1$, $\delta\ell_2 \in L_0$ *are such that*

$$\pi_*\left[(\exp t_i\vec{H}_i)_*\phi_{i-1*}\delta\ell_1\right] = \pi_*\left[(\exp t_i\vec{H}_{i+1})_*\phi_{i*}\delta\ell_2\right],$$

  then

$$\boldsymbol{\sigma}\left((\exp t_i\vec{H}_i)_*\phi_{i-1*}\delta\ell_1, (\exp t_i\vec{H}_{i+1})_*\phi_{i*}\delta\ell_2\right) \geq 0.$$

Proof. Since $\exp t\vec{H}_i$ transforms horizontal submanifolds into horizontal submanifolds, then (4.8)–(4.9) imply that the map $\pi_*\mathcal{H}_{t*}$ is onto for $t \in [0, T]$ if and only if it is onto for $t = t_i$, $i = 1, 2, \ldots, r$.

Let us now check that conditions 1 and 2 are equivalent to

$$(4.10) \qquad \pi_*\mathcal{H}_{t_i*}L_0 = T_{\hat{\xi}(t_i)}M.$$

If $d\,\tau_i(\ell_0) = 0$ on L_0, then the maps $(\exp t_i\vec{H}_i)_*\phi_{i-1*}$ and $(\exp t_i\vec{H}_{i+1})_*\phi_{i*}$ coincide, and hence (4.10) is equivalent to condition 1, and moreover condition 2 holds with the equality sign.

Otherwise we have that (4.10) holds if and only if

$$\pi_* \phi_{i-1\,*} L_0 = \pi_* \phi_{i*} L_0 = T_{x_0} M$$

and the two half-spaces

$$\left\{ \pi_*(\exp t_i \vec{H}_i)_* \phi_{i-1\,*} \delta\ell, \ \langle d\,\tau_i(\ell_0), \delta\ell \rangle \leq 0 \right\},$$
$$\left\{ \pi_*(\exp t_i \vec{H}_{i+1})_* \phi_{i*} \delta\ell, \ \langle d\,\tau_i(\ell_0), \delta\ell \rangle \geq 0 \right\}$$

do not coincide. To check this we can show that for every $\delta\ell_1,\ \delta\ell_2 \in L_0$ such that

$$\pi_* \left[(\exp t_i \vec{H}_i)_* \phi_{i-1\,*} \delta\ell_1 \right] = \pi_* \left[(\exp t_i \vec{H}_{i+1})_* \phi_{i*} \delta\ell_2 \right]$$

one has

$$\langle d\,\tau_i(\ell_0), \delta\ell_1 \rangle \langle d\,\tau_i(\ell_0), \delta\ell_2 \rangle \geq 0.$$

By (4.3) we obtain that

$$\boldsymbol{\sigma} \left((\exp t_i \vec{H}_i)_* \phi_{i-1\,*} \delta\ell_1, (\exp t_i \vec{H}_{i+1})_* \phi_{i*} \delta\ell_2 \right)$$
$$= \boldsymbol{\sigma} \left((\exp t_i \vec{H}_i)_* \phi_{i-1\,*} \delta\ell_1, (\vec{H}_i - \vec{H}_{i+1})(\ell_i) \right) \langle d\,\tau_i(\ell_0), \delta\ell_2 \rangle.$$

Finally by the strict bang-bang Legendre condition and by (4.2) we obtain that

$$\boldsymbol{\sigma} \left((\exp t_i \vec{H}_i)_* \phi_{i-1\,*} \delta\ell_1, (\vec{H}_i - \vec{H}_{i+1})(\ell_i) \right)$$

has the same sign as $\langle d\,\tau_i(\ell_0), \delta\ell_1 \rangle$, and the statement is proved. □

Remark 4.9. Theorem 4.8 states that in the bang-bang case, a *conjugate point* can occur only at a switching time; moreover condition 1 states that the projection has full dimension, while condition 2 says that there is not a fold.

THEOREM 4.10. *Let the map $\pi_* \mathcal{H}_{t*} : L_0 \to T_{\hat{\xi}(t)} M$ be onto for $t \in [0, T]$. If the form*

$$\delta x \mapsto \boldsymbol{\sigma} \left(d(\vartheta_T \circ \pi_T^{-1})_* \delta x, d(-\beta)_* \delta x \right) = \boldsymbol{\sigma} \left((\mathcal{H}_T \circ \pi_T^{-1})_* \delta x, d(-\beta)_* \delta x \right)$$

is positive definite on $T_{x_T} N_T$, then $\hat{\xi}$ is a strict strong local minimizer for the problem (P).

Proof. By (4.6) if we prove that x_T is a local minimizer on N_T of $\chi = \vartheta_T \circ \pi_T^{-1} + \beta$, then $\hat{\xi}$ is a strong local minimizer for the problem (P). As we pointed out before, $d\chi(x_T) = 0$; thus the second derivative of χ is well defined at x_T, and we have that

$$D^2 \chi(x_T)[\delta x]^2 = \boldsymbol{\sigma} \left(d(\vartheta_T \circ \pi_T^{-1})_* \delta x, d(-\beta)_* \delta x \right),$$

which ends the first part of the proof.

Let us now prove that the minimum is locally uniquely attained. First of all let us notice that under our assumptions $\hat{\xi}(T)$ is a strict local minimum for the function χ. Assume now by contradiction that there exists another admissible trajectory ξ with the same cost and the same final point $\xi(T) = \hat{\xi}(T)$; denote by $\lambda := t \mapsto \pi_t^{-1}(\xi(t))$ its lift. By (4.6) and (4.5) we have that

$$\int_\lambda \boldsymbol{\omega} = 0;$$

thus, since H is the maximized Hamiltonian, we have that

$$\langle \mathcal{H}_t(\lambda(t)), \dot{\xi}(t) \rangle - H(\mathcal{H}_t(\lambda(t))) = 0 \quad \text{a.e. } t \in [0,T].$$

Thanks to Assumption 2.1, for all t such that $(t, \lambda(t)) \in \text{int } \Omega_i$ we have that

$$\dot{\xi}(t) = h_i(\xi(t));$$

this equation yields that

$$\dot{\lambda}(t) = 0,$$

and hence, from inequality (4.1), it follows that $\lambda(t)$ is constant, say k_i, for $t \in [\tau_{i-1}(k_i), \tau_i(k_i)]$. When $\tau_i(k_i) \leq t \leq \tau_i(k_{i+1})$ we have that $(t, \lambda(t)) \in \Sigma_i$, and hence there is $\mu \in [0,1]$ such that

$$\dot{\xi}(t) = (1-\mu)h_{i+1}(\xi(t)) + \mu h_i(\xi(t)).$$

For $t \in [\tau_r(k_r), T]$ we have that $\lambda(t) \in \Omega_{r+1}$, and hence λ is as smooth as ξ; moreover on the last time interval $[t_r, T]$ we have that $\lambda(t) = \ell_0$. Let us now show that λ cannot remain on the switching surface for a time interval of positive measure or, equivalently, that $\tau_r(k_r) = t_r$. By contradiction, if $k_r \neq \ell_0$, then

$$(t, \xi(t)) \in \pi\mathcal{H}(\Sigma_r) \subseteq \pi\mathcal{H}(\Omega_{r+1}), \quad t \in [\tau_r(k_r), t_r].$$

If we differentiate the identities

$$t = \tau_r(\lambda(t)), \quad \pi\mathcal{H}_t(\lambda(t)) = \xi(t),$$

we obtain, a.e. $t \in [\tau_r(k_r), t_r]$,

(4.11) $$\langle d\,\tau_r(\lambda(t)), \dot{\lambda}(t) \rangle = 1, \quad \dot{\xi}(t) = h_{r+1}(\xi(t)) + \pi_*\mathcal{H}_{t*}\dot{\lambda}(t).$$

From Assumption 2.2 it follows that for a.a. $t \in [\tau_r(k_r), t_r]$ there exists $\mu_t \in [0,1]$ such that

$$\dot{\xi}(t) = (1-\mu_t)\,h_{r+1}(\xi(t)) + \mu_t\,h_r(\xi(t)),$$

and hence

$$\pi_*\mathcal{H}_{t*}\dot{\lambda}(t) = \mu_t\Big(h_r(\xi(t)) - h_{r+1}(\xi(t))\Big).$$

Since $\dot{\lambda}(t)$ and μ_t are bounded we can take a sequence $t_i \to t_r$ such that $\dot{\lambda}(t_i) \to \delta\ell \in L_0$ and $\mu_{t_i} \to \mu$. Taking into account (4.11) we can say that $\delta\ell \neq 0$, and since we have that $\pi_*\mathcal{H}_{t_r*}$ is injective, then from

$$\pi_*\mathcal{H}_{t_r*}\delta\ell = \mu\Big(h_r(\xi(t_r)) - h_{r+1}(\xi(t_r))\Big)$$

it follows that also $\mu \neq 0$. On the other hand from (4.4) and (4.3) we obtain

$$\pi_*\left(\exp t_r\vec{H}_r\right)_*\phi_{r-1*}\delta\ell = \pi_*\mathcal{H}_{t_r*}\delta\ell - \Big(h_r(\xi(t_r)) - h_{r+1}(\xi(t_r))\Big)$$

$$= (\mu-1)\Big(h_r(\xi(t_r)) - h_{r+1}(\xi(t_r))\Big).$$

Once again the last term has to be nonzero. If we set $\delta\ell_1 := \frac{\mu}{\mu-1}\delta\ell$, then

$$\langle d\,\tau_r(\ell_0), \delta\ell \rangle \langle d\,\tau_r(\ell_0), \delta\ell_1 \rangle = \frac{\mu}{\mu-1} < 0,$$

and from Theorem 4.8 we obtain a contradiction. Therefore $k_r = \ell_0$, and hence $\tau_r(k_r) = \tau_r(\ell_0) = t_r$. We can do the same proof on each interval proceeding backwards in time to prove that the trajectory is constant. $\quad\square$

4.4. The proof. As we pointed out in Remark 4.6 if we have a regular bang-bang extremal with simple switching times and the strict bang-bang Legendre condition is satisfied, then

$$L_k^- = (\exp t_k \vec{H}_k)_* \varphi_{k-1\,*} L_0 \quad \text{for } k = 1, 2, \ldots, r+1,$$
$$L_k^+ = (\exp t_k \vec{H}_{k+1})_* \varphi_{k\,*} L_0 \quad \text{for } k = 1, 2, \ldots, r.$$

On the other hand we are considering a free initial point problem; therefore if $J''_{|V} > 0$, then from STEP 3 of the algorithm described in Corollary 2.10 it follows that we can apply Theorem 4.8, and hence we have that $\pi_* \mathcal{H}_{t*} L_0 = T_{\hat{\xi}(t)} M$ for all $t \in [0, T]$ and

$$L_{r+1}^- = \left(\mathcal{H}_T \circ \pi_T^{-1}\right)_* T_{x_T} M.$$

Therefore STEP 4 of the algorithm described in Corollary 2.10 yields that we can apply Theorem 4.10.

In the abnormal case the cost is zero, and hence the existence of a *strict* strong local minimizer is equivalent to the fact that the reference trajectory is isolated among the admissible trajectories in the C^0 topology.

REFERENCES

[AG90]   A. A. AGRACHEV AND R. V. GAMKRELIDZE, *Symplectic geometry for optimal control*, in Nonlinear Controllability and Optimal Control, Monogr. Textbooks Pure Appl. Math. 133, H. J. Sussmann, ed., Marcel Dekker, New York, 1990, pp. 263–277.

[AG97]   A. A. AGRACHEV AND R. V. GAMKRELIDZE, *Symplectic methods for optimization and control*, in Geometry of Feedback and Optimal Control, Monogr. Textbooks Pure Appl. Math. 207, B. Jacubzick and W. Respondek, eds., Marcel Dekker, New York, 1998, pp. 19–77.

[Arn80]   V. I. ARNOLD, *Mathematical Methods in Classical Mechanics*, Springer, New York, 1980.

[ASZ98a]   A. A. AGRACHEV, G. STEFANI, AND P. ZEZZA, *An invariant second variation in optimal control*, Internat. J. Control, 71 (1998), pp. 689–715.

[ASZ98b]   A. A. AGRACHEV, G. STEFANI, AND P. ZEZZA, *Strong minima in optimal control*, Proc. Steklov Inst. Math., 220 (1998), pp. 4–26.

[ASZ99]   A. A. AGRACHEV, G. STEFANI, AND P. ZEZZA, *A Hamiltonian approach to strong minima in optimal control*, in Differential Geometry and Control (Boulder, CO, 1997), AMS, Providence, RI, 1999, pp. 11–22.

[GH96a]   M. GIAQUINTA AND S. HILDEBRANDT, *Calculus of Variations* I, Springer, Berlin, 1996.

[GH96b]   M. GIAQUINTA AND S. HILDEBRANDT, *Calculus of Variations* II, Springer, Berlin, 1996.

[PS00]   B. PICCOLI AND H. J. SUSSMANN, *Regular synthesis and sufficient conditions for optimality*, SIAM J. Control Optim., 39 (2000), pp. 359–410.

[Sar92]   A. V. SARYCHEV, *Sufficient optimality conditions for Pontryagin extremals*, Systems Control Lett., 19 (1992), pp. 451–460.

[Sar97]   A. V. SARYCHEV, *First- and second-order sufficient optimality conditions for bang-bang controls*, SIAM J. Control Optim., 35 (1997), pp. 315–340.