



**HAL**  
open science

## Quadratic conditions for bang-singular extremals

Maria Soledad Aronna, J. Frederic Bonnans, Andrei V. Dmitruk, Pablo Lotito

► **To cite this version:**

Maria Soledad Aronna, J. Frederic Bonnans, Andrei V. Dmitruk, Pablo Lotito. Quadratic conditions for bang-singular extremals. [Research Report] RR-7664, 2011, pp.39. inria-00605128v1

**HAL Id: inria-00605128**

**<https://inria.hal.science/inria-00605128v1>**

Submitted on 30 Jun 2011 (v1), last revised 2 Jul 2012 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

*Quadratic order conditions for bang-singular  
extremals*

M. Soledad Aronna — J. Frédéric Bonnans — Andrei V. Dmitruk — Pablo A. Lotito

**N° 7664**

June 2011

Thème NUM

 *Rapport  
de recherche*

ISSN 0249-6399 | ISRN INRIA/RR--7664--FR+ENG



## Quadratic order conditions for bang-singular extremals

M. Soledad Aronna<sup>\*†</sup>, J. Frédéric Bonnans<sup>†</sup>, Andrei V. Dmitruk<sup>‡</sup>  
, Pablo A. Lotito<sup>§</sup>

Thème NUM — Systèmes numériques  
Équipes-Projets Commands

Rapport de recherche n° 7664 — June 2011 — 39 pages

**Abstract:** This paper deals with optimal control problems for systems affine in the control variable. We have nonnegativity constraints on the control, and finitely many equality and inequality constraints on the final state. First, we obtain second order necessary optimality conditions. Secondly, we get a second order sufficient condition for the scalar control case. The results use in an essential way the Goh transformation.

**Key-words:** Optimal control, second order conditions, control constraints, singular arc, bang-singular solutions

The first two authors acknowledge the support of EADS chair.

\* CIFASIS-CONICET Argentina (aronna@cmap.polytechnique.fr)

† INRIA-Saclay and CMAP, École Polytechnique, 91128 Palaiseau, France (Fred-eric.Bonnans@inria.fr)

‡ Russian Academy of Sciences, CEMI and Moscow State University (avdmi@cemi.rssi.ru)

§ PLADEMA CONICET and Universidad Nacional de Centro de la Provincia de Buenos Aires, Argentina (plotito@exa.unicen.edu.ar)

## Conditions quadratiques pour des extrémales bang-singulières

**Résumé :** Dans ce travail nous étudions le problème de commande optimale avec des contrôles affines dans la dynamique. On considère des contraintes de non-négativité sur la commande et une quantité finie de contraintes d'égalité et d'inégalité sur la valeur finale de l'état. Premièrement on obtient des conditions nécessaires d'optimalité de second ordre. Ensuite, on présente une condition suffisante pour le cas d'une commande scalaire. Les résultats s'appuient sur la transformation de Goh.

**Mots-clés :** Commande optimale, conditions du second ordre, contraintes sur la commande, arc singulier, solutions bang-singulières.

## 1 Introduction

In this article we obtain second order conditions for an optimal control problem affine in the control. First we consider pointwise nonnegativity constraint on the control, end-point state constraints and fixed time interval. Then we extend the result to bound constraints on the control, initial-final state constraints and problems involving parameters. We do not assume that the multipliers are unique. We study weak and Pontryagin minima.

There is already an important literature on this subject. The case without control constraints, i.e., when the extremal is totally singular, has been extensively studied since middle 60's. Kelley in [23] treated the scalar control case and presented a necessary condition involving the second order derivative of the switching function. The result was extended by Kopp and Moyer [25] for higher order derivatives, and in [24] it was shown that the order had to be even. Goh in [18] proposed a special change of variables obtained via a linear ODE and in [17] used this transformation to show a necessary condition for the vector control problem. An extensive survey of these articles can be found in Gabasov and Kirillova [16]. Jacobson and Speyer in [21], and together with Lele in [22] obtained necessary conditions by penalizing the cost functional. Gabasov and Kirillova [16], Krener [26], Agrachev and Gamkrelidze [1] obtained a countable series of necessary conditions that in fact exploit the idea of Goh transformation. Milyutin in [31] discovered an abstract essence of this approach and obtained even stronger necessary conditions. The main specificity of this kind of problems, where the control enters linearly, is that the corresponding second variation does not contain the Legendre term, so the methods of classical calculus of variations are not applicable for obtaining sufficient conditions. This is why the literature was mostly devoted to necessary conditions, that actually are consequence of the nonnegativity of the second variation. Goh's transformation mentioned above allows one to convert the second variation into another functional that hopefully turns out to be coercive with respect to the  $L_2$ -norm of some state variable. Dmitruk in [5] proved that this coercivity is a sufficient condition for the weak optimality, and presented a closely related necessary condition. He used the abstract approach developed by Levitin, Milyutin and Osmolovskii in [28]. In [6, 8] he also obtained necessary and sufficient conditions for this norm, again closely related, for Pontryagin minimality.

On the other hand, the case with linear control constraints and a "purely" bang-bang control without singular subarcs, was deeply investigated in the past 15 years. Milyutin and Osmolovskii in [32] provided necessary and sufficient conditions based on the general theory of [28]. Osmolovskii in [33] completed some of the proofs of the latter article. Agrachev, Stefani, Zezza [2] reduced the problem to a finite dimensional one having the switching instants as variables and obtained a sufficient condition for strong optimality. The result was recently extended by Poggiolini and Spadini in [34]. On the other hand, Maurer and Osmolovskii in [30, 29] gave a second order sufficient condition that is suitable for practical verifications and presented a numerical procedure that allows to verify the positivity of certain quadratic forms. Felgenhauer in [13, 14, 15] studied both second order optimality conditions and sensitivity of the optimal solution.

The mixed case, where the control is partly bang-bang, partly singular, had not been decoded up to recent time. Poggiolini and Stefani obtained in [35]

a second order sufficient condition with an additional geometrical hypothesis , that we do not need here, and in [36, 37] a second order sufficient condition for the special case of a time-optimal problem. Our approach is quite different.

The article is organised as follows: in the second section we present the problem and give basic definitions. In the third section we make the second order analysis: we obtain the second variation of the Lagrangian functions and a necessary condition. Afterwards we present the Goh transformation and a new necessary condition in the transformed variables. In the fifth section we show a sufficient condition for scalar control. Finally, we give an example with a scalar control where the second order sufficient condition can be verified. The appendix is devoted to a series of technical properties that are used to prove the main results.

## 2 Statement of the problem and assumptions

### 2.1 Statement of the problem

Consider the spaces  $\mathcal{U} := L_\infty(0, T; \mathbb{R}^m)$  and  $\mathcal{X} := W_\infty^1(0, T; \mathbb{R}^n)$ , as control and state spaces respectively. Denote with  $u$  and  $x$  their elements, respectively. When needed, put  $w = (x, u)$  for a point in  $\mathcal{W} := \mathcal{X} \times \mathcal{U}$ . In this paper we investigate the optimal control problem

$$J := \varphi_0(x(T)) \rightarrow \min, \quad (1)$$

$$\dot{x}(t) = \sum_{i=0}^m u_i f_i(x), \quad x(0) = x_0, \quad (2)$$

$$u(t) \geq 0, \text{ for a.a. } t \in (0, T), \quad (3)$$

$$\varphi_i(x(T)) \leq 0, \text{ for } i = 1, \dots, d_\varphi, \quad \eta_j(x(T)) = 0, \text{ for } j = 1, \dots, d_\eta. \quad (4)$$

where  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$  for  $i = 0, \dots, m$ ,  $\varphi_i : \mathbb{R}^n \rightarrow \mathbb{R}$  for  $i = 0, \dots, d_\varphi$ ,  $\eta_j : \mathbb{R}^n \rightarrow \mathbb{R}$  for  $j = 1, \dots, d_\eta$  and  $u_0 \equiv 1$ . Assume that data functions  $f_i$  are twice continuously differentiable. Functions  $\varphi_i$  and  $\eta_j$  are asked to be twice differentiable.

A *trajectory* is an element  $w \in \mathcal{W}$  that satisfies the state equation (2). If in addition constraints (3) and (4) hold, say that  $w$  is a *feasible point* of the problem (1)-(4). Denote by  $\mathcal{A}$  the *set of feasible points*. Call *feasible variation* for  $\hat{w} \in \mathcal{A}$  an element  $\delta w \in \mathcal{W}$  such that  $\hat{w} + \delta w \in \mathcal{A}$ .

*Definition 2.1.* A pair  $w^0 = (x^0, u^0) \in \mathcal{W}$  is said to be a *weak minimum* of problem (1)-(4) if there exists  $\varepsilon > 0$  such that the cost function attains at  $w^0$  its minimum on the set:

$$\{w = (x, u) \in \mathcal{W} : \|x - x^0\|_\infty < \varepsilon, \|u - u^0\|_\infty < \varepsilon\}.$$

And  $w^0$  is a *Pontryagin minimum* of problem (1)-(4) if for any positive  $N$  there exists  $\varepsilon_N > 0$  such that  $w^0$  is a minimum point on the set:

$$\{w = (x, u) \in \mathcal{W} : \|x - x^0\|_\infty < \varepsilon_N, \|u - u^0\|_\infty \leq N, \|u - u^0\|_1 < \varepsilon_N\}.$$

Consider  $\lambda = (\alpha, \beta, \psi) \in \mathbb{R}^{d_\varphi+1,*} \times \mathbb{R}^{d_\eta,*} \times W_\infty^1(0, T; \mathbb{R}^{n,*})$ , i.e.,  $\psi$  is a Lip-continuous function with values in the  $n$ -dimensional space of row-vectors

with real components  $\mathbb{R}^{n,*}$ . Define the *pre-Hamiltonian* function

$$H[\lambda](x, u, t) := \psi(t) \sum_{i=0}^m u_i f_i(x),$$

the *terminal Lagrangian* function

$$\ell[\lambda](q) := \sum_{i=0}^{d_\varphi} \alpha_i \varphi_i(q) + \sum_{j=1}^{d_n} \beta_j \eta_j(q),$$

and the *Lagrangian* function

$$\Phi[\lambda](w) := \ell[\lambda](x(T)) + \int_0^T \psi(t) \left( \sum_{i=0}^m u_i(t) f_i(x(t)) - \dot{x}(t) \right) dt. \quad (5)$$

In this article the optimality of a given feasible trajectory  $\hat{w} = (\hat{x}, \hat{u})$  is studied. Whenever some argument of  $f_i$ ,  $H$ ,  $\ell$ ,  $\Phi$  or their derivatives is omitted, assume that they are evaluated over this trajectory. Without loss of generality suppose that

$$\varphi_i(\hat{x}(T)) = 0, \text{ for all } i = 0, 1, \dots, d_\varphi. \quad (6)$$

## 2.2 First order analysis

*Definition 2.2.* Denote by  $\Lambda \subset \mathbb{R}^{d_\varphi+1,*} \times \mathbb{R}^{d_n,*} \times W_\infty^1(0, T; \mathbb{R}^{n,*})$  the set of *Pontryagin multipliers* associated to  $\hat{w}$  consisting of the elements  $\lambda = (\alpha, \beta, \psi)$  satisfying the *Pontryagin Maximum Principle*, i.e., having the following properties:

$$|\alpha| + |\beta| = 1, \quad (7)$$

$$\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{d_\varphi}) \geq 0, \quad (8)$$

function  $\psi$  is solution of the *costate equation* and satisfies the *transversality condition* in the end-point  $T$ , i.e.,

$$-\dot{\psi}(t) = H_x[\lambda](\hat{x}(t), \hat{u}(t), t), \quad \psi(T) = \ell'[\lambda](\hat{x}(T)), \quad (9)$$

and the following *minimum condition* holds:

$$H[\lambda](\hat{x}(t), \hat{u}(t), t) = \min_{v \geq 0} H[\lambda](\hat{x}(t), v, t), \text{ a.e. on } [0, T]. \quad (10)$$

*Remark 2.3.* For every  $\lambda \in \Lambda$  holds:

- (i)  $H_{u_i}[\lambda]$  is continuous in time,
- (ii)  $H_{u_i}[\lambda](t) \geq 0$ , a.e. on  $(0, T)$ .

Recall the following well known result whose proof can be found e.g. in Alekseev and Tikhomirov [3], Kurcyusz and Zowe [27]:

**Theorem 2.4.** *The set  $\Lambda$  is not empty.*



*Remark 2.5.* Since  $\psi$  may be expressed as a linear continuous mapping of  $(\alpha, \beta)$  and since (7) holds,  $\Lambda$  is a finite-dimensional compact set, thus it can be identified with a compact subset of  $\mathbb{R}^s$ , where  $s := d_\varphi + d_\eta + 1$ .

The following expression for the derivative of the Lagrangian function holds:

$$\Phi_u[\lambda](\hat{w})v = \int_0^T H_u[\lambda](\hat{x}(t), \hat{u}(t), t)v(t)dt. \quad (11)$$

Consider  $v \in \mathcal{U}$  and the *linearized state equation*:

$$\begin{cases} \dot{z}(t) = \sum_{i=0}^m \hat{u}_i(t)f'_i(\hat{x}(t))z(t) + \sum_{i=1}^m v_i(t)f_i(\hat{u}(t)), & \text{a.e. on } [0, T], \\ z(0) = 0. \end{cases} \quad (12)$$

Its solution  $z$  is called *linearized state variable*.

To each index  $i = 1, \dots, m$  we associate the sets:

$$I_0^i := \left\{ t \in [0, T] : \max_{\lambda \in \Lambda} H_{u_i}[\lambda](t) > 0 \right\}, \quad I_+^i := [0, T] \setminus I_0^i, \quad (13)$$

and the *active set*:

$$\tilde{I}_0^i := \{t \in [0, T] : \hat{u}_i(t) = 0\}. \quad (14)$$

Notice that  $I_0^i \subset \tilde{I}_0^i$ , and that  $I_0^i$  is relatively open in  $[0, T]$  as each  $H_{u_i}[\lambda]$  is continuous.

*Assumption 2.6.* Assume *strict complementarity for control constraint*, i.e., for every  $i = 1, \dots, m$  :

$$I_0^i = \tilde{I}_0^i, \text{ up to a set of null measure.} \quad (15)$$

Observe then that for any index  $i = 1, \dots, m$ , control  $\hat{u}_i(t) > 0$  a.e. on  $I_+^i$ , and given  $\lambda \in \Lambda$ ,

$$H_{u_i}[\lambda](t) = 0, \text{ a.e. on } I_+^i. \quad (16)$$

*Assumption 2.7.* For every  $i = 1, \dots, m$  the active set  $I_0^i$  is a finite union of intervals, i.e.,

$$I_0^i = \bigcup_{j=1}^{N_i} I_j^i, \quad (17)$$

for  $I_j^i$  subintervals of  $[0, T]$  of the form  $[0, d)$ ,  $(c, T]$ ; or  $(c, d)$  if  $c \neq 0$  and  $d \neq T$ . Call  $c_1^i < d_1^i < c_2^i < \dots < c_{N_i}^i < d_{N_i}^i$  the endpoints of these intervals. Consequently,  $I_+^i$  is a finite union of intervals as well.

### 2.3 Critical cones

Let  $1 \leq p \leq \infty$ , and call  $\mathcal{U}_p := L_p(0, T; \mathbb{R}^m)$ ,  $\mathcal{U}_p^+ := L_p(0, T; \mathbb{R}_+^m)$  and  $\mathcal{X}_p := W_p^1(0, T; \mathbb{R}^m)$ . Recall that given a topological vector space  $E$ , a subset  $D \subset E$  and  $x \in E$ , a *tangent direction* to  $D$  at  $x$  is an element  $d \in E$  such that there exist sequences  $(\sigma_k) \subset \mathbb{R}_+$  and  $(x_k) \subset D$  such that

$$\frac{x_k - x}{\sigma_k} \rightarrow d. \quad (18)$$

It is a well known result, see e.g. [4], that the tangent cone to  $\mathcal{U}_2^+$  at  $\hat{u}$  is:

$$\{v \in \mathcal{U}_2 : v_i \geq 0 \text{ on } I_0^i, \text{ for } i = 1, \dots, m\}. \quad (19)$$

Given  $v \in \mathcal{U}_p$  and  $z$  solution of (12), consider the *linearization of the final constraints*:

$$\begin{cases} \varphi'_i(\hat{x}(T))z(T) \leq 0, & i = 0, \dots, d_\varphi, \\ \eta'_j(\hat{x}(T))z(T) = 0, & j = 1, \dots, d_\eta. \end{cases} \quad (20)$$

For  $p \in \{2, \infty\}$ , define the  $L_p$ -critical cone as

$$\mathcal{C}_p := \{(z, v) \in \mathcal{X}_p \times \mathcal{U}_p : v \text{ tangent to } \mathcal{U}_p^+, (12) \text{ and } (20) \text{ hold}\}. \quad (21)$$

Certain relations of inclusion and density between some approximate critical cones are needed. Given  $\varepsilon \geq 0$  and  $i = 1, \dots, m$ , define the  $\varepsilon$ -active sets, up to a set of null measure:

$$I_\varepsilon^i := \{t \in (0, T) : \hat{u}_i(t) \leq \varepsilon\}, \quad (22)$$

and the sets:

$$\mathcal{W}_{p,\varepsilon} := \{(z, v) \in \mathcal{X}_p \times \mathcal{U}_p : v_i = 0 \text{ on } I_\varepsilon^i, (12) \text{ holds}\}. \quad (23)$$

By Assumption 2.6, the following explicit expression of  $\mathcal{C}_2$  holds:

$$\mathcal{C}_2 = \{(z, v) \in \mathcal{W}_{2,0} : (20) \text{ holds}\}. \quad (24)$$

Consider the  $\varepsilon$ -critical cones:

$$\mathcal{C}_{p,\varepsilon} := \{(z, v) \in \mathcal{W}_{p,\varepsilon} : (20) \text{ holds}\}. \quad (25)$$

Let  $\varepsilon > 0$ . Note that by (24),  $\mathcal{C}_{2,\varepsilon} \subset \mathcal{C}_2$ . On the other hand, given  $(z, v) \in \mathcal{C}_{\infty,\varepsilon}$ , easily follows that  $\hat{u} + \sigma v \in \mathcal{U}^+$  for small positive  $\sigma$ . Thus  $v$  is tangent to  $\mathcal{U}^+$  at  $\hat{u}$ , and this yields  $\mathcal{C}_{\infty,\varepsilon} \subset \mathcal{C}_\infty$ .

Recall the following technical result, Dmitruk [9].

**Lemma 2.8** (on density). *Consider a locally convex topological space  $X$ , a finite-faced cone  $C \subset X$ , and a linear manifold  $L$  dense in  $X$ . Then the cone  $C \cap L$  is dense in  $C$ .*

**Lemma 2.9.** *Given  $\varepsilon > 0$  the following properties hold:*

- (a)  $\mathcal{C}_{\infty,\varepsilon} \subset \mathcal{C}_{2,\varepsilon}$  with dense inclusion,
- (b)  $\bigcup_{\varepsilon>0} \mathcal{C}_{2,\varepsilon} \subset \mathcal{C}_2$  with dense inclusion.

*Proof.* (a) The inclusion is immediate. As  $\mathcal{U}$  is dense in  $\mathcal{U}_2$ ,  $\mathcal{W}_{\infty,\varepsilon}$  is a dense subspace of  $\mathcal{W}_{2,\varepsilon}$ . By Lemma 2.8,  $\mathcal{C}_{2,\varepsilon} \cap \mathcal{W}_{\infty,\varepsilon}$  is dense in  $\mathcal{C}_{2,\varepsilon}$ , as it was to be proved.

(b) The inclusion is immediate. In order to prove density, consider the following dense subspace of  $\mathcal{W}_{2,0}$ :

$$\mathcal{W}_{2,\bigcup} := \bigcup_{\varepsilon>0} \mathcal{W}_{2,\varepsilon},$$

and the finite-faced cone in  $\mathcal{C}_2 \subset \mathcal{W}_{2,0}$ . By Lemma 2.8,  $\mathcal{C}_2 \cap \mathcal{W}_{2,\bigcup}$  is dense in  $\mathcal{C}_2$ , as it was to be proved.  $\square$

### 3 Second order analysis

#### 3.1 Second variation

Consider the following quadratic mapping on  $\mathcal{W}$ ;

$$\begin{aligned} \Omega[\lambda](\delta x, \delta u) &:= \frac{1}{2} \ell''[\lambda](\hat{x}(T))(\delta x(T))^2 \\ &+ \frac{1}{2} \int_0^T [(H_{xx}[\lambda]\delta x, \delta x) + 2(H_{ux}[\lambda]\delta x, \delta u)] dt. \end{aligned} \quad (26)$$

The next lemma provides a second order expansion for the Lagrangian function involving operator  $\Omega$ . Recall the following notation: given two functions  $h : \mathbb{R}^n \rightarrow \mathbb{R}^{n_h}$  and  $k : \mathbb{R}^n \rightarrow \mathbb{R}^{n_k}$ , say that  $h$  is a *big-O* of  $k$  around 0 and denote it by

$$h(x) = \mathcal{O}(k(x)),$$

if there exists positive  $\delta$  and  $M$  such that  $|h(x)| \leq M|k(x)|$  for  $|x| < \delta$ . It is a *small-o* if  $M$  goes to 0 as  $|x|$  goes to 0. Denote this by

$$h(x) = o(k(x)).$$

**Lemma 3.1.** *Let  $\delta w = (\delta x, \delta u) \in \mathcal{W}$ . Then for every multiplier  $\lambda \in \Lambda$ , the function  $\Phi$  has the following expansion (skipping time arguments):*

$$\begin{aligned} \Phi[\lambda](\hat{w} + \delta w) &= \int_0^T H_u[\lambda]\delta u dt + \Omega[\lambda](\delta x, \delta u) + \frac{1}{2} \int_0^T (H_{uxx}[\lambda]\delta x, \delta x, \delta u) dt \\ &+ \mathcal{O}(|\delta x(T)|^3) + \mathcal{O}(\|\delta x\|_\infty^3) \int_0^T (1 + |(\hat{u} + \delta u)(t)|) dt. \end{aligned}$$

*Proof.* Omit the dependence on  $\lambda$  for the sake of simplicity. Use the Taylor expansions:

$$\ell(\hat{x}(T) + \delta x(T)) = \ell(\hat{x}(T)) + \ell'(\hat{x}(T))\delta x(T) + \frac{1}{2} \ell''(\hat{x}(T))(\delta x(T))^2 + \mathcal{O}(|\delta x(T)|^3),$$

$$f_i(\hat{x}(t) + \delta x(t)) = f_i(\hat{x}(t)) + f'_i(\hat{x}(t))\delta x(t) + \frac{1}{2} f''_i(\hat{x}(t))(\delta x(t))^2 + \mathcal{O}(\|\delta x\|_\infty^3),$$

in the expression

$$\Phi(\hat{w} + \delta w) = \ell(\hat{x} + \delta x(T)) + \int_0^T \psi \left[ \sum_{i=0}^m (\hat{u}_i + \delta u_i) f_i(\hat{x} + \delta x) - \dot{\hat{x}} - \dot{\delta x} \right] dt.$$

Afterwards, use the identity

$$\int_0^T \psi \sum_{i=1}^m \hat{u}_i f'_i(\hat{x}) \delta x dt = -\ell'(\hat{x}(T))\delta x(T) + \int_0^T \psi \dot{\delta x} dt, \quad (27)$$

obtained by integration by parts and equation (2) to get the desired result.  $\square$

The previous lemma yields the following identity for every  $(\delta x, \delta u) \in \mathcal{W}$ :

$$\Omega[\lambda](\delta x, \delta u) = \frac{1}{2} D^2 \Phi[\lambda](\hat{w})(\delta x, \delta u)^2. \quad (28)$$

### 3.2 Necessary condition

This section provides the following result:

**Theorem 3.2.** *If  $\hat{w}$  is a weak minimum then*

$$\max_{\lambda \in \Lambda} \Omega[\lambda](z, v) \geq 0, \quad \text{for all } (z, v) \in \mathcal{C}_2. \quad (29)$$

For the sake of simplicity, define  $\bar{\varphi} : \mathcal{U} \rightarrow \mathbb{R}^{d_\varphi+1}$ , and  $\bar{\eta} : \mathcal{U} \rightarrow \mathbb{R}^{d_\eta}$  as

$$\begin{aligned} \bar{\varphi}_i(u) &:= \varphi_i(x(T)), \quad \text{for } i = 0, 1, \dots, d_\varphi, \\ \bar{\eta}_j(u) &:= \eta_j(x(T)), \quad \text{for } j = 1, \dots, d_\eta, \end{aligned} \quad (30)$$

where  $x$  is the solution of (2) corresponding to  $u$ .

*Definition 3.3.* We say that the equality constraints are nondegenerate if

$$\bar{\eta}'(\hat{u}) \text{ is onto from } \mathcal{U} \text{ to } \mathbb{R}^{d_\eta}. \quad (31)$$

Call them *degenerate* if (31) does not hold.

Write the problem in the following way

$$\bar{\varphi}_0(u) \rightarrow \min; \quad \bar{\varphi}_i(u) \leq 0, \quad i = 1, \dots, d_\varphi, \quad \bar{\eta}(u) = 0, \quad u \in \mathcal{U}_+. \quad (\text{P})$$

**Suppose that  $\hat{u}$  is a local weak solution of (P).** Next we prove Theorem 3.2. Its proof is divided in two cases: degenerate and nondegenerate equality constraints. For the first case the result is immediate and is tackled in the next Lemma. In order to show Theorem 3.2 for the latter case we introduce an auxiliary problem parameterized by certain critical directions  $(z, v)$ , denoted by  $(QP_v)$ . We prove that  $\text{val}(QP_v) \geq 0$  and by a result on duality, the desired second order condition will be derived.

**Lemma 3.4.** *If equality constraints are degenerate, then (29) holds.*

*Proof.* Notice that there exists  $\beta \neq 0$  such that  $\sum_{j=1}^{d_\eta} \beta_j \eta'_j(\hat{x}(T)) = 0$ , since  $\bar{\eta}'(\hat{u})$  is not onto. Consider  $\alpha = 0$  and  $\psi = 0$ . Take  $\lambda := (\alpha, \beta, \psi)$  and notice that both  $\lambda$  and  $-\lambda$  are in  $\Lambda$ . As

$$\Omega[\lambda](z, v) = \frac{1}{2} \sum_{j=1}^{d_\eta} \beta_j \eta''_j(\hat{x}(T))(z(T))^2, \quad (32)$$

thus  $\Omega[\lambda](z, v) \geq 0$  either for  $\lambda$  or  $-\lambda$ . The wanted result follows.  $\square$

Take  $\varepsilon > 0$ ,  $(z, v) \in \mathcal{C}_{\infty, \varepsilon}$ , and rewrite (25) using the notation in (30):

$$\begin{aligned} \mathcal{C}_{\infty, \varepsilon} &= \{(z, v) \in \mathcal{X} \times \mathcal{U} : v_i(t) = 0 \text{ on } I_\varepsilon^i, \quad i = 1, \dots, m, \\ &\quad (12) \text{ holds, } \bar{\varphi}'_i(\hat{u})v \leq 0, \quad i = 0, \dots, d_\varphi, \quad \bar{\eta}'(\hat{u})v = 0\}. \end{aligned} \quad (33)$$

Consider the problem:

$$\begin{aligned} \delta\zeta &\rightarrow \min \\ \bar{\varphi}'_i(\hat{u})r + \bar{\varphi}''_i(\hat{u})(v, v) &\leq \delta\zeta, \quad \text{for } i = 0, \dots, d_\varphi, \\ \bar{\eta}'(\hat{u})r + \bar{\eta}''(\hat{u})(v, v) &= 0, \\ -r_i(t) &\leq \delta\zeta, \quad \text{on } I_0^i, \quad \text{for } i = 1, \dots, m. \end{aligned} \quad (\text{QP}_v)$$

**Proposition 3.5.** *Let  $(z, v) \in \mathcal{C}_{\infty, \varepsilon}$ . If the equality constraints are nondegenerate, problem  $(QP_v)$  is feasible and  $\text{val}(QP_v) \geq 0$ .*

*Proof.* Let us first prove feasibility. As  $\bar{\eta}'(\hat{u})$  is onto, there exists  $r \in \mathcal{U}$  such that the equality constraint in  $(QP_v)$  is satisfied. Take

$$\delta\zeta := \max(\|r\|_{\infty}, \bar{\varphi}'_i(\hat{u})r + \bar{\varphi}''(\hat{u})(v, v)). \quad (34)$$

Thus the pair  $(r, \delta\zeta)$  is feasible for  $(QP_v)$ .

Let us now prove that  $\text{val}(QP_v) \geq 0$ . On the contrary suppose there exists a feasible solution  $(r, \delta\zeta)$  with  $\delta\zeta < 0$ . The last constraint in  $(QP_v)$  implies  $\|r\|_{\infty} \neq 0$ . Set, for  $\sigma > 0$ ,

$$\tilde{u}(\sigma) := \hat{u} + \sigma v + \frac{1}{2}\sigma^2 r, \quad \tilde{\zeta}(\sigma) := \frac{1}{2}\sigma^2 \delta\zeta. \quad (35)$$

The goal is finding  $u(\sigma)$  feasible for (P) such that for small  $\sigma$ ,

$$u(\sigma) \xrightarrow{\mathcal{U}} \hat{u}, \text{ and } \bar{\varphi}_0(u(\sigma)) < \bar{\varphi}_0(\hat{u}), \quad (36)$$

contradicting the weak optimality of  $\hat{u}$ .

Notice that  $\hat{u}_i(t) > \varepsilon$  a.e. on  $[0, T] \setminus I_{\varepsilon}^i$ , and then  $\tilde{u}(\sigma)_i(t) > -\tilde{\zeta}(\sigma)$  for sufficiently small  $\sigma$ . On  $I_{\varepsilon}^i$ , if  $\tilde{u}(\sigma)_i(t) < -\tilde{\zeta}(\sigma)$  then necessarily  $\hat{u}_i(t) < \frac{1}{2}\sigma^2(\|r\|_{\infty} + |\delta\zeta|)$ , as  $v_i(t) = 0$ . Thus, defining the set

$$J_{\sigma}^i := \{t : 0 < \hat{u}_i(t) < \frac{1}{2}\sigma^2(\|r\|_{\infty} + |\delta\zeta|)\}, \quad (37)$$

we get  $\{t \in [0, T] : \tilde{u}(\sigma)_i(t) < -\tilde{\zeta}(\sigma)\} \subset J_{\sigma}^i$ . Observe that on  $J_{\sigma}^i$ , the function  $|\tilde{u}(\sigma)_i(t) + \tilde{\zeta}(\sigma)|/\sigma^2$  is dominated by  $\|r\|_{\infty} + |\delta\zeta|$ . Since  $\text{meas}(J_{\sigma}^i)$  goes to 0 by the Dominated Convergence Theorem, we obtain

$$\int_{J_{\sigma}^i} |\tilde{u}(\sigma)_i(t) + \tilde{\zeta}(\sigma)| dt = o(\sigma^2). \quad (38)$$

Take

$$\tilde{\tilde{u}}(\sigma) := \begin{cases} \tilde{u}(\sigma) & \text{on } [0, T] \setminus J_{\sigma}^i, \\ -\tilde{\zeta}(\sigma) & \text{on } J_{\sigma}^i. \end{cases} \quad (39)$$

Thus,  $\tilde{\tilde{u}}$  satisfies

$$\tilde{\tilde{u}}(\sigma)(t) \geq -\tilde{\zeta}(\sigma), \quad \text{a.e. on } [0, T], \quad (40)$$

$$\|\tilde{\tilde{u}}(\sigma) - \hat{u}\|_1 = o(\sigma^2), \quad \|\tilde{\tilde{u}}(\sigma) - \hat{u}\|_{\infty} = O(\sigma^2), \quad (41)$$

and the following estimates hold

$$\begin{aligned} \bar{\varphi}_i(\tilde{\tilde{u}}(\sigma)) &= \bar{\varphi}_i(\hat{u}) + \sigma \bar{\varphi}'_i(\hat{u})v + \frac{1}{2}\sigma^2 [\bar{\varphi}'_i(\hat{u})r + \bar{\varphi}''_i(\hat{u})(v, v)] + o(\sigma^2) \\ &< \bar{\varphi}_i(\hat{u}) + \tilde{\zeta}(\sigma) + o(\sigma^2), \end{aligned} \quad (42)$$

$$\bar{\eta}(\tilde{\tilde{u}}(\sigma)) = \sigma \bar{\eta}'(\hat{u})v + \frac{1}{2}\sigma^2 [\bar{\eta}'(\hat{u})r + \bar{\eta}''(\hat{u})(v, v)] + o(\sigma^2) = o(\sigma^2). \quad (43)$$

As  $\bar{\eta}'(\hat{u})$  is onto on  $\mathcal{U}$  we can find a corrected control  $u(\sigma)$  satisfying the equality constraint and such that  $\|u(\sigma) - \tilde{\tilde{u}}(\sigma)\|_{\infty} = o(\sigma^2)$ . Deduce by (40) that  $u(\sigma) \geq 0$  a.e. on  $[0, T]$ , and by (42) that it satisfies the terminal inequality constraints. Thus  $u(\sigma)$  is feasible for (P) and it satisfies (35). This contradicts the weak optimality of  $\hat{u}$ .  $\square$

Recall that a *Lagrange multiplier* associated with  $\hat{w}$  is a pair  $(\lambda, \mu) \in \mathbb{R}^{d_\varphi+1} \times \mathbb{R}^{d_\eta} \times W_\infty^1(0, T; \mathbb{R}^{n,*}) \times \mathcal{U}^*$  with  $\lambda = (\alpha, \beta, \psi)$  satisfying (7), (8),  $\mu \geq 0$  and the *stationarity* condition:

$$\int_0^T H_u[\lambda](t)v(t)dt + \int_0^T v(t)d\mu(t) = 0, \quad \text{for every } v \in \mathcal{U}. \quad (44)$$

Here  $\mathcal{U}^*$  denotes the dual space of  $\mathcal{U}$ . Simple computations show that  $(\lambda, \mu)$  is a Lagrange multiplier if and only if  $\lambda$  is a Pontryagin multiplier and  $\mu = H_u[\lambda]$ . Thus  $\mu \in L_\infty(0, T; \mathbb{R}^{m,*})$ .

Let us come back to Theorem 3.2:

*Proof.* [of Theorem 3.2] Lemma 3.4 covers the degenerate case. Assume thus that  $\bar{\eta}'(\hat{u})$  is onto. Take  $\varepsilon > 0$  and  $(z, v) \in \mathcal{C}_{\infty, \varepsilon}$ . Applying Proposition 3.5 we see that there can not exist  $r$  and  $\delta\zeta < 0$  such that

$$\begin{aligned} \bar{\varphi}'_i(\hat{u})r + \bar{\varphi}''_i(\hat{u})(v, v) &\leq \delta\zeta, \quad i = 0, \dots, d_\varphi, \\ \bar{\eta}'(\hat{u})r + \bar{\eta}''(\hat{u})(v, v) &= 0, \\ -r_i(t) &\leq \delta\zeta, \quad \text{on } I_0^i, \quad \text{for } i = 1, \dots, m. \end{aligned} \quad (45)$$

By Dubovitskii-Milyutin Theorem (see [11]) we obtain the existence of  $(\alpha, \beta) \in \mathbb{R}^s$  and  $\mu \in \mathcal{U}^*$  with  $\text{supp } \mu_i \subset I_0^i$ , and  $(\alpha, \beta, \mu) \neq 0$  such that

$$\sum_{i=0}^{d_\varphi} \alpha_i \bar{\varphi}'_i(\hat{u}) + \sum_{i=1}^{d_\eta} \beta_j \bar{\eta}'_j(\hat{u}) - \mu = 0, \quad (46)$$

and denoting  $\lambda := (\alpha, \beta, \psi)$ , with  $\psi$  solution of (9), the following holds:

$$\sum_{i=0}^{d_\varphi} \alpha_i \bar{\varphi}''_i(\hat{u})(v, v) + \sum_{i=1}^{d_\eta} \beta_j \bar{\eta}''_j(\hat{u})(v, v) \geq 0. \quad (47)$$

By Lemma 7.2 we obtain

$$\Omega[\lambda](z, v) \geq 0. \quad (48)$$

Observe that (46) implies that  $\lambda \in \Lambda$ . Consider now  $(\bar{z}, \bar{v}) \in \mathcal{C}_2$ , and note that Lemma 2.9 guarantees the existence of a sequence  $\{(z_\varepsilon, v_\varepsilon)\} \subset \mathcal{C}_{\infty, \varepsilon}$  converging to  $(\bar{z}, \bar{v})$  in  $\mathcal{X}_2 \times \mathcal{U}_2$ . Recall Remark 2.5. Let  $\lambda_\varepsilon \in \Lambda$  be such that (48) holds for  $(\lambda_\varepsilon, z_\varepsilon, v_\varepsilon)$ . Since  $(\lambda_\varepsilon)$  is bounded, it contains a limit point  $\bar{\lambda} \in \Lambda$ . Thus (48) holds for  $(\bar{\lambda}, \bar{z}, \bar{v})$ , as it was to be proved.  $\square$

## 4 Goh Transformation

Consider an arbitrary linear system:

$$\begin{cases} \dot{z}(t) = A(t)z(t) + B(t)v(t), & t \in [0, T], \\ z(0) = 0. \end{cases} \quad (49)$$

where  $A(t) \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$  is an essentially bounded function of  $t$ , and  $B(t) \in \mathcal{L}(\mathbb{R}^m; \mathbb{R}^n)$  is a Lipschitz-continuous function of  $t$ . To each  $v \in \mathcal{U}$  associate the state variable  $z \in \mathcal{X}$  solution of (12). Let us present a transformation of

the variables  $(z, v) \in \mathcal{W}$ , first introduced by Goh in [18]. Define two new state variables as follows:

$$\begin{cases} y(t) := \int_0^t v(s) ds, \\ \xi(t) := z(t) - B(t)y(t). \end{cases} \quad (50)$$

Thus  $y \in \mathcal{Y} := W_\infty^1(0, T; \mathbb{R}^m)$ ,  $y(0) = 0$  and  $\xi$  is an element of space  $\mathcal{X}$ . Easily follows that  $\xi$  is solution of the linear differential equation

$$\dot{\xi}(t) = A(t)\xi(t) + B_1(t)y(t), \quad \xi(0) = 0, \quad (51)$$

where

$$B_1(t) := A(t)B(t) - \dot{B}(t). \quad (52)$$

For the purposes of this article take

$$A(t) := \sum_{i=0}^m \hat{u}_i f'_i(\hat{x}(t)), \quad \text{and} \quad B(t)v(t) := \sum_{i=1}^m v_i(t) f_i(\hat{u}(t)), \quad (53)$$

and then (49) coincides with linearized equation (12).

#### 4.1 Transformed critical directions

As optimality conditions on the variables obtained by the Goh Transformation will be derived, a new set of critical directions is needed. Take a point  $(z, v)$  in  $\mathcal{C}_\infty$ , and define  $\xi$  and  $y$  by transformation (50). Let  $h := y(T)$  and notice that since (20) is satisfied, the following inequalities hold:

$$\begin{aligned} \varphi'_i(\hat{x}(T))(\xi(T) + B(T)h) &\leq 0, \quad \text{for } i = 0, \dots, d_\varphi, \\ \eta'_j(\hat{x}(T))(\xi(T) + B(T)h) &= 0, \quad \text{for } j = 1, \dots, d_\eta. \end{aligned} \quad (54)$$

Define the set of *transformed critical directions* as follows:

$$\mathcal{P} := \left\{ (\xi, y, h) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m : \dot{y}_i = 0 \text{ over } I_0^i, y(0) = 0, h := y(T), \right. \\ \left. (51) \text{ and } (54) \text{ hold} \right\}. \quad (55)$$

Observe that for every  $(\xi, y, h) \in \mathcal{P}$  and  $1 \leq i \leq m$ :

$$y_i \text{ is constant over each connected component of } I_0^i, \quad (56)$$

and on the endpoints it holds:

$$\begin{aligned} y_i &= 0 \text{ on } [0, d_1^i], \text{ if } 0 \in I_0^i, \text{ and} \\ y_i &= h_i \text{ on } (c_{N_i}^i, T], \text{ if } T \in I_0^i, \end{aligned} \quad (57)$$

where  $c_1^i$  and  $d_1^i$  were introduced in Assumption 2.7. Define the set

$$\mathcal{P}_2 := \{(\xi, y, h) \in \mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m : (51), (54), (56) \text{ and } (57) \text{ hold}\}. \quad (58)$$

**Lemma 4.1.**  $\mathcal{P}$  is a dense subset of  $\mathcal{P}_2$  in the  $\mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m$ -topology.

*Proof.* The inclusion is immediate. In order to prove the density, consider the following sets:

$$\begin{aligned} X &:= \{(\xi, y, h) \in \mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m : (51), (56) \text{ and } (57) \text{ hold}\}, \\ L &:= \{(\xi, y, y(T)) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m : y(0) = 0, (51) \text{ and } (56) \text{ hold}\}, \\ C &:= \{(\xi, y, h) \in X : (54) \text{ holds}\}. \end{aligned}$$

By Lemma 7.1,  $L$  is a dense subset of  $X$ . Conclude with Lemma 2.8.  $\square$

## 4.2 Transformed second variation

We are interested in writing  $\Omega$  in terms of variables  $y$  and  $\xi$  defined in (50). Introduce the following notation for the sake of simplifying the presentation:

*Definition 4.2.* Consider the following matrices of sizes  $n \times n$ ,  $m \times n$  and  $m \times n$  respectively:

$$Q[\lambda] := H_{xx}[\lambda], \quad C[\lambda] := H_{ux}[\lambda], \quad M[\lambda] := B^\top Q[\lambda] - \dot{C}[\lambda] - C[\lambda]A, \quad (59)$$

where  $A$  and  $B$  were defined in (53). Notice that  $M$  is well-defined as  $C$  is Lipschitz-continuous on  $t$ . Decompose matrix  $C[\lambda]B$  in its symmetric and skew-symmetric parts, i.e., consider

$$S[\lambda] := \frac{1}{2}(C[\lambda]B + (C[\lambda]B)^\top), \quad V[\lambda] := \frac{1}{2}(C[\lambda]B - (C[\lambda]B)^\top). \quad (60)$$

*Remark 4.3.* Observe that  $S[\lambda]$  and  $V[\lambda]$  are Lipschitz-continuous as  $C[\lambda]$  and  $B$  are. In fact, simple computations yield:

$$S_{ij}[\lambda] = \frac{1}{2}\psi(f'_i f_j + f'_j f_i), \quad V_{ij}[\lambda] = \frac{1}{2}\psi[f_i, f_j], \quad \text{for } i, j = 1, \dots, m, \quad (61)$$

where

$$[f_i, f_j] := f'_i f_j - f'_j f_i. \quad (62)$$

With this notation  $\Omega$  takes the form:

$$\Omega[\lambda](\delta x, v) = \frac{1}{2}\ell''[\lambda](\hat{x}(T))(\delta x(T))^2 + \frac{1}{2}\int_0^T [(Q[\lambda]\delta x, \delta x) + 2(C[\lambda]\delta x, v)]dt.$$

Define the  $m \times m$ -matrix

$$R[\lambda] := B^\top Q[\lambda]B - C[\lambda]B_1 - (C[\lambda]B_1)^\top - \dot{S}[\lambda], \quad (63)$$

where  $B_1$  was introduced in equation (52). Consider the function  $g[\lambda]$  from  $\mathbb{R}^n \times \mathbb{R}^m$  to  $\mathbb{R}$  defined by:

$$g[\lambda](\zeta, h) := \ell''[\lambda](\hat{x}(T))(\zeta + B(T)h)^2 + (C[\lambda](T)(2\zeta + B(T)h), h). \quad (64)$$

*Remark 4.4.* (i) We will use the same notation for the matrices  $Q[\lambda]$ ,  $C[\lambda]$ ,  $M[\lambda]$ ,  $\ell''[\lambda](\hat{x}(T))$  and for the bilinear mapping they define.

(ii) Observe that when  $m = 1$ , the function  $V[\lambda] \equiv 0$  since it becomes a skew-symmetric scalar.

*Definition 4.5.* Define the mapping over  $\mathcal{X} \times \mathcal{Y} \times \mathcal{U}$ :

$$\begin{aligned} \Omega_{\mathcal{P}}[\lambda](\xi, y, v) &:= g[\lambda](\xi(T), y(T)) \\ &+ \int_0^T \{(Q[\lambda]\xi, \xi) + 2(M[\lambda]\xi, y) + (R[\lambda]y, y) + (V[\lambda]y, v)\}dt, \end{aligned} \quad (65)$$

with  $g[\lambda]$ ,  $Q[\lambda]$ ,  $M[\lambda]$ ,  $R[\lambda]$  and  $V[\lambda]$  defined in (59)-(64).

The following theorem shows that  $\Omega_{\mathcal{P}}$  coincides with  $\Omega$ . See e.g. [8].

**Theorem 4.6.** *Let  $(z, v) \in \mathcal{W}$  satisfying (12) and  $(\xi, y)$  defined by (50). Then*

$$\Omega[\lambda](z, v) = \Omega_{\mathcal{P}}[\lambda](\xi, y, v).$$



*Proof.* We omit the dependence on  $\lambda$  for the sake of simplicity. Replace  $z$  by its expression in (50) and obtain:

$$\begin{aligned} \Omega(z, v) &= \frac{1}{2} \ell''(\hat{x}(T))(\xi(T) + B(T)y(T))^2 \\ &+ \frac{1}{2} \int_0^T [(Q(\xi + By), \xi + By) + (C(\xi + By), v) + (C^\top v, \xi + By)] dt. \end{aligned} \quad (66)$$

Integrating by parts yields

$$\int_0^T (C\xi, v) dt = [(C\xi, y)]_0^T - \int_0^T (\dot{C}\xi + C(A\xi + B_1y), y) dt, \quad (67)$$

and

$$\begin{aligned} \int_0^T (CB_y, v) dt &= \int_0^T ((S + V)y, v) dt \\ &= \frac{1}{2} [(Sy, y)]_0^T + \int_0^T (-\frac{1}{2}(\dot{S}y, y) + (Vy, v)) dt. \end{aligned} \quad (68)$$

Combining (66), (67) and (68) we get the desired result.  $\square$

**Corollary 4.7.** *If  $V[\lambda] \equiv 0$  then  $\Omega$  does not involve explicitly  $v$ , and it can be expressed in terms of  $(\xi, y, y(T))$ .*

In view of (61), previous corollary holds in particular if  $[f_i, f_j] = 0$  on the reference trajectory for each pair  $1 \leq i < j \leq m$ .

**Corollary 4.8.** *If  $\hat{w}$  is a weak minimum, then*

$$\max_{\lambda \in \Lambda} \Omega_{\mathcal{P}}[\lambda](\xi, y, v) \geq 0, \quad (69)$$

for every  $(z, v) \in \mathcal{C}_2$ , and  $(\xi, y)$  defined by (50).

### 4.3 New second order condition

In this paragraph we present a necessary condition involving variable  $(\xi, y, h)$  in  $\mathcal{P}_2$ . To achieve this we remove the explicit dependence on  $v$  from the second variation, for certain multipliers. Recall that we consider  $\lambda = (\alpha, \beta)$  as elements of  $\mathbb{R}^s$ .

*Definition 4.9.* Given  $M \subset \mathbb{R}^s$ , define

$$G(M) := \{\lambda \in M : V_{ij}[\lambda](t) = 0 \text{ on } I_+^i \cap I_+^j, \text{ for any pair } 1 \leq i < j \leq m\}.$$

**Theorem 4.10.** *Let  $M \subset \mathbb{R}^s$  be convex and compact, and assume that*

$$\max_{\lambda \in M} \Omega_{\mathcal{P}}[\lambda](\xi, y, \dot{y}) \geq 0, \quad \text{for all } (\xi, y, h) \in \mathcal{P}. \quad (70)$$

Then

$$\max_{\lambda \in G(M)} \Omega_{\mathcal{P}}[\lambda](\xi, y, \dot{y}) \geq 0, \quad \text{for all } (\xi, y, h) \in \mathcal{P}. \quad (71)$$

The proof is based on some techniques introduced in Dmitruk [5, 8] for the proof of similar theorems.

Let  $1 \leq i < j \leq m$  and  $t^* \in \text{int } I_+^i \cap I_+^j$ . Take  $y \in \mathcal{Y}$  satisfying

$$y(0) = y(T) = 0, \quad y_k = 0, \text{ for } k \neq i, k \neq j. \quad (72)$$

Such functions define a linear continuous mapping  $r : \mathbb{R}^{s,*} \rightarrow \mathbb{R}$  by

$$\lambda \mapsto r[\lambda] := \int_0^T (V[\lambda](t^*)y, \dot{y}) dt. \quad (73)$$

By condition (72), and since  $V[\lambda]$  is skew-symmetric,

$$\int_0^T (V[\lambda](t^*)y, \dot{y}) dt = V_{ij}[\lambda](t^*) \int_0^T (y_i \dot{y}_j - y_j \dot{y}_i) dt. \quad (74)$$

Each  $r$  is an element of the dual space of  $\mathbb{R}^{s,*}$ , and then it can be identified with an element of  $\mathbb{R}^s$ . Consequently, the subset of  $\mathbb{R}^s$  defined by

$$R_{ij}(t^*) := \{r \in \mathbb{R}^s : y \in \mathcal{Y} \text{ satisfies (72), } r \text{ is defined by (73)}\}, \quad (75)$$

is a linear subspace of  $\mathbb{R}^s$ . Now, considering all the finite collections

$$\Theta_{ij} := \{\theta = \{t^1 < \dots < t^{N_\theta}\} : t^k \in \text{int } I_+^i \cap I_+^j \text{ for } k = 1, \dots, N_\theta\}, \quad (76)$$

define

$$\mathcal{R} := \sum_{i < j} \bigcup_{\theta \in \Theta_{ij}} \sum_{k=1}^{N_\theta} R_{ij}(t^k). \quad (77)$$

Note that  $\mathcal{R}$  is a linear subspace of  $\mathbb{R}^s$ . Given  $(\xi, y, y(T)) \in \mathcal{P}$ , let the mapping  $p_y : \mathbb{R}^{s,*} \rightarrow \mathbb{R}$  be given by

$$\lambda \mapsto p_y[\lambda] := \Omega_{\mathcal{P}}[\lambda](\xi, y, \dot{y}). \quad (78)$$

Thus,  $p_y$  is an element of  $\mathbb{R}^s$ .

**Lemma 4.11.** *Let  $(\bar{\xi}, \bar{y}, \bar{y}(T)) \in \mathcal{P}$  and  $r \in \mathcal{R}$ . Then there exists a sequence  $\{(\xi^\nu, y^\nu, y^\nu(T))\}$  in  $\mathcal{P}$  such that*

$$\Omega_{\mathcal{P}}[\lambda](\xi^\nu, y^\nu, \dot{y}^\nu) \longrightarrow p_{\bar{y}}[\lambda] + r[\lambda]. \quad (79)$$

*Proof.* Take  $(\bar{\xi}, \bar{y}, \bar{y}(T)) \in \mathcal{P}$ , its corresponding critical direction  $(\bar{z}, \bar{v}) \in \mathcal{C}$  related by (50) and  $p_{\bar{y}}$  defined in (78). Assume that  $r \in R_{ij}(t^*)$  for some  $1 \leq i < j \leq m$  and  $t^* \in \text{int } I_+^i \cap I_+^j$ , i.e.,  $r$  is associated via (73) to some function  $\tilde{y}$  verifying (72). Take  $\tilde{y}(t) = 0$  when  $t \notin [0, T]$ . Consider

$$\tilde{y}^\nu(t) := \tilde{y}(\nu(t - t^*)), \quad \check{y}^\nu := \bar{y} + \tilde{y}^\nu. \quad (80)$$

Let  $\check{\xi}^\nu$  be the solution of (51) corresponding to  $\check{y}^\nu$ . Observe that for large enough  $\nu$ , as  $t^* \in \text{int } I_+^i \cap I_+^j$ ,

$$\check{y}_k^\nu = 0, \text{ a.e. on } I_0^k, \text{ for } k = 1, \dots, m. \quad (81)$$

Let  $(\tilde{z}^\nu, \tilde{v}^\nu)$  and  $(\check{z}^\nu, \check{v}^\nu)$  be the points associated by transformation (50) with  $(\tilde{\xi}^\nu, \tilde{y}^\nu, \dot{y}^\nu(T))$  and  $(\check{\xi}^\nu, \check{y}^\nu, \dot{y}^\nu(T))$ , respectively. By (81), we get

$$\check{v}_k^\nu = 0, \text{ a.e. on } I_0^k, \text{ for } k = 1, \dots, m. \quad (82)$$

Note, however, that  $(\check{z}^\nu, \check{v}^\nu)$  can violate the terminal constraints defining  $\mathcal{C}_\infty$ , i.e., constraints in (20). Let us look for an estimate of the magnitude of this violation. Since

$$\|\tilde{y}^\nu\|_1 = \mathcal{O}(1/\nu), \quad (83)$$

and  $(\tilde{\xi}^\nu, \tilde{y}^\nu)$  is solution of (51), Gronwall's Lemma implies

$$|\tilde{\xi}^\nu(T)| = \mathcal{O}(1/\nu). \quad (84)$$

On the other hand, notice that  $\check{z}^\nu(T) = \bar{z}(T) + \tilde{\xi}^\nu(T)$ , and thus

$$|\check{z}^\nu(T) - \bar{z}(T)| = \mathcal{O}(1/\nu). \quad (85)$$

By Hoffman's Lemma (see [20]), there exists  $(\Delta z^\nu, \Delta v^\nu) \in \mathcal{W}$  satisfying  $\|\Delta v^\nu\|_\infty + \|\Delta z^\nu\|_\infty = \mathcal{O}(1/\nu)$ , and such that  $(z^\nu, v^\nu) := (\check{z}^\nu, \check{v}^\nu) + (\Delta z^\nu, \Delta v^\nu)$  belongs to  $\mathcal{C}_\infty$ . Let  $(\xi^\nu, y^\nu, \dot{y}^\nu(T)) \in \mathcal{P}$  be defined by (50). Let us show that for each  $\lambda \in M$ ,

$$\lim_{\nu \rightarrow \infty} \Omega_{\mathcal{P}}[\lambda](\xi^\nu, y^\nu, \dot{y}^\nu) = p_{\bar{y}}[\lambda] + r[\lambda]. \quad (86)$$

Observe that

$$\lim_{\nu \rightarrow \infty} \Omega_{\mathcal{P}}[\lambda](\xi^\nu, y^\nu, \dot{y}^\nu) - p_{\bar{y}}[\lambda] = \lim_{\nu \rightarrow \infty} \int_0^T \{(V[\lambda]\bar{y}, \dot{y}^\nu) + (V[\lambda]\tilde{y}^\nu, \dot{y}^\nu)\} dt, \quad (87)$$

since the terms involving  $\xi^\nu - \bar{\xi}$ ,  $y^\nu - \bar{y}$  or  $\Delta v^\nu$  vanish as  $\|\xi^\nu - \bar{\xi}\|_\infty \rightarrow 0$  and  $\|y^\nu - \bar{y}\|_1 \rightarrow 0$ . Integrating by parts the first term in the right hand-side of (87) we obtain

$$\int_0^T (V[\lambda]\bar{y}, \dot{y}^\nu) dt = [(V[\lambda]\bar{y}, \dot{y}^\nu)]_0^T - \int_0^T \{(\dot{V}[\lambda]\bar{y}, \dot{y}^\nu) + (V[\lambda]\dot{\bar{y}}, \dot{y}^\nu)\} dt \xrightarrow{\nu \rightarrow \infty} 0,$$

by (83) and since  $\tilde{y}^\nu(0) = \tilde{y}^\nu(T) = 0$ . Coming back to (87) we have

$$\lim_{\nu \rightarrow \infty} \Omega_{\mathcal{P}}[\lambda](\xi^\nu, y^\nu, \dot{y}^\nu) - p_{\bar{y}}[\lambda] = \lim_{\nu \rightarrow \infty} \int_0^T (V[\lambda]\tilde{y}^\nu, \dot{y}^\nu) dt \quad (88)$$

$$= \lim_{\nu \rightarrow \infty} \int_0^T (V[\lambda](t)\tilde{y}(\nu(t-t^*)), \dot{\tilde{y}}(\nu(t-t^*))) d\nu t \quad (89)$$

$$= \lim_{\nu \rightarrow \infty} \int_{-\nu t^*}^{\nu(T-t^*)} (V[\lambda](t^* + s/\nu)\tilde{y}(s), \dot{\tilde{y}}(s)) ds = r[\lambda], \quad (90)$$

and thus (79) holds when  $r \in R_{ij}(t^*)$ .

Consider the general case when  $r \in \mathcal{R}$ , i.e.,  $r = \sum_{i < j} \sum_{k=1}^{N_{ij}} r_{ij}^k$ , with each  $r_{ij}^k$  in  $R_{ij}(t_{ij}^k)$ . Let  $\tilde{y}_{ij}^k$  associated to  $r_{ij}^k$  by (73). Define  $\tilde{y}_{ij}^{k,\nu}$  as in (80), and follow the previous procedure for  $\bar{y} + \sum_{i < j} \sum_{k=1}^{N_{ij}} \tilde{y}_{ij}^{k,\nu}$  to get the desired result.  $\square$

*Proof.* [of **Theorem 4.10**] Take  $(\bar{\xi}, \bar{y}, \bar{y}(T)) \in \mathcal{P}$  and  $r \in \mathcal{R}$ . By Lemma 4.11 there exists a sequence  $\{(\xi^\nu, y^\nu, y^\nu(T))\}$  in  $\mathcal{P}$  such that for each  $\lambda \in M$ ,

$$\Omega_{\mathcal{P}}[\lambda](\xi^\nu, y^\nu, \dot{y}^\nu) \rightarrow \Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{y}, \dot{\bar{y}}) + r[\lambda]. \quad (91)$$

This convergence being uniform over  $M$ , from (70) we get that

$$\max_{\lambda \in M} (\Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{y}, \dot{\bar{y}}) + r[\lambda]) \geq 0, \quad \text{for all } r \in \mathcal{R}. \quad (92)$$

Hence

$$\inf_{r \in \mathcal{R}} \max_{\lambda \in M} (\Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{y}, \dot{\bar{y}}) + r[\lambda]) \geq 0, \quad (93)$$

where the expression in brackets is linear both in  $\lambda$  and  $r$ . Furthermore, note that  $M$  and  $\mathcal{R}$  are convex, and  $M$  is compact. In light of MinMax Theorem [38, Corollary 37.3.2, page 39] we can invert the order of inf and max in (93) and obtain

$$\max_{\lambda \in M} \inf_{r \in \mathcal{R}} (\Omega_{\mathcal{P}}[\lambda](\bar{\xi}, \bar{y}, \dot{\bar{y}}) + r[\lambda]) \geq 0. \quad (94)$$

Observe that if for certain  $\lambda \in M$ , there exists  $r \in \mathcal{R}$  with  $r[\lambda] \neq 0$ , then the infimum in the previous equation is  $-\infty$ . Since this infimum is nonnegative we can restrict the maximum to the set of  $\lambda \in M$  for which  $r[\lambda] = 0$  for every  $r \in \mathcal{R}$ . Note that this set is  $G(M)$ , and thus the conclusion follows.  $\square$

Consider for  $i, j = 1, \dots, m$  :

$$I_{ij} := \{t \in (0, T) : \hat{u}_i(t) = 0, \hat{u}_j(t) > 0\}. \quad (95)$$

By Assumption 2.7,  $I_{ij}$  can be expressed as a finite union of intervals, i.e.,

$$I_{ij} = \bigcup_{k=1}^{K_{ij}} I_{ij}^k, \quad \text{where } I_{ij}^k := (c_{ij}^k, d_{ij}^k). \quad (96)$$

Let  $(z, v) \in \mathcal{C}_\infty$ ,  $i \neq j$ ,  $y$  be defined by (50). Notice that  $y_i$  is constant on each  $(c_{ij}^k, d_{ij}^k)$ . Denote with  $y_{i,j}^k$  its value on this interval.

**Proposition 4.12.** *Let  $(z, v) \in \mathcal{C}_\infty$ ,  $y$  defined by (50) and  $\lambda \in G(\Lambda)$ . Then*

$$\int_0^T (V[\lambda]y, v) dt = \sum_{\substack{i \neq j \\ i, j=1}}^m \sum_{k=1}^{K_{ij}} y_{i,j}^k \left\{ [V_{ij}[\lambda]y_j]_{c_{ij}^k}^{d_{ij}^k} - \int_{c_{ij}^k}^{d_{ij}^k} \dot{V}_{ij}[\lambda]y_j dt \right\}. \quad (97)$$

*Proof.* Observe that

$$\int_0^T (V[\lambda]y, v) dt = \sum_{\substack{i \neq j \\ i, j=1}}^m \int_0^T V_{ij}[\lambda]y_i v_j dt, \quad (98)$$

since  $V_{ii}[\lambda] \equiv 0$ . Fix  $i \neq j$ , and recall that  $V_{ij}[\lambda]$  is differentiable on time (see expression (61)). Since  $(z, v) \in \mathcal{C}_\infty$  and  $\lambda \in G(\Lambda)$ ,

$$\begin{aligned} \int_0^T V_{ij}[\lambda]y_i v_j dt &= \int_{I_{ij}} V_{ij}[\lambda]y_i v_j dt = \sum_{k=1}^{K_{ij}} \int_{c_{ij}^k}^{d_{ij}^k} V_{ij}[\lambda]y_i v_j dt \\ &= \sum_{k=1}^{K_{ij}} y_{i,j}^k \left\{ [V_{ij}[\lambda]y_j]_{c_{ij}^k}^{d_{ij}^k} - \int_{c_{ij}^k}^{d_{ij}^k} \dot{V}_{ij}[\lambda]y_j dt \right\}, \end{aligned} \quad (99)$$

where the last equality was obtained integrating by parts and knowing that  $y_i$  is constant on  $I_{ij}$ . The desired result follows from (98) and (99).  $\square$

Given a real function  $h$  and  $c \in \mathbb{R}$ , define

$$h(c+) := \lim_{t \rightarrow c+} h(t), \text{ and } h(c-) := \lim_{t \rightarrow c-} h(t). \quad (100)$$

*Definition 4.13.* Let  $(\xi, y, h) \in \mathcal{P}_2$  and  $\lambda \in G(\Lambda)$ . Define

$$\Xi[\lambda](\xi, y, h) := 2 \sum_{\substack{i \neq j \\ i, j=1}}^m \sum_{\substack{k=1 \\ c_{ij}^k \neq 0}}^{K_{ij}} y_{i,j}^k \left\{ V_{ij}[\lambda](d_{ij}^k) y_j(d_{ij}^k+) - V_{ij}[\lambda](c_{ij}^k) y_j(c_{ij}^k-) - \int_{c_{ij}^k}^{d_{ij}^k} \dot{V}_{ij}[\lambda] y_j dt \right\},$$

where the above expression is interpreted as follows:

- (i)  $y_j(d_{ij}^k+) := h_j$ , if  $d_{ij}^k = T$ ,
- (ii)  $V_{ij}[\lambda](c_{ij}^k) y_j(c_{ij}^k-) := 0$ , if  $\hat{u}_i > 0$  and  $\hat{u}_j > 0$  for  $t < c_{ij}^k$ ,
- (iii)  $V_{ij}[\lambda](d_{ij}^k) y_j(d_{ij}^k+) := 0$ , if  $\hat{u}_i > 0$  and  $\hat{u}_j > 0$  for  $t > d_{ij}^k$ .

**Proposition 4.14.** *The following properties for  $\Xi$  hold:*

- (i)  $\Xi[\lambda](\xi, y, h)$  is well-defined for each  $(\xi, y, h) \in \mathcal{P}_2$ , and  $\lambda \in G(\Lambda)$ .
- (ii) If  $\{(\xi^\nu, y^\nu, y^\nu(T))\} \subset \mathcal{P}$  converges in the  $\mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m$ - topology to  $(\xi, y, h) \in \mathcal{P}_2$ , then

$$\int_0^T (V[\lambda] y^\nu, \dot{y}^\nu) dt \xrightarrow{\nu \rightarrow \infty} \Xi[\lambda](\xi, y, h). \quad (101)$$

*Proof.* (i) Take  $(\xi, y, h) \in \mathcal{P}_2$ . First observe that  $y_i \equiv y_{i,j}^k$  over  $(c_{ij}^k, d_{ij}^k)$ . As  $c_{ij}^k \neq 0$ , two possible situations can arise:

- (a) For  $t < c_{ij}^k$  :  $\hat{u}_j = 0$ , thus  $y_j$  is constant, and consequently  $y_j(c_{ij}^k-)$  is well-defined.
- (b) For  $t < c_{ij}^k$  :  $\hat{u}_i > 0$  and  $\hat{u}_j > 0$ , thus  $V_{ij}[\lambda](c_{ij}^k) = 0$  since  $\lambda \in G(\Lambda)$ .

The same analysis can be done for  $t > d_{ij}^k$  when  $d_{ij}^k \neq T$ . Conclude that  $\Xi$  is correctly defined.

(ii) Observe that since  $y^\nu$  converges to  $y$  in the  $\mathcal{U}_2$ -topology and since  $y_i^\nu$  is constant over  $I_{ij}$ , then  $y_i$  is constant as well, and  $y_i^\nu$  goes to  $y_i$  pointwise on  $I_{ij}$ . Thus,  $y_i^\nu(c_{ij}^k) \rightarrow y_{i,j}^k$ , and  $y_i^\nu(d_{ij}^k) \rightarrow y_{i,j}^k$ . Now, for the terms on  $y_j$  the same analysis can be made and yields to either  $y_j^\nu(c_{ij}^k) \rightarrow y_j(c_{ij}^k-)$  or  $V_{ij}[\lambda](c_{ij}^k) = 0$ ; and, either  $y_j^\nu(d_{ij}^k) \rightarrow y_j(d_{ij}^k+)$  or  $V_{ij}[\lambda](d_{ij}^k) = 0$ , when  $d_{ij}^k < T$ . For  $d_{ij}^k = T$ , it holds  $y_j^\nu(T) \rightarrow h_j$ .  $\square$

*Definition 4.15.* For  $(\xi, y, h) \in \mathcal{P}_2$  and  $\lambda \in G(\Lambda)$  define

$$\begin{aligned} \Omega_{\mathcal{P}_2}[\lambda](\xi, y, h) &:= g[\lambda](\xi(T), h) + \Xi[\lambda](\xi, y, h) \\ &+ \int_0^T ((Q[\lambda]\xi, \xi) + 2(M[\lambda]\xi, y) + (R[\lambda]y, y)) dt. \end{aligned} \quad (102)$$

*Remark 4.16.* Observe that when  $m = 1$ , the mapping  $\Xi \equiv 0$  since  $V \equiv 0$ . Thus, in this case,  $\Omega_{\mathcal{P}_2}$  can be defined for any element  $(\xi, y, h) \in \mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}$  and any  $\lambda \in \Lambda$ . If we take  $(z, v) \in \mathcal{W}$  satisfying (12), and define  $(\xi, y)$  by (50), then

$$\Omega[\lambda](z, v) = \Omega_{\mathcal{P}}[\lambda](\xi, y, \dot{y}) = \Omega_{\mathcal{P}_2}[\lambda](\xi, y, y(T)). \quad (103)$$

For  $m > 1$ , previous equality holds for  $(z, v) \in \mathcal{C}_\infty$ .

**Lemma 4.17.** *Let  $\{(\xi^\nu, y^\nu, y^\nu(T))\} \subset \mathcal{P}$  be a sequence converging to  $(\xi, y, h) \in \mathcal{P}_2$  in the  $\mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m$ -topology. Then*

$$\lim_{\nu \rightarrow \infty} \Omega_{\mathcal{P}}[\lambda](\xi^\nu, y^\nu, \dot{y}^\nu) = \Omega_{\mathcal{P}_2}[\lambda](\xi, y, h). \quad (104)$$

Denote with  $\text{co } \Lambda$  the convex hull of  $\Lambda$ .

**Theorem 4.18.** *Let  $\hat{w}$  be a weak minimum, then*

$$\max_{\lambda \in G(\text{co } \Lambda)} \Omega_{\mathcal{P}_2}[\lambda](\xi, y, h) \geq 0, \quad \text{for all } (\xi, y, h) \in \mathcal{P}_2. \quad (105)$$

*Proof.* Corollary 4.8 together with Theorem 4.10 applied to  $M := \text{co } \Lambda$  yield

$$\max_{\lambda \in G(\text{co } \Lambda)} \Omega_{\mathcal{P}}[\lambda](\xi, y, \dot{y}) \geq 0, \quad \text{for all } (\xi, y, y(T)) \in \mathcal{P}. \quad (106)$$

The result follows from density result 4.1 and Lemma 4.17.  $\square$

*Remark 4.19.* Notice that in case (31) is not satisfied condition (105) does not provide any useful information as  $0 \in \text{co } \Lambda$ . On the other hand, in case (31) holds, every  $\lambda = (\alpha, \beta, \psi) \in \Lambda$  necessarily has  $\alpha \neq 0$ , and thus  $0 \notin \text{co } \Lambda$ .

## 5 Sufficient condition

Consider the problem for a scalar control, i.e., let  $m = 1$ . This section provides a sufficient condition for Pontryagin optimality.

*Definition 5.1.* Given  $(y, h) \in \mathcal{U}_2 \times \mathbb{R}$ , let

$$\gamma(y, h) := \int_0^T y(t)^2 dt + |h|^2. \quad (107)$$

*Definition 5.2.* A sequence  $\{v_k\} \subset \mathcal{U}$  converges to 0 in the Pontryagin sense if  $\|v_k\|_1 \rightarrow 0$  and there exists  $N$  such that  $\|v_k\|_\infty < N$ .

*Definition 5.3.* We say that  $\hat{w}$  satisfies  $\gamma$ -quadratic growth condition in the Pontryagin sense if there exists  $\rho > 0$  such that, for every sequence of feasible variations  $\{(\delta x_k, v_k)\}$  with  $\{v_k\}$  converging to 0 in the Pontryagin sense,

$$J(\hat{u} + v_k) - J(\hat{u}) \geq \rho \gamma(y_k, y_k(T)), \quad (108)$$

holds for big enough  $k$ , where  $y_k$  is defined by (50). Equivalently, for all  $N > 0$ , there exists  $\varepsilon > 0$  such that if  $\|v\|_\infty < N$  and  $\|v\|_1 < \varepsilon$ , then (108) holds.

*Definition 5.4.* We say that  $\hat{w}$  is normal if  $\alpha_0 > 0$  for every  $\lambda \in \Lambda$ .

**Theorem 5.5.** *Suppose that there exists  $\rho > 0$  such that*

$$\max_{\lambda \in \Lambda} \Omega_{\mathcal{P}_2}[\lambda](\xi, y, h) \geq \rho \gamma(y, h), \quad \text{for all } (\xi, y, h) \in \mathcal{P}_2. \quad (109)$$

*Then  $\hat{w}$  is a Pontryagin minimum satisfying  $\gamma$ -quadratic growth. And in case  $\hat{w}$  is normal, the converse holds.*

*Remark 5.6.* In case the bang arcs are absent, i.e., the control is totally singular, this theorem reduces to one proved in Dmitruk [6, 8].

Recall that  $\Phi$  is defined in (5). We will use the following technical result:

**Lemma 5.7.** *Consider  $\{v_k\} \subset \mathcal{U}$  converging to 0 in the Pontryagin sense. Let  $u_k := \hat{u} + v_k$  and  $x_k$  the corresponding solution of equation (2). Then for every  $\lambda \in \Lambda$ :*

$$\Phi[\lambda](x_k, u_k) = \Phi[\lambda](\hat{x}, \hat{u}) + \int_0^T H_u[\lambda](t)v_k(t)dt + \Omega[\lambda](z_k, v_k) + o(\gamma_k), \quad (110)$$

where  $z_k$  is defined by (12),  $\gamma_k := \gamma(y_k, y_k(T))$ , and  $y_k$  is defined by (50).

*Proof.* By Lemma 3.1 we can write

$$\Phi[\lambda](x_k, u_k) = \Phi[\lambda](\hat{x}, \hat{u}) + \int_0^T H_u[\lambda](t)v_k(t)dt + \Omega[\lambda](z_k, v_k) + R_k, \quad (111)$$

where, in view of Lemma 7.5,

$$R_k := \Delta_k \Omega[\lambda] + \int_0^T (H_{uxx}[\lambda](t)\delta x_k(t), \delta x_k(t), v_k(t))dt + o(\gamma_k), \quad (112)$$

with  $\delta x_k := x_k - \hat{x}$ , and

$$\Delta_k \Omega[\lambda] := \Omega[\lambda](\delta x_k, v_k) - \Omega[\lambda](z_k, v_k). \quad (113)$$

Next, prove that

$$R_k = o(\gamma_k). \quad (114)$$

Note that  $\mathcal{Q}(a, a) - \mathcal{Q}(b, b) = \mathcal{Q}(a + b, a - b)$ , for any bilinear mapping  $\mathcal{Q}$ , and any pair  $a, b$ . Put  $\eta_k := \delta x_k - z_k$ . Hence, from (113) we get,

$$\begin{aligned} \Delta_k \Omega[\lambda] &= \frac{1}{2} \ell''[\lambda](\hat{x}(T))(\delta x_k(T) + z_k(T), \eta_k(T)) \\ &\quad + \frac{1}{2} \int_0^T (H_{xx}[\lambda](\delta x_k + z_k), \eta_k)dt + \int_0^T (H_{ux}[\lambda]\eta_k, v_k)dt. \end{aligned}$$

By Lemmas 7.5 and 7.13 in the Appendix, the first and the second terms are of order  $o(\gamma_k)$ . Integrate by parts the last term\* and obtain:

$$\int_0^T (H_{ux}[\lambda]\eta_k, v_k)dt \quad (115)$$

$$= [(H_{ux}[\lambda]\eta_k, y_k)]_0^T - \int_0^T \{(\dot{H}_{ux}[\lambda]\eta_k, y_k) + (H_{ux}[\lambda]\dot{\eta}_k, y_k)\}dt. \quad (116)$$

---

\*observe that  $H_{ux}, \eta_k$  and  $y_k$  are functions of bounded variation

Thus by Lemma 7.13 we get that the first two terms in (116) are of order  $o(\gamma_k)$ . It remains to deal with last term in the integral. Replace  $\dot{\eta}_k$  by its expression in equation (241) of Lemma 7.13:

$$\begin{aligned} \int_0^T (H_{ux}[\lambda]\dot{\eta}_k, y_k)dt &= \int_0^T (H_{ux}[\lambda] \left( \sum_{i=0}^1 \hat{u}_i f'_i(\hat{x})\eta_k + v_k f'_1(\hat{x})\delta x_k + \zeta_k \right), y_k)dt \\ &= o(\gamma_k) + \int_0^T \frac{d}{dt} \left( \frac{y_k^2}{2} \right) H_{ux}[\lambda] f'_1(\hat{x}) \delta x_k dt, \end{aligned} \quad (117)$$

where the second equality follows from Lemmas 7.5 and 7.13. Integrating by parts the last term, we get

$$\begin{aligned} \int_0^T \frac{d}{dt} \left( \frac{y_k^2}{2} \right) H_{ux}[\lambda] f'_1(\hat{x}) \delta x_k dt &= \left[ \frac{y_k^2}{2} H_{ux}[\lambda] f'_1(\hat{x}) \delta x_k \right]_0^T \\ &- \int_0^T \frac{y_k^2}{2} \frac{d}{dt} (H_{ux}[\lambda] f'_1(\hat{x})) \delta x_k dt - \int_0^T \frac{y_k^2}{2} H_{ux}[\lambda] f'_1(\hat{x}) \delta \dot{x}_k dt \\ &= o(\gamma_k) - \int_0^T \frac{d}{dt} \left( \frac{y_k^3}{6} \right) H_{ux}[\lambda] f'_1(\hat{x}) f_1(\hat{x}) dt \\ &= o(\gamma_k) - \left[ \frac{y_k^3}{6} H_{ux}[\lambda] f'_1(\hat{x}) f_1(\hat{x}) \right]_0^T + \int_0^T \frac{y_k^3}{6} \frac{d}{dt} (H_{ux}[\lambda] f'_1(\hat{x}) f_1(\hat{x})) dt \\ &= o(\gamma_k), \end{aligned} \quad (118)$$

where we used Lemma 7.13, and in particular equation (242). From (117) and (118) follows that the term in (115) is of order  $o(\gamma_k)$ . Thus,

$$\Delta_k \Omega[\lambda] \leq o(\gamma_k). \quad (119)$$

Study now the third order term in (112):

$$\begin{aligned} \int_0^T (H_{uxx}[\lambda] \delta x_k, \delta x_k, v_k) dt &= [y_k \delta x_k^\top H_{uxx}[\lambda] \delta x_k]_0^T \\ &- \int_0^T y_k \delta x_k^\top \dot{H}_{uxx}[\lambda] \delta x_k dt - 2 \int_0^T y_k \delta x_k^\top H_{uxx}[\lambda] \delta \dot{x}_k dt \\ &= o(\gamma_k) - \int_0^T \frac{d}{dt} (y_k^2) \delta x_k^\top H_{uxx}[\lambda] f_1(\hat{x}) dt \\ &= o(\gamma_k) - [y_k^2 \delta x_k^\top H_{uxx}[\lambda] f_1(\hat{x})]_0^T - \int_0^T y_k^2 v_k f_1(\hat{x})^\top H_{uxx}[\lambda] f_1(\hat{x}) dt = o(\gamma_k), \end{aligned} \quad (120)$$

by Lemmas 7.5 and 7.13. Last inequality follows integrating by parts one more time as it was done in (118). Consider expression (112). By inequality (119) and equation (120), equality (114) is obtained, and thus the desired results follows.  $\square$

*Proof.* [of Theorem 5.5] *Part 1.* First we prove that if  $\hat{w}$  is a normal Pontryagin minimum satisfying  $\gamma$ -quadratic growth condition in the Pontryagin sense then



(109) holds for some  $\rho > 0$ . Here the necessary condition of Theorem 3.2 is used. Define  $\hat{y}(t) := \int_0^t \hat{u}(s)ds$ , and note that  $(\hat{w}, \hat{y})$  is, for some  $\rho' > 0$ , a Pontryagin minimum of

$$\begin{aligned} \tilde{J} &:= J - \rho' \gamma(y, y(T)) \rightarrow \min, \\ (2)-(4), \dot{y} &= u, y(0) = 0. \end{aligned} \quad (121)$$

Observe that the critical cone  $\tilde{\mathcal{C}}_2$  for (121) consists of the points  $(z, v, \delta y)$  in  $\mathcal{X}_2 \times \mathcal{U}_2 \times W_2^1(0, T; \mathbb{R})$  verifying  $(z, v) \in \mathcal{C}_2$ ,  $\delta \dot{y} = v$  and  $\delta y(0) = 0$ . Since the pre-Hamiltonian at point  $(\hat{w}, \hat{y})$  coincides with the original pre-Hamiltonian, the set of multipliers for (121) consists of the points  $(\lambda, \psi_y)$  with  $\lambda \in \Lambda$ .

Applying second order necessary condition of Theorem 3.2 at the point  $(\hat{w}, \hat{y})$  we get that for every  $(z, v) \in \mathcal{C}_2$  and  $\delta y(t) := \int_0^t v(s)ds$ , there exists  $\lambda \in \Lambda$  such that

$$\Omega[\lambda](z, v) - \alpha_0 \rho' (\|\delta y\|_2^2 + \delta y^2(T)) \geq 0, \quad (122)$$

where  $\alpha_0 > 0$  since  $\hat{w}$  is normal. Take  $\rho := \min_{\lambda \in \Lambda} \alpha_0 \rho' > 0$ . Performing Goh transformation in (122), condition (109) for the constant  $\rho$  follows.

*Part 2.* We shall prove that if (109) holds for some  $\rho > 0$ , then  $\hat{w}$  satisfies  $\gamma$ -quadratic growth in the Pontryagin sense. By *reductio ad absurdum* assume that the quadratic growth condition (108) is not valid. Consequently, there exists a sequence  $\{v_k\} \subset \mathcal{U}$  converging to 0 in the Pontryagin sense such that, denoting  $u_k := \hat{u} + v_k$ ,

$$J(\hat{u} + v_k) \leq J(\hat{u}) + o(\gamma_k), \quad (123)$$

where  $y_k(t) := \int_0^t v_k(s)ds$  and  $\gamma_k := \gamma(y_k, y_k(T))$ . Denote by  $x_k$  the solution of equation (2) corresponding to  $u_k$ ,  $w_k := (x_k, u_k)$  and let  $z_k$  be the solution of (12) associated with  $v_k$ . Take any  $\lambda \in \Lambda$ . Multiply inequality (123) by  $\alpha_0$ , add the nonpositive term  $\sum_{i=0}^{d_\varphi} \alpha_i \varphi_i(x_k(T)) + \sum_{j=1}^{d_\eta} \beta_j \eta_j(x_k(T))$  to its left-hand side, and obtain the inequality

$$\Phi[\lambda](x_k, u_k) \leq \Phi[\lambda](\hat{x}, \hat{u}) + o(\gamma_k). \quad (124)$$

Recall expansion (110). Let  $(\bar{y}_k, \bar{h}_k) := (y_k, y_k(T))/\sqrt{\gamma_k}$ . Note that the elements of this sequence have unit norm in  $\mathcal{U}_2 \times \mathbb{R}$ . By Banach-Alaoglu Theorem, extracting if necessary a sequence, we may assume that there exists  $(\bar{y}, \bar{h}) \in \mathcal{U}_2 \times \mathbb{R}$  such that

$$\bar{y}_k \rightharpoonup \bar{y}, \text{ and } \bar{h}_k \rightarrow \bar{h}, \quad (125)$$

where the first limit is taken in the weak topology of  $\mathcal{U}_2$ . The remainder of the proof is split in two parts:

- (a) Using equations (110) and (124) we prove that  $(\bar{\xi}, \bar{y}, \bar{h}) \in \mathcal{P}_2$ , where  $\bar{\xi}$  is solution of (51).
- (b) We prove that  $(\bar{y}, \bar{h}) = 0$  and that it is the limit of  $\{(\bar{y}_k, \bar{h}_k)\}$  in the strong sense. This leads to a contradiction since each  $(\bar{y}_k, \bar{h}_k)$  has unit norm.

(a) We shall prove that  $(\bar{\xi}, \bar{y}, \bar{h}) \in \mathcal{P}_2$ . From (110) and (124) follows:

$$0 \leq \int_0^T H_u[\lambda](t) v_k(t) dt \leq -\Omega_{\mathcal{P}_2}[\lambda](\xi_k, y_k, h_k) + o(\gamma_k), \quad (126)$$

where  $\xi_k$  is solution of (51) corresponding to  $y_k$ , the first inequality holds as  $H_u[\lambda]v_k \geq 0$  almost everywhere on  $[0, T]$  and we replaced  $\Omega_{\mathcal{P}}$  by  $\Omega_{\mathcal{P}_2}$  in view of Remark 4.16. By the continuity of mapping  $\Omega_{\mathcal{P}_2}[\lambda]$  over  $\mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}$  deduce that

$$0 \leq \int_0^T H_u[\lambda](t)v_k(t)dt \leq O(\gamma_k), \quad (127)$$

and thus, for each composing interval  $(c, d)$  of  $I_0$ ,

$$\lim_{k \rightarrow \infty} \int_c^d H_u[\lambda](t)\varphi(t) \frac{v_k(t)}{\sqrt{\gamma_k}} dt = 0, \quad (128)$$

for every nonnegative Lipschitz continuous function  $\varphi$  with  $\text{supp } \varphi \subset (c, d)$ . Integrating by parts in (128) and by (125) obtain

$$0 = \lim_{k \rightarrow \infty} \int_c^d \frac{d}{dt} (H_u[\lambda](t)\varphi(t)) \bar{y}_k(t) dt = \int_c^d \frac{d}{dt} (H_u[\lambda](t)\varphi(t)) \bar{y}(t) dt.$$

By Lemma 7.6,  $\bar{y}$  is nondecreasing over  $(c, d)$ . Hence, in view of Lemma 7.8, we can integrate by parts in previous equation and get

$$\int_c^d H_u[\lambda](t)\varphi(t) d\bar{y}(t) = 0. \quad (129)$$

Take  $t_0 \in (c, d)$ . By strict complementary assumption 2.6, there exists  $\lambda_0 \in \Lambda$  such that  $H_u[\lambda_0](t_0) > 0$ . Hence, in view of the continuity of  $H_u[\lambda_0]$ , there exists  $\varepsilon > 0$  such that  $H_u[\lambda_0] > 0$  on  $(t_0 - 2\varepsilon, t_0 + 2\varepsilon) \subset (c, d)$ . Choose  $\varphi$  such that  $\text{supp } \varphi \subset (t_0 - 2\varepsilon, t_0 + 2\varepsilon)$ , and  $H_u[\lambda_0](t)\varphi(t) = 1$  on  $(t_0 - \varepsilon, t_0 + \varepsilon)$ . Since  $d\bar{y} \geq 0$ , equation (129) yields

$$\begin{aligned} 0 &= \int_c^d H_u[\lambda](t)\varphi(t) d\bar{y}(t) \geq \int_{t_0 - \varepsilon}^{t_0 + \varepsilon} H_u[\lambda](t)\varphi(t) d\bar{y}(t) \\ &= \int_{t_0 - \varepsilon}^{t_0 + \varepsilon} d\bar{y}(t) = \bar{y}(t_0 + \varepsilon) - \bar{y}(t_0 - \varepsilon). \end{aligned} \quad (130)$$

As  $\varepsilon$  and  $t_0 \in (c, d)$  are arbitrary we get that

$$d\bar{y}(t) = 0, \quad \text{on } I_0, \quad (131)$$

and thus (56) holds. Let us prove condition (57) for  $(\bar{\xi}, \bar{y}, \bar{h})$ . Suppose that  $0 \in I_0$ . Take  $\varepsilon > 0$ , and notice that by Assumption 2.6 there exists  $\lambda' \in \Lambda$  and  $\delta > 0$  such that  $H_u[\lambda'](t) > \delta$  for  $t \in [0, d_1 - \varepsilon]$ , and thus by (128) we get  $\int_0^{d_1 - \varepsilon} v_k(t)/\sqrt{\gamma_k} dt \rightarrow 0$ , as  $v_k \geq 0$ . Then for all  $s \in [0, d_1]$  holds

$$\bar{y}_k(s) \rightarrow 0, \quad (132)$$

and thus

$$\bar{y} = 0, \quad \text{on } [0, d_1], \quad \text{if } 0 \in I_0. \quad (133)$$

Suppose that  $T \in I_0$ . Then, we get that  $\int_{a_N + \varepsilon}^T \bar{v}_k(t) dt \rightarrow 0$  by an analogous argument. Thus, the following pointwise convergence holds for every  $s \in (a_N, T]$ :

$$\bar{h}_k - \bar{y}_k(s) \rightarrow 0, \quad (134)$$

and then,

$$\bar{y} = \bar{h}, \text{ on } (a_N, T], \text{ if } T \in I_0. \quad (135)$$

It remains to check the final conditions (54) for  $\bar{h}$ . Let  $0 \leq i \leq d_\varphi$ ,

$$\varphi'_i(\hat{x}(T))(\bar{\xi}(T) + B(T)\bar{h}) = \lim_{k \rightarrow \infty} \varphi'_i(\hat{x}(T)) \left( \frac{\xi_k(T) + B(T)h_k}{\sqrt{\gamma_k}} \right) \quad (136)$$

$$= \lim_{k \rightarrow \infty} \varphi'_i(\hat{x}(T)) \frac{z_k(T)}{\sqrt{\gamma_k}}. \quad (137)$$

A first order Taylor expansion of function  $\varphi_i$  around  $\hat{x}(T)$  gives

$$\varphi_i(x_k(T)) = \varphi_i(\hat{x}(T)) + \varphi'_i(\hat{x}(T))\delta x_k(T) + O(|\delta x_k(T)|^2).$$

By Lemmas 7.5 and 7.13 in the Appendix we can write

$$\varphi_i(x_k(T)) = \varphi_i(\hat{x}(T)) + \varphi'_i(\hat{x}(T))z_k(T) + o(\sqrt{\gamma_k}). \quad (138)$$

Thus

$$\varphi'_i(\hat{x}(T)) \frac{z_k(T)}{\sqrt{\gamma_k}} = \frac{\varphi_i(x_k(T)) - \varphi_i(\hat{x}(T))}{\sqrt{\gamma_k}} + o(1). \quad (139)$$

Since  $x_k$  satisfies (4), equations (136) and (139) yield, for  $1 \leq i \leq d_\varphi$  :  $\varphi'_i(\hat{x}(T))(\bar{\xi}(T) + B(T)\bar{h}) \leq 0$ . For  $i = 0$  use inequality (123). Analogously,

$$\eta'_j(\hat{x}(T))(\bar{\xi}(T) + B(T)\bar{h}) = 0, \quad \text{for } j = 1, \dots, d_\eta. \quad (140)$$

Thus  $(\bar{\xi}, \bar{y}, \bar{h})$  satisfies (54), and by (131), (133) and (135), we obtain

$$(\bar{\xi}, \bar{y}, \bar{h}) \in \mathcal{P}_2. \quad (141)$$

(b) Come back to the expansion (110). Equation (124) and  $H_u[\lambda] \geq 0$  imply

$$\begin{aligned} \Omega_{\mathcal{P}_2}[\lambda](\xi_k, y_k, y_k(T)) = \\ \Phi[\lambda](x_k, u_k) - \Phi[\lambda](\hat{x}, \hat{u}) - \int_0^T H_u[\lambda]v_k dt - o(\gamma_k) \leq o(\gamma_k). \end{aligned} \quad (142)$$

Thus

$$\liminf_{k \rightarrow \infty} \Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}_k, \bar{y}_k, \bar{h}_k) \leq \limsup_{k \rightarrow \infty} \Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}_k, \bar{y}_k, \bar{h}_k) \leq 0. \quad (143)$$

Recall Assumption 2.7, and let  $N$  be the number of connected components of  $I_0$ . Set  $\varepsilon > 0$ , and for each composing interval  $(c, d)$  of  $I_0$ , consider a smaller interval of the form  $(c + \varepsilon/2N, d - \varepsilon/2N)$ . Call  $I_0^\varepsilon$  their union. Notice that  $I_0 \setminus I_0^\varepsilon$  is of measure  $\varepsilon$ . Put  $I_+^\varepsilon := [0, T] \setminus I_0^\varepsilon$ . Split  $\Omega_{\mathcal{P}_2}$  as follows

$$\begin{aligned} \Omega_{\mathcal{P}_2,0}^\varepsilon[\lambda](\xi, y, h) := \\ \int_0^T \{(Q[\lambda]\xi, \xi) + (M[\lambda]\xi, y)\} dt + \int_{I_0^\varepsilon} (R[\lambda]y, y) dt + g[\lambda](\xi(T), h), \end{aligned} \quad (144)$$

and

$$\Omega_{\mathcal{P}_2,+}^\varepsilon[\lambda](y) := \int_{I_+^\varepsilon} (R[\lambda]y, y) dt. \quad (145)$$

Consider the Hilbert space

$$\Gamma_2 := \{(\xi, y, h) \in \mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R} : (51), (56) \text{ and } (57) \text{ hold}\}. \quad (146)$$

Take  $\rho > 0$  the constant in (109) and define

$$\Lambda^\rho := \{\lambda \in \text{co } \Lambda : \Omega_{\mathcal{P}_2}[\lambda] - \rho\gamma \text{ is weakly l.s.c. on } \Gamma_2\}. \quad (147)$$

Equation (109) and Lemma 7.12 imply that

$$\max_{\lambda \in \Lambda^\rho} \Omega_{\mathcal{P}_2}[\lambda](\bar{\xi}, \bar{y}, \bar{h}) \geq \rho\gamma(\bar{y}, \bar{h}). \quad (148)$$

Denote by  $\bar{\lambda}$  the element in  $\Lambda^\rho$  that reaches the maximum in (148). Observe that  $\Omega_{\mathcal{P}_2}[\bar{\lambda}]$  is weakly continuous. Thus,  $\Omega_{\mathcal{P}_2}[\bar{\lambda}] - \rho\gamma$  is weakly l.s.c. if and only if the quadratic mapping

$$y \mapsto \Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}](y) - \rho \int_{I_+^\varepsilon} |y(t)|^2 dt, \quad (149)$$

is weakly l.s.c. over the subspace  $\mathcal{U}_{2, I_+^\varepsilon}$  of  $\mathcal{U}_2$  consisting of the functions that are constant over each connected component of  $[0, T] \setminus I_+^\varepsilon$ . Applying Lemma 7.11 to the mapping in (149) we get that  $R[\bar{\lambda}](t) \succeq \rho$  on  $I_+^\varepsilon$ . Consequently  $\Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}]$  is a Legendre form on  $\mathcal{U}_{2, I_+^\varepsilon}$ , and thus

$$\Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}](\bar{y}) \leq \liminf_{k \rightarrow \infty} \Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}](\bar{y}_k). \quad (150)$$

Hence,

$$\begin{aligned} \rho\gamma(\bar{y}, \bar{h}) &\leq \Omega_{\mathcal{P}_2}[\bar{\lambda}](\bar{\xi}, \bar{y}, \bar{h}) \leq \lim_{k \rightarrow \infty} \Omega_{\mathcal{P}_2, 0}^\varepsilon[\bar{\lambda}](\xi_k, \bar{y}_k, \bar{h}_k) \\ &\quad + \liminf_{k \rightarrow \infty} \Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}](\bar{y}_k) = \liminf_{k \rightarrow \infty} \Omega_{\mathcal{P}_2}[\bar{\lambda}](\bar{\xi}_k, \bar{y}_k, \bar{h}_k), \end{aligned} \quad (151)$$

where first inequality is due to condition (109), and the second inequality is an immediate consequence of (150). Inequality (143) implies that the right-hand side of the last expression is nonpositive. Hence,

$$(\bar{y}, \bar{h}) = 0, \text{ and } \lim_{k \rightarrow \infty} \Omega_{\mathcal{P}_2}[\bar{\lambda}](\bar{\xi}_k, \bar{y}_k, \bar{h}_k) = 0. \quad (152)$$

As  $\Omega_{\mathcal{P}_2, 0}^\varepsilon[\bar{\lambda}]$  is weakly continuous,  $\lim_{k \rightarrow \infty} \Omega_{\mathcal{P}_2, 0}^\varepsilon[\bar{\lambda}](\bar{\xi}_k, \bar{y}_k, \bar{h}_k) = 0$  and thus

$$\lim_{k \rightarrow \infty} \Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}](\bar{y}_k) = 0. \quad (153)$$

We have:  $\Omega_{\mathcal{P}_2, +}^\varepsilon[\bar{\lambda}]$  is a Legendre form and  $\bar{y}_k \rightarrow 0$  on  $I_+^\varepsilon$ , thus by (153),

$$\bar{y}_k \rightarrow 0, \quad \text{on } \mathcal{U}_{2, I_+^\varepsilon}. \quad (154)$$

As  $\{\bar{y}_k\}$  converges uniformly on  $I_0^\varepsilon$  holds, the strong convergence holds on  $[0, T]$ , and thus

$$(\bar{y}_k, \bar{h}_k) \longrightarrow (0, 0), \quad \text{on } \mathcal{U}_2 \times \mathbb{R}. \quad (155)$$

This leads to a contradiction since  $(\bar{y}_k, \bar{h}_k)$  has unit norm for every  $k \in \mathbb{N}$ . Thus,  $\hat{w}$  is a Pontryagin minimum satisfying quadratic growth.  $\square$

## 6 Extensions and an example

### 6.1 Including parameters

Consider the following optimal control problem, where the initial state is not determined, some parameters are included and a more general control constraint is considered:

$$J := \varphi_0(x(0), x(T), r(0)) \rightarrow \min, \quad (156)$$

$$\dot{x}(t) = \sum_{i=0}^m u_i(t) f_i(x(t), r(t)), \quad (157)$$

$$\dot{r}(t) = 0, \quad (158)$$

$$a_i \leq u_i(t) \leq b_i, \text{ for a.a. } t \in (0, T), \quad i = 1, \dots, m \quad (159)$$

$$\varphi_i(x(0), x(T), r(0)) \leq 0, \text{ for } i = 1, \dots, d_\varphi, \quad (160)$$

$$\eta_j(x(0), x(T), r(0)) = 0, \text{ for } j = 1 \dots, d_\eta, \quad (161)$$

where  $u \in \mathcal{U}$ ,  $x \in \mathcal{X}$ ,  $r \in \mathbb{R}^{n_r}$  is a parameter considered as a state variable with zero-dynamics,  $a, b \in \mathbb{R}^m$ , functions  $f_i : \mathbb{R}^{n+n_r} \rightarrow \mathbb{R}^n$ ,  $\varphi_i : \mathbb{R}^{2n+n_r} \rightarrow \mathbb{R}$ , and  $\eta : \mathbb{R}^{2n+n_r} \rightarrow \mathbb{R}^{d_\eta}$  are twice continuously differentiable. As  $r$  has zero dynamics, the costate variable  $\psi_r$  corresponding to equation (158) does not appear in the pre-Hamiltonian. Denote with  $\psi$  the costate variable associated with (157). The pre-Hamiltonian function for problem (156)-(161) is:

$$H[\lambda](x, r, u, t) = \psi(t) \sum_{i=0}^m u_i f_i(x, r). \quad (162)$$

Let  $(\hat{x}, \hat{r}, \hat{u})$  be a feasible solution for (157)-(161). Since  $\hat{r}(\cdot)$  is constant, we can denote it with  $\hat{r}$ . Assume that

$$\varphi_i(\hat{x}(0), \hat{x}(T), \hat{r}) = 0, \quad \text{for } i = 0, \dots, d_\varphi. \quad (163)$$

An element  $\lambda = (\alpha, \beta, \psi_x, \psi_r) \in \mathbb{R}^{d_\varphi+d_\eta+1} \times W_\infty^1(0, T; \mathbb{R}^{n,*}) \times W_\infty^1(0, T; \mathbb{R}^{n_r,*})$  is a Pontryagin multiplier for  $(\hat{x}, \hat{r}, \hat{u})$  if it satisfies (7), (8), the costate equation for  $\psi$ :

$$\begin{cases} -\dot{\psi}_x(t) = H_x[\lambda](\hat{x}(t), \hat{r}, \hat{u}(t), t), \text{ a.e. on } (0, T) \\ \psi_x(0) = -\ell_{x_0}[\lambda](\hat{x}(0), \hat{x}(T), \hat{r}), \\ \psi_x(T) = \ell_{x_T}[\lambda](\hat{x}(0), \hat{x}(T), \hat{r}), \end{cases} \quad (164)$$

and for  $\psi_r$ :

$$\begin{cases} -\dot{\psi}_r(t) = H_r[\lambda](\hat{x}(t), \hat{r}, \hat{u}(t), t), \text{ a.e. on } (0, T) \\ \psi_r(0) = -\ell_r[\lambda](\hat{x}(0), \hat{x}(T), \hat{r}), \quad \psi_r(T) = 0. \end{cases} \quad (165)$$

Observe that (165) implies the stationarity condition

$$\ell_r(\hat{x}(0), \hat{x}(T), \hat{r}) + \int_0^T H_r[\lambda](t) dt = 0. \quad (166)$$

Take  $v \in \mathcal{U}$  and consider the linearized state equation:

$$\begin{cases} \dot{z}(t) = \sum_{i=0}^m \hat{u}_i(t)[f_{i,x}(\hat{x}(t), \hat{r})z(t) + f_{i,r}(\hat{x}(t), \hat{r})\delta r(t)] + \sum_{i=1}^m v_i(t)f_i(\hat{x}(t), \hat{r}), \\ \dot{\delta r}(t) = 0, \end{cases} \quad (167)$$

where we can see that  $\delta r(\cdot)$  is constant and thus we denote it with  $\delta r$ . Let the linearized initial-final constraints be:

$$\begin{aligned} \varphi'_i(\hat{x}(0), \hat{x}(T), \hat{r})(z(0), z(T), \delta r) &\leq 0, \quad \text{for } i = 1, \dots, d_\varphi, \\ \eta'_j(\hat{x}(0), \hat{x}(T), \hat{r})(z(0), z(T), \delta r) &= 0, \quad \text{for } j = 1, \dots, d_\eta. \end{aligned} \quad (168)$$

Define for each  $i = 1, \dots, m$  the sets:

$$I_a^i := \{t \in [0, T] : \max_{\lambda \in \Lambda} H_{u_i}[\lambda](t) > 0\}, \quad (169)$$

$$I_b^i := \{t \in [0, T] : \max_{\lambda \in \Lambda} H_{u_i}[\lambda](t) < 0\}, \quad (170)$$

$$I_{\text{sing}}^i := [0, T] \setminus (I_a^i \cup I_b^i). \quad (171)$$

*Assumption 6.1.* Consider the natural extension of Assumption 2.7, i.e., for each  $i = 1, \dots, m$ , the sets  $I_a^i$  and  $I_b^i$  are finite unions of intervals, i.e.,

$$I_a^i = \bigcup_{j=1}^{N_a^i} I_{j,a}^i, \quad I_b^i = \bigcup_{j=1}^{N_b^i} I_{j,b}^i \quad (172)$$

for  $I_{j,a}^i$  and  $I_{j,b}^i$  subintervals of  $[0, T]$  of the form  $[0, c)$ ,  $(d, T]$ ; or  $(c, d)$  if  $c \neq 0$  and  $d \neq T$ . Notice that  $I_a^i \cap I_b^i = \emptyset$ . Call  $c_{1,a}^i < d_{1,a}^i < c_{2,a}^i < \dots < c_{N_a^i,a}^i < d_{N_a^i,a}^i$  the endpoints of these intervals corresponding to bound  $a$ , and define them analogously for  $b$ . Consequently,  $I_{\text{sing}}^i$  is a finite union of intervals as well. Assume that a concatenation of a bang arc followed by another bang arc is forbidden.

*Assumption 6.2.* Strict complementarity assumption for control constraints:

$$\begin{cases} I_a^i = \{t \in [0, T] : \hat{u}_i(t) = a_i\}, \text{ up to a set of null measure,} \\ I_b^i = \{t \in [0, T] : \hat{u}_i(t) = b_i\}, \text{ up to a set of null measure.} \end{cases} \quad (173)$$

Consider

$$\mathcal{C}_2 := \left\{ (z, \delta r, v) \in \mathcal{X}_2 \times \mathbb{R}^{n_r} \times \mathcal{U}_2 : \begin{aligned} &(167)-(168) \text{ hold,} \\ &v_i = 0 \text{ on } I_a^i \cup I_b^i, \text{ for } i = 1, \dots, m \end{aligned} \right\}. \quad (174)$$

The Goh transformation allows to obtain variables  $(\xi, y)$  defined by

$$y(t) := \int_0^t v(s)ds, \quad \xi := z - \sum_{i=1}^m y_i f_i. \quad (175)$$

Notice that  $\xi$  satisfies equation

$$\begin{aligned} \dot{\xi} &= A^x \xi + A^r \delta r + B_1^x y, \\ \xi(0) &= z(0), \end{aligned} \quad (176)$$

where, denoting  $[f_i, f_j]^x := f_{i,x}f_j - f_{j,x}f_i$ ,

$$A^x := \sum_{i=0}^m \hat{u}_i f_{i,x}, \quad A^r := \sum_{i=0}^m \hat{u}_i f_{i,r}, \quad B_1^x y := \sum_{j=1}^m y_j \sum_{i=0}^m \hat{u}_i [f_i, f_j]^x. \quad (177)$$

Consider the trasformed (168):

$$\begin{aligned} \varphi'_i(\hat{x}(0), \hat{x}(T), \hat{r})(\xi(0), \xi(T) + B(T)h, \delta r) &\leq 0, \quad i = 1, \dots, d_\varphi, \\ \eta'_j(\hat{x}(0), \hat{x}(T), \hat{r})(\xi(0), \xi(T) + B(T)h, \delta r) &= 0, \quad j = 1, \dots, d_\eta. \end{aligned} \quad (178)$$

Take the cone

$$\mathcal{P} := \left\{ \begin{array}{l} (\xi, \delta r, y, h) \in \mathcal{X} \times \mathbb{R}^{nr} \times \mathcal{Y} \times \mathbb{R}^m : y(0) = 0, \quad h = y(T), \\ (176) \text{ and } (178) \text{ hold, } y'_i = 0 \text{ on } I_a^i \cup I_b^i, \text{ for } i = 1, \dots, m \end{array} \right\}.$$

Observe that each  $(\xi, \delta r, y, h) \in \mathcal{P}$  satisfies:

$$y_i \text{ constant over each composing interval of } I_a^i \cup I_b^i, \quad (179)$$

and on the endpoints:

$$\begin{cases} y_i = 0 \text{ on } [0, d], \text{ if } 0 \in I_a^i \cup I_b^i, \text{ and,} \\ y_i = h_i \text{ on } [c, T], \text{ if } T \in I_a^i \cup I_b^i, \end{cases} \quad (180)$$

where  $[0, d]$  is the first maximal composing interval of  $I_a^i \cup I_b^i$  when  $0 \in I_a^i \cup I_b^i$ , and  $[c, T]$  is its last composing interval when  $T \in I_a^i \cup I_b^i$ . Define

$$\mathcal{P}_2 := \left\{ \begin{array}{l} (\xi, \delta r, y, h) \in \mathcal{X}_2 \times \mathbb{R}^{nr} \times \mathcal{U}_2 \times \mathbb{R}^m : \\ (176), (178), (179) \text{ and } (180) \text{ hold for } i = 1, \dots, m \end{array} \right\}. \quad (181)$$

Recall definitions in equations (59), (60), (63), (64), (65). Minor simplifications appear in the computations of these functions as dynamics of  $r$  is null and  $\delta r$  is constant. We will detail these calculations for an example.

Consider  $M \subset \mathbb{R}^s$  and the subset of  $M \subset \mathbb{R}^s$  defined by

$$G(M) := \{\lambda \in M : V_{ij}[\lambda] = 0 \text{ on } I_{\text{sing}}^i \cap I_{\text{sing}}^j, \text{ for every pair } 1 < i \neq j \leq m\}.$$

Using the same techniques we obtain the equivalent of Theorem 4.18:

**Corollary 6.3.** *Suppose that  $(\hat{x}, \hat{r}, \hat{u})$  is a weak minimum for problem (156)-(161). Then*

$$\max_{\lambda \in G(\text{co}\Lambda)} \Omega_{\mathcal{P}_2}[\lambda](\xi, \delta r, y, h) \geq 0, \quad \text{for all } (\xi, \delta r, y, h) \in \mathcal{P}_2. \quad (182)$$

By a simple adaptation of the proof of Theorem 5.5 we get the equivalent result:

**Corollary 6.4.** *Let  $m = 1$ . Suppose that there exists  $\rho > 0$  such that*

$$\max_{\lambda \in \Lambda} \Omega_{\mathcal{P}_2}[\lambda](\xi, \delta r, y, h) \geq \rho \gamma(y, h), \quad \text{for all } (\xi, \delta r, y, h) \in \mathcal{P}_2. \quad (183)$$

*Then  $(\hat{x}, \hat{r}, \hat{u})$  is a Pontryagin minimum that satisfies  $\gamma$ -quadratic growth.*

## 6.2 Application to minimum-time problem

Consider the problem

$$\begin{aligned} J &:= T \rightarrow \min, \\ \text{s.t.} & \text{ (157) - (161)}. \end{aligned} \quad (184)$$

Observe that by the change of variables:

$$x(s) \leftarrow x(Ts), \quad u(s) \leftarrow u(Ts), \quad (185)$$

we can transform the problem into formulation:

$$J := T(0) \rightarrow \min, \quad (186)$$

$$\dot{x}(s) = T(s) \sum_{i=0}^m u_i(s) f_i(x(s), r(s)), \quad \text{a.e. on } (0, 1), \quad (187)$$

$$\dot{r}(s) = 0, \quad \text{a.e. on } (0, 1), \quad (188)$$

$$\dot{T}(s) = 0, \quad \text{a.e. on } (0, 1), \quad (189)$$

$$a_i \leq u_i(s) \leq b_i, \quad \text{a.e. on } (0, 1), \quad i = 1, \dots, m \quad (190)$$

$$\varphi_i(x(0), x(1), r(0)) \leq 0, \quad \text{for } i = 1, \dots, d_\varphi, \quad (191)$$

$$\eta_j(x(0), x(T), r(0)) = 0, \quad \text{for } j = 1 \dots, d_\eta. \quad (192)$$

To the problem written in this form we can apply Corollaries 6.3 and 6.4. We will detail the calculations in the following example:

### 6.2.1 Example: Markov-Dubins problem

Consider the problem in the interval  $[0, T]$  with free final time  $T$  :

$$\begin{aligned} J &:= T \rightarrow \min, \\ \dot{x}_1 &= -\sin x_3, \quad x_1(0) = 0, \quad x_1(T) = b_1, \\ \dot{x}_2 &= \cos x_3, \quad x_2(0) = 0, \quad x_2(T) = b_2, \\ \dot{x}_3 &= u, \quad x_3(0) = 0, \quad x_3(T) = \theta, \\ &-1 \leq u \leq 1, \end{aligned} \quad (193)$$

with  $0 < \theta < \pi$ ,  $b_1$  and  $b_2$  fixed. We will study the optimality of the extremal

$$\hat{u}(t) := \begin{cases} 1 & \text{on } [0, \theta], \\ 0 & \text{on } (\theta, \hat{T}]. \end{cases} \quad (194)$$

Observe that by the change of variables (185) we can transform (193) into the following problem on the interval  $[0, 1]$  :

$$\begin{aligned} J &:= T(0) \rightarrow \min, \\ \dot{x}_1(s) &= -T(s) \sin x_3(s), \quad x_1(0) = 0, \quad x_1(1) = b_1, \\ \dot{x}_2(s) &= T(s) \cos x_3(s), \quad x_2(0) = 0, \quad x_2(1) = b_2, \\ \dot{x}_3(s) &= T(s)u(s), \quad x_3(0) = 0, \quad x_3(1) = \theta, \\ \dot{T}(s) &= 0, \\ &-1 \leq u(s) \leq 1. \end{aligned} \quad (195)$$



We get for state variables:

$$\hat{x}_3(s) = \begin{cases} \hat{T}s & \text{on } [0, \theta/\hat{T}], \\ \theta & \text{on } (\theta/\hat{T}, 1], \end{cases} \quad (196)$$

$$\hat{x}_1(s) = \begin{cases} \cos(\hat{T}s) - 1 & \text{on } [0, \theta/\hat{T}], \\ \hat{T} \sin \theta (\theta/\hat{T} - s) + \cos \theta - 1 & \text{on } (\theta/\hat{T}, 1], \end{cases} \quad (197)$$

$$\hat{x}_2(s) = \begin{cases} \sin \hat{T}s & \text{on } [0, \theta/\hat{T}], \\ \hat{T} \cos \theta (s - \theta/\hat{T}) + \sin \theta & \text{on } (\theta/\hat{T}, 1]. \end{cases} \quad (198)$$

As terminal values for  $x_1$  and  $x_2$  are fixed, final time  $\hat{T}$  is determined by the previous equalities. The pre-Hamiltonian for problem (195) is

$$H[\lambda](s) := T(s)(-\psi_1(s) \sin x_3(s) + \psi_2(s) \cos x_3(s) + \psi_3(s)u(s)). \quad (199)$$

The final Lagrangian is

$$\ell := \alpha_0 T(1) + \sum_{j=1}^3 (\beta^j x_j(0) + \beta_j x_j(1)). \quad (200)$$

As  $\dot{\psi}_1 \equiv 0$ , and  $\dot{\psi}_2 \equiv 0$ , we get

$$\psi_1 \equiv \beta_1, \quad \psi_2 \equiv \beta_2, \quad \text{on } [0, 1]. \quad (201)$$

Since on  $[\theta/\hat{T}, 1]$ , candidate control  $\hat{u}$  is singular, we have  $H_u[\lambda] \equiv 0$ , and then by (199) we get

$$\psi_3(s) = 0, \quad \text{on } [\theta/\hat{T}, 1]. \quad (202)$$

Thus  $\beta_3 = 0$ . Besides, as costate equation for  $\psi_3$  is

$$-\dot{\psi}_3 = \hat{T}(-\beta_1 \cos \hat{x}_3 - \beta_2 \sin \hat{x}_3), \quad (203)$$

by (196) and (202) we get

$$\beta_1 \cos \theta + \beta_2 \sin \theta = 0. \quad (204)$$

From (196) and (202) and since  $H$  is constantly equal to  $-\alpha_0$ , we get

$$H = \hat{T}(-\beta_1 \sin \theta + \beta_2 \cos \theta) \equiv -\alpha_0. \quad (205)$$

**Proposition 6.5.** *The following properties hold*

- (i)  $\alpha_0 > 0$ ,
- (ii)  $H_u[\lambda](s) < 0$ , on  $[0, \theta/\hat{T}]$  for all  $\lambda \in \Lambda$ .

*Proof.* **Item (i)** Suppose that  $\alpha_0 = 0$ . By (204) and (205) we obtain

$$\beta_1 \cos \theta + \beta_2 \sin \theta = 0, \quad \text{and} \quad -\beta_1 \sin \theta + \beta_2 \cos \theta = 0. \quad (206)$$

Suppose, w.l.g., that  $\cos \theta \neq 0$ , then  $\beta_1 = -\beta_2 \frac{\sin \theta}{\cos \theta}$  and thus

$$\beta_2 \frac{\sin^2 \theta}{\cos \theta} + \beta_2 \cos \theta = 0.$$

Conclude that  $\beta_2 = 0$  as well. This implies  $(\alpha_0, \beta_1, \beta_2, \beta_3) = 0$ , that is in contradiction with the non-triviality condition (7). So,  $\alpha_0 > 0$ , as it was to be proved.

**Item (ii)** Observe that

$$H_u[\lambda](s) \leq 0, \quad \text{on } [0, \theta/\hat{T}], \quad (207)$$

and  $H_u[\lambda] = \psi_3$ . Let us prove that  $\psi_3$  is never 0 on  $[0, \theta/\hat{T}]$ . Suppose there exists  $s_1 \in [0, \theta/\hat{T}]$  such that  $\psi_3(s_1) = 0$ . Thus, since  $\psi_3(\theta/\hat{T}) = 0$  by (202), there exists  $s_2 \in (s_1, \theta/\hat{T})$  such that  $\psi_3(s_2) = 0$ , i.e.,

$$\beta_1 \cos(\hat{T}s_2) + \beta_2 \sin(\hat{T}s_2) = 0. \quad (208)$$

Equations (204) and (208) imply that  $\tan(\theta/\hat{T}) = \tan(s_2/\hat{T})$ . This contradicts  $\theta < \pi$ . Thus  $\psi_3(s) \neq 0$  for every  $s \in [0, \theta/\hat{T}]$ , and consequently

$$H_u[\lambda](s) < 0, \quad \text{for } s \in [0, \theta/\hat{T}]. \quad (209)$$

□

Since  $\alpha_0 > 0$ , then  $\delta T = 0$  for each element of the critical cone, where  $\delta T$  is the linearized state variable  $T$ . Observe that as  $\hat{u} = 1$  on  $[0, \theta/\hat{T}]$ , then

$$y = 0 \text{ and } \xi = 0, \text{ on } [0, \theta/\hat{T}], \text{ for all } (\xi, \delta T, y, h) \in \mathcal{P}_2. \quad (210)$$

We look for the second variation in the interval  $[\theta/\hat{T}, 1]$ . Goh transformation gives

$$\xi_3 = z_3 - \hat{T}y, \quad (211)$$

and since  $\dot{z}_3 = \hat{T}v$ , we get  $z_3 = \hat{T}y$  and thus  $\xi_3 = 0$ . Then, as  $H_{ux} = 0$  and  $\ell'' = 0$ , we get

$$\Omega[\lambda] = \int_{\theta/\hat{T}}^1 (\beta_1 \sin \theta - \beta_2 \cos \theta) y^2 dt = \alpha_0 \int_0^1 y^2 dt. \quad (212)$$

Notice that if  $(\xi, \delta T, y, h) \in \mathcal{P}_2$  then  $h$  satisfies  $\xi_3(T) + \hat{T}h = 0$ , and as  $\xi_3(T) = 0$ , we get  $h = 0$ . Thus

$$\Omega[\lambda](\xi, y, h) = \alpha_0 \int_0^T y^2 dt = \alpha_0 \gamma(y, h), \quad \text{on } \mathcal{P}_2. \quad (213)$$

Since Assumptions 6.1 and 6.2 hold, we conclude by Corollary 6.4 that  $(\hat{x}, \hat{T}, \hat{u})$  is a Pontryagin minimum satisfying quadratic growth.

## 7 Appendix

**Lemma 7.1.** *Let*

$$X := \{(\xi, y, h) \in \mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m : (51), (56)-(57) \text{ hold}\},$$

$$L := \{(\xi, y, y(T)) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m : y(0) = 0, (51) \text{ and } (56)\}.$$

*Then  $L$  is a dense subset of  $X$  in the  $\mathcal{X}_2 \times \mathcal{U}_2 \times \mathbb{R}^m$ -topology.*

*Proof.* (See Lemma 6 in [10].) Let us prove the result for  $m = 1$ . The general case is a trivial extension. Let  $(\bar{\xi}, \bar{y}, \bar{h}) \in X$  and  $\varepsilon, \delta > 0$ . Consider  $\phi \in \mathcal{Y}$  such that  $\|\bar{y} - \phi\|_2 < \varepsilon/2$ . In order to satisfy condition (57) take

$$\begin{cases} y_\delta(t) := 0, & \text{for } t \in [0, d_1], & \text{if } c_1 = 0, \\ y_\delta(t) := h, & \text{for } t \in [c_N, T], & \text{if } d_N = T, \end{cases} \quad (214)$$

where  $c_j, d_j$  were introduced in Assumption 2.7. Since  $\bar{y}$  is constant on each  $I_j$ , define  $y_\delta$  constant over these intervals with the same constant value as  $\bar{y}$ . It remains to define  $y_\delta$  over  $I_+$ . Over each maximal composing interval  $(a, b)$  of  $I_+$ , define  $y_\delta$  as it is described next. Take  $c := \bar{y}(a-)$  if  $a > 0$ , or  $c := 0$  if  $a = 0$ ; and let  $d := \bar{y}(b+)$  if  $b < T$ , or  $d := h$  when  $b = T$ . Define two affine functions  $\ell_{1,\delta}$  and  $\ell_{2,\delta}$  satisfying

$$\begin{aligned} \ell_{1,\delta}(a) &= c, & \ell_{1,\delta}(a + \delta) &= \phi(a + \delta), \\ \ell_{2,\delta}(b) &= d, & \ell_{2,\delta}(b - \delta) &= \phi(b - \delta). \end{aligned} \quad (215)$$

Take

$$y_\delta(t) := \begin{cases} \ell_{1,\delta}(t), & \text{for } t \in [a, a + \delta], \\ \phi(t), & \text{for } t \in (a + \delta, b - \delta), \\ \ell_{2,\delta}(t), & \text{for } t \in [b - \delta, b], \end{cases} \quad (216)$$

and notice that  $\|\phi - y_\delta\|_{2,[a,b]} \leq \frac{1}{k} \max(|c|, |d|, M)$ , where  $M := \sup_{t \in [a,b]} |\phi(t)|$ . Finally, observe that  $y_\delta(T) = h$ , and for sufficiently small  $\delta$ ,

$$\|\bar{y} - y_\delta\|_2 \leq \|\bar{y} - \phi\|_2 + \|\phi - y_\delta\|_2 < \varepsilon. \quad (217)$$

Thus, the result follows.  $\square$

**Lemma 7.2.** *Let  $\lambda \in \Lambda$  and  $(z, v) \in \mathcal{C}_2$ . Then*

$$\sum_{i=0}^{d_\varphi} \alpha_i \bar{\varphi}_i''(\hat{u})(v, v) + \sum_{i=1}^{d_\eta} \beta_i \bar{\eta}_i''(\hat{u})(v, v) = \Omega[\lambda](z, v). \quad (218)$$

*Proof.* Let us compute the left-hand side of (218). Notice that

$$\sum_{i=0}^{d_\varphi} \alpha_i \bar{\varphi}_i(\hat{u}) + \sum_{i=1}^{d_\eta} \beta_i \bar{\eta}_i(\hat{u}) = \ell[\lambda](\hat{x}(T)). \quad (219)$$

Let us look for a second order expansion for  $\ell$ . Consider first a second order expansion of the state variable:

$$x = \hat{x} + z + \frac{1}{2} z_{vv} + o(\|v\|_\infty^2), \quad (220)$$

where  $z_{vv}$  satisfies

$$\dot{z}_{vv} = Az_{vv} + D_{(x,u)^2}^2 F(\hat{x}, \hat{u})(z, v)^2, \quad z_{vv}(0) = 0, \quad (221)$$

with  $F(x, u) := \sum_{i=0}^m u_i f_i(x)$ . Consider the second order expansion for  $\ell$ :

$$\begin{aligned} \ell[\lambda](x(T)) &= \ell[\lambda](\hat{x} + z + \frac{1}{2} z_{vv})(T) + o(\|v\|_1^2) \\ &= \ell[\lambda](\hat{x}(T)) + \ell'[\lambda](\hat{x}(T))(z(T) + \frac{1}{2} z_{vv}(T)) \\ &\quad + \frac{1}{2} \ell''[\lambda](\hat{x}(T))(z(T) + \frac{1}{2} z_{vv}(T))^2 + o(\|v\|_1^2). \end{aligned} \quad (222)$$

**Step 1.** Compute

$$\begin{aligned} \ell'[\lambda](\hat{x}(T))z_{vv}(T) &= \psi(T)z_{vv}(T) - \psi(0)z_{vv}(0) \\ &= \int_0^T [\dot{\psi}z_{vv} + \psi\dot{z}_{vv}]dt = \int_0^T \{-\psi Az_{vv} + \psi(Az_{vv} + D^2F_{(x,u)^2}(z, v)^2)\}dt \\ &= \int_0^T D^2H[\lambda](z, v)^2dt. \end{aligned}$$

**Step 2.** Compute  $\ell''[\lambda](\hat{x}(T))(z(T), z_{vv}(T))$ . Applying Gronwall's Lemma we get  $\|z\|_\infty = O(\|v\|_1)$ , and  $\|z_{vv}\|_\infty = O(\|v^2\|_1)$ . Thus

$$|(z(T), z_{vv}(T))| = O(\|v\|_1^3), \quad (223)$$

and we conclude that

$$|\ell''[\lambda](\hat{x}(T))(z(T), z_{vv}(T))| = O(\|v\|_1^3). \quad (224)$$

**Step 3.** See that  $\ell''[\lambda](\hat{x}(T))(z_{vv}(T))^2 = O(\|v\|_1^4)$ . Then we get by (222),

$$\begin{aligned} \ell[\lambda](x(T)) &= \ell[\lambda](\hat{x}(T)) + \ell'[\lambda](\hat{x}(T))z(T) \\ &\quad + \frac{1}{2}\ell''[\lambda](\hat{x}(T))z^2(T) + \frac{1}{2}\int_0^T D_{(x,u)^2}^2H[\lambda](z, v)^2dt + o(\|v\|_1^2) \\ &= \ell[\lambda](\hat{x}(T)) + \ell'[\lambda](\hat{x}(T))z(T) + \Omega[\lambda](z, v) + o(\|v\|_1^2). \end{aligned}$$

Conclude by (219).  $\square$

**Lemma 7.3.** Given  $(z, v) \in \mathcal{W}$  satisfying (12), the following estimation holds for some  $\rho > 0$ :

$$\|z\|_2^2 + |z(T)|^2 \leq \rho\gamma(y, y(T)), \quad (225)$$

where  $y$  is defined by (50).

*Remark 7.4.*  $\rho$  depends on  $\hat{w}$ , i.e., it does not vary for different  $(z, v)$ .

*Proof.* Every time we put  $\rho_i$  we refer to a constant depending on  $\|A\|_\infty$ ,  $\|B\|_\infty$  or both. Consider  $\xi$  the solution of equation (51) corresponding to  $y$ . Gronwall's Lemma and Cauchy-Schwartz inequality imply

$$\|\xi\|_\infty \leq \rho_1\|y\|_2. \quad (226)$$

This last inequality together with expression (50) imply

$$\|z\|_2 \leq \|\xi\|_2 + \|B\|_\infty\|y\|_2 \leq \rho_2\|y\|_2. \quad (227)$$

On the other hand, equations (50) and (226) lead to

$$|z(T)| \leq |\xi(T)| + \|B\|_\infty|y(T)| \leq \rho_1\|y\|_2 + \|B\|_\infty|y(T)|. \quad (228)$$

Then by inequality ' $ab \leq \frac{a^2+b^2}{2}$ ' we get

$$|z(T)|^2 \leq \rho_3(\|y\|_2^2 + |y(T)|^2). \quad (229)$$

Conclude by equations (227) and (229).  $\square$

The next lemma is a generalization of previous result to the nonlinear case. See Lemma 6.1 in Dmitruk [8].

**Lemma 7.5.** *Let  $w = (x, u)$  solution of (2) with  $\|u\|_2 \leq c$  for some constant  $c$ . Put  $(\delta x, v) := w - \hat{w}$ . Then*

$$|\delta x(T)|^2 + \|\delta x\|_2^2 \leq \rho\gamma(y, y(T)), \quad (230)$$

where  $y$  is defined by (50) and  $\rho$  depends on  $c$ .

**Lemma 7.6.** *Let  $\{y_k\} \subset L_2(a, b)$  be a sequence of continuous non-decreasing functions that converges weakly to  $y \in L_2(a, b)$ . Then  $y$  is non-decreasing.*

*Proof.* Let  $s, t \in (a, b)$  such that  $s < t$ , and  $\varepsilon_0 > 0$  such that  $s + \varepsilon_0 < t - \varepsilon_0$ . For every  $k \in \mathbb{N}$ , and every  $0 < \varepsilon < \varepsilon_0$  holds:

$$\int_{s-\varepsilon}^{s+\varepsilon} y_k(\nu) d\nu \leq \int_{t-\varepsilon}^{t+\varepsilon} y_k(\nu) d\nu. \quad (231)$$

Taking limit when  $k$  goes to infinity and multiplying by  $\frac{1}{2\varepsilon}$  deduce that

$$\frac{1}{2\varepsilon} \int_{s-\varepsilon}^{s+\varepsilon} y(\nu) d\nu \leq \frac{1}{2\varepsilon} \int_{t-\varepsilon}^{t+\varepsilon} y(\nu) d\nu. \quad (232)$$

As  $(a, b)$  is a finite measure space,  $y$  is a function of  $L_1(a, b)$ , and almost all points in  $(a, b)$  are Lebesgue points (see Rudin [39] - Theorem 7.7). Thus, taking  $\varepsilon$  to 0 follows from the previous inequality that

$$y(s) \leq y(t), \quad (233)$$

as we wanted to prove.  $\square$

**Lemma 7.7.** *Consider a sequence  $\{y_k\}$  of non-decreasing continuous functions in a compact real interval  $I$ , and assume that  $\{y_k\}$  converges weakly to 0 in  $\mathcal{U}_2$ . Then it converges uniformly to 0 on any interval  $(a, b) \subset I$ .*

*Proof.* Take an arbitrary interval  $(a, b) \subset I$ . First prove the pointwise convergence of  $\{y_k\}$  to 0. By *reductio ad absurdum* suppose that there exists  $c \in (a, b)$  such that  $\{y_k(c)\}$  does not converge to 0. Thus there exist  $\varepsilon > 0$  and a subsequence  $\{y_{k_j}\}$  such that  $y_{k_j}(c) > \varepsilon$  for each  $j \in \mathbb{N}$ , or  $y_{k_j}(c) < -\varepsilon$  for each  $j \in \mathbb{N}$ . Suppose, without loss of generality, the first statement. Thus

$$0 < \varepsilon(b - c) < y_{k_j}(c)(b - c) \leq \int_c^b y_{k_j}(t) dt, \quad (234)$$

where the last inequality stands since  $y_{k_j}$  is nondecreasing. But the right-hand side of (234) goes to 0 as  $j$  goes to infinity. This contradicts the hypothesis, and thus the pointwise convergence of  $\{y_k\}$  to 0 follows. The uniform convergence is a direct consequence of the monotonicity of functions  $y_k$ .  $\square$

**Lemma 7.8.** [12, Theorem 22, Page 154 - Volume I] *Let  $a$  and  $b$  be two functions of bounded variation in  $[0, T]$ . Suppose that one is continuous and the other is right-continuous. Then*

$$\int_0^T a(t) db(t) + \int_0^T b(t) da(t) = [ab]_0^T. \quad (235)$$

**Lemma 7.9.** *Let  $m = 1$ , i.e., consider a scalar control variable. Then, for any  $\lambda \in \Lambda$  the function  $R[\lambda](t)$  defined in (63) is continuous in  $t$ .*

*Proof.* Consider definition (60). Condition  $V[\lambda] \equiv 0$  yields  $S[\lambda] = C[\lambda]B$ , and since  $R[\lambda]$  is scalar we can write

$$R[\lambda] = B^\top Q[\lambda]B - 2C[\lambda]B_1 - \dot{C}[\lambda]B - C[\lambda]\dot{B}.$$

Note that  $B = f_1$ ,  $B_1 = [f_0, f_1]$ ,  $C[\lambda] = -\psi f_1'$ , and  $Q[\lambda] = -\psi(f_0'' + \hat{u}f_1'')$ , thus

$$\begin{aligned} R[\lambda] &= \psi(f_0'' + \hat{u}f_1'')(f_1, f_1) - 2\psi f_1'(f_0'f_1 - f_1'f_0) \\ &\quad + \psi(f_0' + \hat{u}f_1')f_1'f_1 - \psi f_1''(f_0 + \hat{u}f_1)f_1 - \psi f_1'f_1'(f_0 + \hat{u}f_1) \\ &= \psi[f_1, [f_1, f_0]]. \end{aligned}$$

Since  $f_0$  and  $f_1$  are twice continuously differentiable, we conclude that  $R[\lambda]$  is continuous in time.  $\square$

**Lemma 7.10.** [19] *Consider a quadratic form  $\mathcal{Q} = \mathcal{Q}_1 + \mathcal{Q}_2$  where  $\mathcal{Q}_1$  is a Legendre form and  $\mathcal{Q}_2$  is weakly continuous over some Hilbert space. Then  $\mathcal{Q}$  is a Legendre form.*

**Lemma 7.11.** [19, Theorem 3.2] *Consider a real interval  $I$  and a quadratic form  $\mathcal{Q}$  over the Hilbert space  $L_2(I)$ , given by*

$$\mathcal{Q}(y) := \int_I y^\top(t)R(t)y(t)dt.$$

*Then  $\mathcal{Q}$  is weakly l.s.c. over  $L_2(I)$  iff*

$$R(t) \succeq 0, \quad \text{a.e. on } I. \quad (236)$$

**Lemma 7.12.** [7, Theorem 5] *Given a Hilbert space  $H$ , and  $a_1, a_2, \dots, a_p \in H$ , set*

$$K := \{x \in H : (a_i, x) \leq 0, \text{ for } i = 1, \dots, p\}. \quad (237)$$

*Let  $M$  be a convex and compact subset of  $\mathbb{R}^s$ , and let  $\{Q^\psi : \psi \in M\}$  be a family of continuous quadratic forms over  $H$ , the mapping  $\psi \rightarrow Q^\psi$  being affine. Set  $M_\# := \{\psi \in M : Q^\psi \text{ is weakly l.s.c.}\}$  and assume that*

$$\max_{\psi \in M} Q^\psi(x) \geq 0, \quad \text{for all } x \in K. \quad (238)$$

*Then*

$$\max_{\psi \in M_\#} Q^\psi(x) \geq 0, \quad \text{for all } x \in K. \quad (239)$$

The following result is an adaptation of Lemma 6.5 in [8]:

**Lemma 7.13.** *Consider a sequence  $\{v_k\} \subset \mathcal{U}$  and  $\{y_k\}$  their primitives defined by (50). Call  $u_k := \hat{u} + v_k$ ,  $x_k$  its corresponding solution of (2), and  $z_k$  the linearized state corresponding to  $v_k$ , i.e., solution of (12). Define for each  $k \in \mathbb{N}$ ,*

$$\delta x_k := x_k - \hat{x}, \quad \eta_k := \delta x_k - z_k, \quad \gamma_k := \gamma(y_k, y_k(T)). \quad (240)$$

*Suppose that  $\{v_k\}$  converges to 0 in the Pontryagin sense. Then*

(i)

$$\dot{\eta}_k = \sum_{i=0}^m \hat{u}_i f'_i(\hat{x}) \eta_k + \sum_{i=1}^m v_{i,k} f'_i(\hat{x}) \delta x_k + \zeta_k, \quad (241)$$

$$\dot{\delta x}_k = \sum_{i=0}^m u_{i,k} f'_i(\hat{x}) \delta x_k + \sum_{i=1}^m v_{i,k} f_i(\hat{x}) + \zeta_k, \quad (242)$$

where  $\|\zeta_k\|_2 \leq o(\sqrt{\gamma_k})$ , and  $\|\zeta_k\|_\infty \rightarrow 0$ ,

(ii)  $\|\eta_k\|_\infty \leq o(\sqrt{\gamma_k})$ .

*Proof.* (i,ii) Consider the Taylor expansions of second order of  $f_i$  :

$$f_i(x_k) = f_i(\hat{x}) + f'_i(\hat{x}) \delta x_k + \frac{1}{2} f''_i(\hat{x}) (\delta x_k, \delta x_k) + o(\|\delta x_k\|_\infty^2).$$

We can write

$$\dot{\delta x}_k = \sum_{i=0}^m u_{i,k} f'_i(\hat{x}) \delta x_k + \sum_{i=1}^m v_{i,k} f_i(\hat{x}) + \zeta_k, \quad (243)$$

with

$$\zeta_k = \frac{1}{2} \sum_{i=0}^m u_{i,k} f''_i(\hat{x}) (\delta x_k, \delta x_k) + o(|\delta x_k(t)|^2) \sum_{i=0}^m u_{i,k}. \quad (244)$$

As  $\{u_k\}$  is bounded in  $L_\infty$ , and  $\|\delta x_k\|_\infty \rightarrow 0$  we get  $\|\zeta_k\|_\infty \rightarrow 0$  and the following  $L_2$ -norm bound:

$$\begin{aligned} \|\zeta_k\|_2 &\leq \text{const.} \sum_{i=0}^m \|u_{i,k} (\delta x_k, \delta x_k)\|_2 + o(\gamma_k) \left\| \sum_{i=0}^m u_{i,k} \right\|_1 \\ &\leq \text{const.} \|u_k\|_\infty \|\delta x_k\|_2^2 = O(\gamma_k) \leq o(\sqrt{\gamma_k}). \end{aligned} \quad (245)$$

Let us look for the differential equation of  $\eta_k$  defined in (240). By (243), and adding and subtracting the term  $\sum_{i=1}^m \hat{u}_i f'_i(\hat{x}) \delta x_k$  we obtain

$$\dot{\eta}_k = \sum_{i=0}^m \hat{u}_i f'_i(\hat{x}) \eta_k + \sum_{i=1}^m v_{i,k} f'_i(\hat{x}) \delta x_k + \zeta_k,$$

Thus we obtain (i). Applying Gronwall's Lemma to this last differential equation we get

$$\|\eta_k\|_\infty \leq \left\| \sum_{i=1}^m v_{i,k} f'_i(\hat{x}) \delta x_k + \zeta_k \right\|_1. \quad (246)$$

As  $\|v_k\|_\infty < N$  and  $\|v_k\|_1 \rightarrow 0$ , we have also that  $\|v_k\|_2 \rightarrow 0$ . Applying Cauchy-Schwartz inequality to (246), from (245) we get (ii).  $\square$

## References

- [1] A.A. Agrachev and R.V. Gamkrelidze. Second order optimality principle for a time- optimal problem. *Math. USSR, Sbornik*, 100(4), 1976.
- [2] A.A. Agrachev, G. Stefani, and P.L. Zezza. Strong optimality for a bang-bang trajectory. *SIAM J. Control and Optimization*, 41:991–1014, 2002.

- 
- [3] V.M. Alekseev, V.M. Tikhomirov, and S.V. Fomin. *Optimal Control*. Nauka, Moscow, 1979. [in Russian].
- [4] R. Cominetti and J.-Penot. Tangent sets of order one and two to the positive cones of some functional spaces. *Applied Mathematics and Optimization*, (36):291–312, 1997.
- [5] A.V. Dmitruk. Quadratic conditions for a weak minimum for singular regimes in optimal control problems. *Soviet Math. Doklady*, 18(2), 1977.
- [6] A.V. Dmitruk. Quadratic conditions for a Pontryagin minimum in an optimal control problems, linear in the control, with a constraint on the control. *Soviet Math. Doklady*, 28(2):364–368, 1983.
- [7] A.V. Dmitruk. Jacobi-type conditions for the problem of Bolza with inequalities. *Math. Notes*, 35(6):427–435, 1984.
- [8] A.V. Dmitruk. Quadratic order conditions for a Pontryagin minimum in an optimal control problem linear in the control. *Math. USSR Izvestiya*, 28:275–303, 1987.
- [9] A.V. Dmitruk. Jacobi type conditions for singular extremals. *Control & Cybernetics*, 37(2):285–306, 2008.
- [10] A.V. Dmitruk and K.K. Shishov. Analysis of a quadratic functional with a partly singular Legendre condition. *Moscow University Comput. Math. and Cybernetics*, 34(1):16–25, 2010.
- [11] A.Ya. Dubovitskii and A.A. Milyutin. Extremum problems with constraints. *USSR Comp. Math. and Math. Phys.*, 5(3):1–80, 1965.
- [12] N. Dunford and J. Schwartz. *Linear operators, Vol I and II*. Interscience, New York, 1958, 1963.
- [13] U. Felgenhauer. On stability of bang-bang type controls. *SIAM J. Control Optim.*, 41(6):1843–1867 (electronic), 2003.
- [14] U. Felgenhauer. Optimality and sensitivity for semilinear bang-bang type optimal control problems. *Int. J. Appl. Math. Comput. Sci.*, 14(4):447–454, 2004.
- [15] U. Felgenhauer. Optimality properties of controls with bang-bang components in problems with semilinear state equation. *Control Cybernet.*, 34(3):763–785, 2005.
- [16] R. Gabasov and F.M. Kirillova. High-order necessary conditions for optimality. *J. SIAM Control*, 10:127–168, 1972.
- [17] B.S. Goh. Necessary conditions for singular extremals involving multiple control variables. *J. SIAM Control*, 4:716–731, 1966.
- [18] B.S. Goh. The second variation for the singular Bolza problem. *J. SIAM Control*, 4(2):309–325, 1966.
- [19] M.R. Hestenes. Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. *Pacific J. Math.*, 1(4):525–581, 1951.



- [20] A. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards, Section B, Mathematical Sciences*, 49:263–265, 1952.
- [21] D. H. Jacobson and Speyer J.L. Necessary and sufficient conditions for optimality for singular control problems: A limit approach. *J. Math. Anal. Appl.*, 34:239–266, 1971.
- [22] D.H. Jacobson, M.M. Lele, and J.L. Speyer. New necessary conditions of optimality for control problems with state-variable inequality constraints. *Journal of Mathematical Analysis and Applications*, 35:255–284, 1971.
- [23] H.J. Kelley. A second variation test for singular extremals. *AIAA Journal*, 2:1380–1382, 1964.
- [24] H.J. Kelley, R.E. Kopp, and H.G. Moyer. Singular extremals. In *Topics in Optimization*, pages 63–101. Academic Press, New York, 1967.
- [25] R.E. Kopp and H.G. Moyer. Necessary conditions for singular extremals. *AIAA Journal*, 3:1439–1444, 1965.
- [26] A.J. Krener. The high order maximal principle and its application to singular extremals. *SIAM J. on Control*, 15(2):256–293, 1977.
- [27] S. Kurcyusz and J. Zowe. Regularity and stability for the mathematical programming problem in Banach spaces. In *Applied Mathematics and Optimization*, pages 49–62. Springer, New York, 1979.
- [28] E. S. Levitin, A. A. Milyutin, and N. P. Osmolovskii. Higher order conditions for local minima in problems with constraints. *Uspekhi Mat. Nauk*, 33(6(204)):85–148, 272, 1978.
- [29] H. Maurer and N.P. Osmolovskii. Second order optimality conditions for bang-bang control problems. *Control and Cybernetics*, 32:555–584, 2003.
- [30] H. Maurer and N.P. Osmolovskii. Second order sufficient conditions for time-optimal bang-bang control. *SIAM J. Control Optim.*, 42:2239–2263, 2003.
- [31] A.A. Milyutin. On quadratic conditions for an extremum in smooth problems with a finite-dimensional range. *Methods of the Theory of Extremal Problems in Economics*, pages 138–177, 1981.
- [32] A.A. Milyutin and N.P. Osmolovskii. Calculus of variations and optimal control. *American Mathematical Society*, 1998.
- [33] N. P. Osmolovskii. Quadratic extremality conditions for broken extremals in the general problem of the calculus of variations. *J. Math. Sci. (N. Y.)*, 123(3):3987–4122, 2004. Optimal control and dynamical systems.
- [34] L. Poggiolini and M. Spadini. Strong local optimality for a bang-bang trajectory in a Mayer problem. *SIAM J. Control Optimization*, 49:140–161, 2011.

- [35] L. Poggiolini and G. Stefani. On second order sufficient conditions for a bang-singular arc. *Proceedings of science - SISSA*, 2005.
- [36] L. Poggiolini and G. Stefani. Minimum time optimality of a partially singular arc: second order conditions. In *Lagrangian and Hamiltonian methods for nonlinear control 2006*, volume 366 of *Lecture Notes in Control and Inform. Sci.*, pages 281–291. Springer, Berlin, 2007.
- [37] L. Poggiolini and G. Stefani. Sufficient optimality conditions for a bang-singular extremal in the minimum time problem. *Control Cybernet.*, 37(2):469–490, 2008.
- [38] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, New Jersey, 1970.
- [39] W. Rudin. *Real and complex analysis*. Mc Graw-Hill, New York, 1987.

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Statement of the problem and assumptions</b>	<b>4</b>
2.1	Statement of the problem . . . . .	4
2.2	First order analysis . . . . .	5
2.3	Critical cones . . . . .	6
<b>3</b>	<b>Second order analysis</b>	<b>8</b>
3.1	Second variation . . . . .	8
3.2	Necessary condition . . . . .	9
<b>4</b>	<b>Goh Transformation</b>	<b>11</b>
4.1	Transformed critical directions . . . . .	12
4.2	Transformed second variation . . . . .	13
4.3	New second order condition . . . . .	14
<b>5</b>	<b>Sufficient condition</b>	<b>19</b>
<b>6</b>	<b>Extensions and an example</b>	<b>26</b>
6.1	Including parameters . . . . .	26
6.2	Application to minimum-time problem . . . . .	29
6.2.1	Example: Markov-Dubins problem . . . . .	29
<b>7</b>	<b>Appendix</b>	<b>31</b>



---

Centre de recherche INRIA Saclay – Île-de-France  
Parc Orsay Université - ZAC des Vignes  
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex  
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier  
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq  
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique  
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex  
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex  
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex  
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

---

Éditeur  
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)  
<http://www.inria.fr>  
ISSN 0249-6399