Convergence of the shooting algorithm for singular optimal control problems

Maria Soledad Aronna

Convergence of the shooting algorithm for singular optimal control problems*

M. Soledad Aronna[1]

***Abstract*— In this article we propose a shooting algorithm for optimal control problems governed by systems that are affine in one part of the control variable. Finitely many equality constraints on the initial and final state are considered. We recall a second order sufficient condition for weak optimality, and show that it guarantees the local quadratic convergence of the algorithm. We show an example and solve it numerically.**

***Index Terms*— optimal control, singular control, second order optimality condition, weak optimality, shooting algorithm, Gauss-Newton method**

I. INTRODUCTION

We investigate optimal control problems governed by ordinary differential equations that are affine in one part of the control variable. This class of system includes both the totally affine and the nonlinear cases. This study is motivated by many models that are found in practice. Among them we can cite the followings: the Goddard's problem analyzed in Martinon et al. [4], [5], [16], other models concerning the motion of a rocket in Lawden [15], Bell and Jacobson [3], Goh [13], Oberle [19], and an optimal production process in Cho et al. [6].

We can find shooting-like methods applied to the numerical solution of partially affine problems in, for instance, Oberle [18], [20] and Oberle-Taubert [21], where the authors use a generalization of the algorithm that Maurer [16] suggested for totally affine systems. These works present interesting implementations of a shooting-like algorithm, but they do not link the convergence of the method with sufficient conditions of optimality as it is done in this article.

In this paper we propose a shooting algorithm which can be also used to solve problems with bounds on the controls. We give a theoretical support to this method, by showing that a second order sufficient condition for optimality proved in Aronna [1] ensures the local quadratic convergence of the algorithm.

The article is organized as follows. In Section II we give the statement of the problem, the main definitions and assumptions, and a first order optimality condition. The shooting algorithm is described in Section III. In Section IV we recall a second order sufficient condition for weak optimality. We state the main result of the article in Section V. In Section VI we work out an example and solve it numerically.

*This article will appear in the Proceedings of the European Control Conference to be held in Zurich, Switzerland, 2013.

[1]M.S. Aronna is a Marie Curie Fellow within the Network SADCO at the Department of Electrical and Electronic Engineering, Imperial College of London, South Kensington campus, London, SW7 2AZ, UK, aronna@cmap.polytechnique.fr

NOTATIONS. Let h_t denote the value of function h at time t if h is a function that depends only on t, and $h_{i,t}$ the ith component of h evaluated at t. Partial derivatives of a function h of (t,x) are referred as $D_t h$ or $\dot{h}$ for the derivative in time, and $D_x h$ or h_x for the differentiations with respect to space variables. The same convention is extended to higher order derivatives. By $L^p(0,T;\mathbb{R}^k)$ we mean the Lebesgue space with domain equal to the interval $[0,T]\subset\mathbb{R}$ and with values in $\mathbb{R}^k$. The notation $W^{q,s}(0,T;\mathbb{R}^k)$ refers to the Sobolev spaces.

II. STATEMENT OF THE PROBLEM

We study the optimal control problem (P) given by

$$J := \varphi_0(x_0, x_T) \to \min, \tag{1}$$

$$\dot{x}_t = F(x_t, u_t, v_t) = \sum_{i=0}^{m} v_{i,t} f_i(x_t, u_t), \text{ a.e. on } [0,T], \tag{2}$$

$$\eta_j(x_0, x_T) = 0, \quad \text{for } j = 1\dots, d_\eta. \tag{3}$$

Here $f_i : \mathbb{R}^{n+l} \to \mathbb{R}^n$ for $i = 0,\dots,m$, $\varphi_0 : \mathbb{R}^{2n} \to \mathbb{R}$, $\eta_j : \mathbb{R}^{2n} \to \mathbb{R}$ for $j = 1,\dots,d_\eta$ and we put, in sake of simplicity of notation, $v_0 \equiv 1$ which is not a variable. The *nonlinear control* u belongs to $\mathcal{U} := L^\infty(0,T;\mathbb{R}^l)$, while $\mathcal{V} := L^\infty(0,T;\mathbb{R}^m)$ denotes the space of *affine controls* v, and $\mathcal{X} := W^{1,\infty}(0,T;\mathbb{R}^n)$ refers to the state space. When needed, we write $w = (x,u,v)$ for a point in $\mathcal{W} := \mathcal{X}\times\mathcal{U}\times\mathcal{V}$. Assume throughout the article that data functions φ_0, f_i and η_j have Lipschitz-continuous second derivatives. A *trajectory* is an element $w\in\mathcal{W}$ that satisfies the state equation (2). If in addition, the constraints in (3) hold, we say that w is a *feasible trajectory* of problem (P).

Set $\mathcal{X}_* := W^{1,\infty}(0,T;\mathbb{R}^{n,*})$ the space of Lipschitz-continuous functions with values in the $n-$dimensional space of row-vectors with real components $\mathbb{R}^{n,*}$. Consider an element $\lambda := (\beta, p) \in \mathbb{R}^{d_\eta,*}\times\mathcal{X}_*$ and define the *pre-Hamiltonian* function

$$H[\lambda](x,u,v,t) := p_t F(x,u,v), \tag{4}$$

the *initial-final Lagrangian* function

$$\ell[\lambda](\zeta_0,\zeta_T) := \varphi_0(\zeta_0,\zeta_T) + \sum_{j=1}^{d_\eta} \beta_j \eta_j(\zeta_0,\zeta_T), \tag{5}$$

and the *Lagrangian* function

$$Ł[\lambda](w) := \ell[\lambda](x_0,x_T) + \int_0^T p_t\left(F(x_t,u_t,v_t) - \dot{x}_t\right)\mathrm{d}t.$$

Throughout the article we study a nominal feasible trajectory $\hat{w} = (\hat{x},\hat{u},\hat{v})$, that we assume to be smooth. We present now an hypothesis for the endpoint constraints. Consider the mapping

$$\begin{aligned} G\colon \mathbb{R}^n \times \mathcal{U} \times \mathcal{V} &\rightarrow \mathbb{R}^{d_\eta} \\ (x_0, u, v) &\mapsto \eta(x_0, x_T), \end{aligned}$$

where x_t is the solution of (2) associated with (x_0, u, v).

Assumption 1: The derivative of G at $(\hat{x}_0, \hat{u}, \hat{v})$ is onto.

The latter hypothesis is usually known as *qualification of the endpoint equality constraints.*

Definition 2.1: It is said that the feasible trajectory $\hat{w}$ is a *weak minimum* of problem (P) if there exists $\varepsilon > 0$ such that $\hat{w}$ is a minimum in the set of feasible trajectories $w = (x, u, v)$ satisfying

$$\|x - \hat{x}\|_\infty < \varepsilon, \quad \|u - \hat{u}\|_\infty < \varepsilon \quad \|v - \hat{v}\|_\infty < \varepsilon.$$

The following first order necessary condition holds for $\hat{w}$. See the book by Pontryagin et al. [22] for a proof.

Theorem 2.1: Let $\hat{w}$ be a weak solution satisfying Assumption 1, then there exists a unique $\hat{\lambda} = (\hat{\beta}, \hat{p}) \in \mathbb{R}^{d_\eta,*} \times \mathcal{X}_*$ such that $\hat{p}$ is solution of the *costate equation*

$$-\dot{\hat{p}}_t = D_x H[\hat{\lambda}](\hat{x}_t, \hat{u}_t, \hat{v}_t, t), \quad \text{a.e. on } [0, T], \tag{6}$$

with *transversality conditions*

$$\hat{p}_0 = -D_{x_0}\ell[\hat{\lambda}](\hat{x}_0, \hat{x}_T), \tag{7}$$

$$\hat{p}_T = D_{x_T}\ell[\hat{\lambda}](\hat{x}_0, \hat{x}_T), \tag{8}$$

and the *stationarity condition*

$$\begin{cases} H_u[\hat{\lambda}](\hat{x}_t, \hat{u}_t, \hat{v}_t, t) = 0, \\ H_v[\hat{\lambda}](\hat{x}_t, \hat{u}_t, \hat{v}_t, t) = 0, \end{cases} \quad \text{a.e. on } [0, T], \tag{9}$$

is verified.

Throughout this article $\hat{w}$ is considered to be a weak solution and thus, it satisfies (9) for its unique associated multiplier $\hat{\lambda}$. Furthermore, note that since v appears linearly in H we have that $D^2_{(u,v)^2}H[\hat{\lambda}](\hat{x}_t, \hat{u}_t, \hat{v}_t, t)$ is a singular matrix on $[0, T]$. Therefore, $\hat{w}$ is a *singular solution* (as defined in [3] and [5]).

III. The shooting algorithm

The purpose of this section is to present an appropriate numerical scheme to solve the problem (P). More precisely, we investigate the formulation and the convergence of an algorithm that approximates an optimal solution provided an initial estimate exists.

A. Optimality system

In what follows we use the first order optimality conditions (9) to provide a set of equations from which we can determine $\hat{w}$. We obtain an optimality system in the form of a *two-point boundary value problem* (TPBVP).

Throughout the rest of the article we assume, in sake of simplicity, that whenever some argument of f_i, H, ℓ, Ł or their derivatives is omitted, they are evaluated at $\hat{w}$ and $\hat{\lambda}$.

We shall recall that for the case where all the control variables appear nonlinearly ($m = 0$), the classical technique is using the stationarity equation

$$H_u[\hat{\lambda}](\hat{w}) = 0, \tag{10}$$

to write $\hat{u}$ as a function of $(\hat{x}, \hat{\lambda})$. This procedure is also detailed in [17] and [24]. One is able to do this by assuming, for instance, the *strengthened Legendre-Clebsch condition*

$$H_{uu}[\hat{\lambda}](\hat{w}) \succ 0. \tag{11}$$

In case (11) holds, due to the Implicit Function Theorem, we can write $\hat{u} = U[\hat{\lambda}](\hat{x})$ with U being a smooth function. Hence, replacing the occurrences of $\hat{u}$ by $U[\hat{\lambda}](\hat{x})$ in the state and costate equations yields a two-point boundary value problem.

On the other hand, when the system is affine in all the control variables ($l = 0$), we cannot eliminate the control from the equation $H_v = 0$ and, therefore, a different technique is employed (see e.g. [16], [2], [24]). The idea is to consider an index $1 \leq i \leq m$, and to take $\mathrm{d}^{M_i} H_v / \mathrm{d}t^{M_i}$ to be the lowest order derivative of H_v in which $\hat{v}_i$ appears with a coefficient that is not identically zero. Goh [11], [10], Kelley et al. [14] and Robbins [23] proved that M_i is even. This implies that the control does not appear the first time we derive H_v with respect to time, i.e. $\dot{H}_v$ depends only on $\hat{x}$ and $\hat{\lambda}$ and consequently, it is differentiable in time. Thus the expression

$$\ddot{H}_v[\hat{\lambda}](\hat{w}) = 0 \tag{12}$$

is well-defined. The control $\hat{v}$ can be retrieved from (12) provided that, for instance, the *strengthened generalized Legendre-Clebsch condition*

$$-\frac{\partial \ddot{H}_v}{\partial v}[\hat{\lambda}](\hat{w}) \succ 0 \tag{13}$$

holds (see Goh [10], [12], [13]). In this case, we can write $\hat{v} = V[\hat{\lambda}](\hat{x})$ with V being differentiable. By replacing $\hat{v}$ by $V[\hat{\lambda}](\hat{x})$ in the state-costate equations, we get an optimality system in the form of a boundary value problem.

In the problem studied here, where $l > 0$ and $m > 0$, we aim to use both equations (10) and (12) to retrieve the control $(\hat{u}, \hat{v})$ as a function of the state $\hat{x}$ and the multiplier $\hat{\lambda}$. We next describe a procedure to achieve this elimination that was proposed in Goh [12], [13]. First let as recall a necessary condition proved in Goh [9] and in [1, Lemma 3.10 and Corollary 5.2]. Define $[f_i, f_j]^x := (D_x f_j) f_i - (D_x f_i) f_j$, which is referred as the *Lie bracket* in the variable x of f_i and f_j.

Lemma 3.1 (Necessary conditions for weak optimality): If $\hat{w}$ is a smooth weak minimum for (P) satisfying Assumption 1, then

$$H_{uv} \equiv 0, \tag{14}$$

$$\hat{p}[f_i, f_j]^x = 0, \quad \text{for } i, j = 1, \ldots, m. \tag{15}$$

Let us show that H_v can be differentiated twice with respect to the time variable, as it was done in the totally affine case. Observe that (10) may be used to write $\dot{\hat{u}}$ as a function of $(\hat{\lambda}, \hat{w})$. In fact, in view of Lemma 3.1, the coefficient of $\dot{\hat{v}}$ in $\dot{H}_u$ is zero. Consequently,

$$\dot{H}_u = \dot{H}_u[\hat{\lambda}](\hat{x}, \hat{u}, \hat{v}, \dot{\hat{u}}) = 0 \tag{16}$$

and, if the strengthened Legendre-Clebsch condition (11) holds, $\dot{\hat{u}}$ can be eliminated from (16) yielding

$$\dot{\hat{u}} = \Gamma[\hat{\lambda}](\hat{x}, \hat{u}, \hat{v}). \tag{17}$$

Take now an index $i = 1, \ldots, m$ and observe that

$$0 = \dot{H}_{v_i} = \frac{\mathrm{d}}{\mathrm{d}t}\hat{p}\hat{f}_i = \hat{p}\sum_{j=0}^{m} \hat{v}_j [f_j, f_i]^x + H_{v_i u}\dot{\hat{u}} = \hat{p}\,[f_0, f_i]^x,$$

where Lemma 3.1 is used in the last equality. Therefore, $\dot{H}_v = \dot{H}_v[\hat{\lambda}](\hat{x}, \hat{u})$. We can then differentiate $\dot{H}_v$ one more time, replace the occurrence of $\dot{\hat{u}}$ by Γ and obtain (12) as it was desired. See that (12) together with the boundary conditions

$$H_v[\hat{\lambda}](\hat{w}_T) = 0, \tag{18}$$
$$\dot{H}_v[\hat{\lambda}](\hat{w}_0) = 0, \tag{19}$$

guarantee the second identity in the stationarity condition (9).

Notation: Denote by (OS) the set of equations consisting of (2)-(3), (6)-(8), (10), (12) and the boundary conditions (18)-(19).

Remark 3.1: Instead of (18)-(19), we could choose another pair of endpoint conditions among the four possible ones: $H_{v,0} = 0$, $H_{v,T} = 0$, $\dot{H}_{v,0} = 0$ and $\dot{H}_{v,T} = 0$, always including at least one of order zero. The choice we made will simplify the presentation of the result afterwards.

Observe now that the derivative with respect to (u, v) of the mapping $(w, \lambda) \mapsto \begin{pmatrix} H_u[\lambda](w) \\ -\ddot{H}_v[\lambda](w) \end{pmatrix}$ is given by

$$\mathcal{J} := \begin{pmatrix} H_{uu} & H_{uv} \\ -\dfrac{\partial \ddot{H}_v}{\partial u} & -\dfrac{\partial \ddot{H}_v}{\partial v} \end{pmatrix}. \tag{20}$$

On the other hand, if (11) and (13) are verified, $\mathcal{J}$ is definite positive along $(\hat{w}, \hat{\lambda})$ and, consequently, it is nonsingular. In this case we may write $\hat{u} = U[\hat{\lambda}](\hat{x})$ and $\hat{v} = V[\hat{\lambda}](\hat{x})$ from (10) and (12). Thus (OS) can be regarded as a TPBVP whenever the following hypothesis is verified.

Assumption 2: $(\hat{w}, \hat{\lambda})$ satisfies (11) and (13).

Summing up we get the following result.

Proposition 3.1 (Elimination of the control): If $\hat{w}$ is a smooth weak minimum verifying Assumptions 1 and 2, then

$$\hat{u} = U[\hat{\lambda}](\hat{x}), \quad \hat{v} = V[\hat{\lambda}](\hat{x}),$$

with smooth functions U and V.

Remark 3.2: When the linear and nonlinear controls are uncoupled, this elimination of the controls is much simpler. An example is shown in Oberle [20] where a nonlinear control variable can be eliminated by the stationarity of the pre-Hamiltonian, and the remaining problem has two uncoupled controls, one linear and one nonlinear. Another example is the one presented in Section VI.

B. The algorithm

The aim of this section is to present a numerical scheme to solve system (OS). In view of Proposition 3.1 we can define the following mapping.

Definition 3.1: Let $\mathcal{S} : \mathbb{R}^n \times \mathbb{R}^{n+d_\eta,*} =: \mathrm{D}(\mathcal{S}) \to \mathbb{R}^{d_\eta} \times \mathbb{R}^{2n+2m,*}$ be the *shooting function* given by

$$\big(x_0, p_0, \beta\big) =: \nu \quad \mapsto \mathcal{S}(\nu) := \begin{pmatrix} \eta(x_0, x_T) \\ p_0 + D_{x_0}\ell[\lambda](x_0, x_T) \\ p_T - D_{x_T}\ell[\lambda](x_0, x_T) \\ H_v[\lambda](w_T) \\ \dot{H}_v(w_0) \end{pmatrix},$$

where (x, p) is a solution of (2),(6),(10),(12) with initial conditions x_0 and p_0, $\lambda := (p, \beta)$; and where the occurrences of u and v were replaced by $u = U[\lambda](x)$ and $v = V[\lambda](x)$.

Note that solving (OS) consists of finding $\hat{\nu} \in \mathrm{D}(\mathcal{S})$ such that

$$\mathcal{S}(\hat{\nu}) = 0. \tag{21}$$

Since the number of equations in (21) is greater than the number of unknowns, the Gauss-Newton method is a suitable approach to solve it. The *shooting algorithm* we propose here consists of solving the equation (21) by the Gauss-Newton method.

C. The Gauss-Newton Method

This algorithm solves the equivalent least squares problem

$$\min_{\nu \in \mathrm{D}(\mathcal{S})} \left|\mathcal{S}\left(\nu\right)\right|^2.$$

At each iteration k, given the approximate value ν^k, it looks for Δ^k that gives the minimum of the linear approximation of problem

$$\min_{\Delta \in \mathrm{D}(\mathcal{S})} \left|\mathcal{S}(\nu^k) + \mathcal{S}'(\nu^k)\Delta\right|^2. \tag{22}$$

Afterwards it updates

$$\nu^{k+1} \leftarrow \nu^k + \Delta^k.$$

In order to solve the linear approximation of problem (22) at each iteration k, we look for Δ^k in the kernel of the derivative of the objective function, i.e. Δ^k satisfying

$$\mathcal{S}'(\nu^k)^\top \mathcal{S}'(\nu^k)\Delta^k + \mathcal{S}'(\nu^k)^\top \mathcal{S}(\nu^k) = 0.$$

Hence, to compute direction Δ^k the matrix $\mathcal{S}'(\nu^k)^\top \mathcal{S}'(\nu^k)$ must be nonsingular. Thus, Gauss-Newton method will be applicable provided that $\mathcal{S}'(\hat{\nu})^\top \mathcal{S}'(\hat{\nu})$ is invertible, where $\hat{\nu} := (\hat{x}_0, \hat{p}_0, \hat{\beta})$. It follows easily that $\mathcal{S}'(\hat{\nu})^\top \mathcal{S}'(\hat{\nu})$ is nonsingular if and only if $\mathcal{S}'(\hat{\nu})$ is one-to-one.

Furthermore, since the right hand-side of system (21) is zero, it can be proved that the Gauss-Newton algorithm converges locally quadratically if the function $\mathcal{S}$ has Lipschitz continuous derivative. The latter holds true here given the regularity hypotheses on the data functions. This convergence result is stated in the proposition below. See e.g. Fletcher [8] for a proof.

Proposition 3.2: If $\mathcal{S}'(\hat{\nu})$ is one-to-one then the shooting algorithm is locally quadratically convergent.

IV. Second order sufficient condition

In this section we present a sufficient condition for optimality proved in [1], and we state in Section V afterwards that this condition guarantees the local quadratic convergence of the shooting algorithm proposed above.

Given $(\bar{x}_0, \bar{u}, \bar{v}) \in \mathbb{R}^n \times \mathcal{U} \times \mathcal{V}$, consider the *linearized state equation*

$$\dot{\bar{x}}_t = F_{x,t}\bar{x}_t + F_{u,t}\bar{u}_t + F_{v,t}\bar{v}_t, \quad \text{a.e. on } [0,T], \tag{23}$$

$$\bar{x}(0) = \bar{x}_0, \tag{24}$$

where $F_{x,t}$ refers to the partial derivative of F with respect to x, i.e. $D_x F_t$; and equivalent notations hold for the other involved derivatives. Take an element $\bar{w} \in \mathcal{W}$ and define the second variation of the Lagrangian function

$$\Omega(\bar{w}) := \tfrac{1}{2} D^2 Ł[\hat{\lambda}](\hat{w})\, \bar{w}^2.$$

It can be proved that Ω can be written as

$$\Omega(\bar{x},\bar{u},\bar{v}) = \tfrac{1}{2} D^2 \ell(\hat{x}_0, \hat{x}_T)(\bar{x}_0, \bar{x}_T)^2 + \int_0^T \big[\tfrac{1}{2}\bar{x}^\top H_{xx}\bar{x} + \bar{u}^\top H_{ux}\bar{x} + \bar{v}^\top H_{vx}\bar{x} + \tfrac{1}{2}\bar{u}^\top H_{uu}\bar{u} + \bar{v}^\top H_{vu}\bar{u}\big]\,\mathrm{d}t.$$

Note that this mapping Ω does not contain a quadratic term on $\bar{v}$ since $H_{vv} \equiv 0$. Hence, one cannot state a sufficient condition in terms of the uniform positivity of Ω on the set of critical directions, as it is done in the totally nonlinear case. Therefore, we use a change of variables introduced by Goh in [11] and transform Ω into a quadratic mapping that may result uniformly positive in an associated transformed set of critical directions.

Consider hence the linear differential system in (23) and the change of variables

$$\begin{cases} \bar{y}_t := \displaystyle\int_0^t \bar{v}_s \mathrm{d}s, \\ \bar{\xi}_t := \bar{x}_t - F_{v,t}\,\bar{y}_t, \end{cases} \quad \text{for } t \in [0,T]. \tag{25}$$

This change of variables can be done in any linear system of differential equations, and it is often called *Goh's transformation*. Observe that $\bar{\xi}$ defined in that way satisfies the linear equation

$$\dot{\bar{\xi}} = F_x\bar{\xi} + F_u\bar{u} + B\bar{y}, \quad \bar{\xi}_0 = \bar{x}_0, \tag{26}$$

where $B := F_xF_v - \frac{\mathrm{d}}{\mathrm{d}t}F_v$.

A. Critical cones

We define now the sets of critical directions associated with $\hat{w}$. Even if we are working with control variables in L^∞ and hence the control perturbations are naturally taken in L^∞, the second order analysis involves quadratic mappings and it is useful to extend them continuously to L^2. Given $\bar{w} \in \mathcal{W}_2 := W^{1,2}(0,T;\mathbb{R}^n) \times L^2(0,T;\mathbb{R}^l) \times L^2(0,T;\mathbb{R}^m)$ satisfying (23)-(24), consider the *linearization of the endpoint constraints and cost function,*

$$D\eta_j(\hat{x}_0, \hat{x}_T)(\bar{x}_0, \bar{x}_T) = 0, \quad \text{for } j = 1, \dots, d_\eta, \tag{27}$$

$$D\varphi_0(\hat{x}_0, \hat{x}_T)(\bar{x}_0, \bar{x}_T) \le 0. \tag{28}$$

Define the *critical cone* in $\mathcal{W}_2$ by

$$\mathcal{C}_2 := \{\bar{w} \in \mathcal{W}_2 : \text{(23)-(24), (27)-(28) hold}\}. \tag{29}$$

Since we aim to state an optimality condition in terms of the variables after Goh's transformation, we transform the equations defining $\mathcal{C}_2$. Let $(\bar{x},\bar{u},\bar{v}) \in \mathcal{C}_2$ be a critical direction. Define $(\bar{\xi},\bar{y})$ by transformation (25) and set $\bar{h} := \bar{y}_T$. Then the transformed of (27)-(28) is

$$D\eta_j(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + B_T\bar{h}) = 0, \text{ for } j = 1, \dots, d_\eta, \tag{30}$$

$$D\varphi_0(\hat{x}_0, \hat{x}_T)(\bar{\xi}_0, \bar{\xi}_T + B_T\bar{h}) \le 0. \tag{31}$$

Consequently, the transformed critical cone is given by

$$\mathcal{P}_2 := \{(\bar{\xi},\bar{u},\bar{y},\bar{h}) \in \mathcal{W}_2 \times \mathbb{R}^m : \text{(26), (30)-(31) hold}\}. \tag{32}$$

B. Second variation

Next we state that if $\hat{w}$ is a weak minimum, then the transformation of Ω yields the quadratic mapping

$$\bar{\Omega}(\bar{\xi},\bar{u},\bar{y},\bar{h}) := g(\bar{\xi}_0, \bar{\xi}_T, \bar{h}) + \int_0^T \big(\tfrac{1}{2}\bar{\xi}^\top H_{xx}\bar{\xi} + \bar{u}^\top H_{ux}\bar{\xi} + \bar{y}^\top M[\lambda]\bar{\xi} + \tfrac{1}{2}\bar{u}^\top H_{uu}[\lambda]\bar{u} + \bar{y}^\top J[\lambda]\bar{u} + \tfrac{1}{2}\bar{y}^\top R[\lambda]\bar{y}\big)\,\mathrm{d}t, \tag{33}$$

with

$$M := F_v^\top H_{xx} - \dot{H}_{vx} - H_{vx}F_x, \;\; J := F_v^\top H_{ux}^\top - H_{vx}F_u,$$
$$S := \tfrac{1}{2}(H_{vx}F_v + (H_{vx}F_v)^\top),$$
$$V := \tfrac{1}{2}(H_{vx}F_v - (H_{vx}F_v)^\top),$$
$$R := F_v^\top H_{xx}F_v - (H_{vx}B + (H_{vx}B)^\top) - \dot{S},$$

$$g(\zeta_0, \zeta_T, h) := \tfrac{1}{2}\ell''(\zeta_0, \zeta_T + F_{v,T}h)^2 + h^\top(H_{vx,T}\zeta_T + \tfrac{1}{2}S_Th).$$

Easy computations show that $V_{ij} = \hat{p}[f_j, f_i]^x$, for $i, j = 1, \dots, m$. Thus, in view of Lemma 3.1, one has that $V \equiv 0$ if $\hat{w}$ is a weak minimum. Furthermore, we get the following result, which also uses [1, Theorem 4.4].

Theorem 4.1: If $\hat{w}$ is a smooth weak minimum, then

$$\Omega(\bar{x},\bar{u},\bar{v}) = \bar{\Omega}(\bar{\xi},\bar{u},\bar{y},\bar{y}_T),$$

for all $(\bar{x},\bar{u},\bar{v}) \in \mathcal{W}$ and $(\bar{\xi},\bar{u},\bar{y})$ given by (25).

C. The sufficient condition

We state now a second order sufficient condition for strict weak optimality.

Define the $\gamma-$*order* by

$$\bar{\gamma}(\bar{\zeta}_0,\bar{u},\bar{y},\bar{h}) := |\bar{\zeta}_0|^2 + |\bar{h}|^2 + \int_0^T (|\bar{u}_t|^2 + |\bar{y}_t|^2)\mathrm{d}t,$$

for $(\bar{\zeta}_0,\bar{u},\bar{y},\bar{h}) \in \mathbb{R}^n \times L^2(0,T;\mathbb{R}^l) \times L^2(0,T;\mathbb{R}^m) \times \mathbb{R}^m$. It can also be considered as a function of $(\bar{\zeta}_0,\bar{u},\bar{v}) \in \mathbb{R}^n \times L^2(0,T;\mathbb{R}^l) \times L^2(0,T;\mathbb{R}^m)$ by setting

$$\gamma(\bar{\zeta}_0,\bar{u},\bar{v}) := \bar{\gamma}(\bar{\zeta}_0,\bar{u},\bar{y},\bar{y}_T), \tag{34}$$

with $\bar{y}$ being the primitive of $\bar{v}$ defined in (25).

Definition 4.1: [$\gamma-$growth] We say that $\hat{w}$ satisfies $\gamma-$*growth condition in the weak sense* if there exist $\varepsilon, \rho > 0$ such that

$$J(w) \geq J(\hat{w}) + \rho\gamma(x_0 - \hat{x}_0, u - \hat{u}, v - \hat{v}), \qquad (35)$$

for every feasible trajectory w with $\|w - \hat{w}\|_\infty < \varepsilon$.

Theorem 4.2 (Sufficient condition for weak optimality): Let $\hat{w}$ be a smooth feasible trajectory such that Assumption 1 is satisfied. Then the following assertions hold.

(i) Assume that there exists $\rho > 0$ such that

$$\bar{\Omega}(\bar{\xi}, \bar{u}, \bar{y}, \bar{h}) \geq \rho\bar{\gamma}(\bar{\xi}_0, \bar{u}, \bar{y}, \bar{h}), \quad \text{on } \mathcal{P}_2. \qquad (36)$$

Then $\hat{w}$ is a weak minimum satisfying $\gamma-$growth in the weak sense.

(ii) Conversely, if $\hat{w}$ is a weak solution satisfying $\gamma-$growth in the weak sense then (36) holds for some $\rho > 0$.

V. Main result: Convergence of the shooting algorithm

The main result of this article is the theorem below that gives a condition guaranteeing the quadratic convergence of the shooting method near an optimal local solution.

Theorem 5.1: Suppose that $\hat{w}$ is a smooth weak minimum satisfying Assumptions 1 and 2, and such that (36) holds. Then the shooting algorithm is locally quadratically convergent.

Remark 5.1: The complete proof of this theorem can be found in [1]. The idea of the proof is to show that (36) yields the injectivity of $\mathcal{S}'(\hat{\nu})$ and then use Proposition 3.2. In order to prove that (36) implies that $\mathcal{S}'(\hat{\nu})$ is one-to-one, the following elements are employed: the linearization of (OS) which gives an expression of the derivative $\mathcal{S}'(\hat{\nu})$, the Goh's transformed of this linearized system and an associated linear-quadratic optimal control problem in the variables $(\bar{\xi}, \bar{u}, \bar{v}, \bar{h})$ involving (26) and (33).

Remark 5.2 (Bang-singular solutions): Finally we claim that the formulation of the shooting algorithm above and the proof of its local convergence (Theorem 5.1) can be done also for problems where the controls are subject to bounds of the type

$$0 \leq u_t \leq 1, \quad 0 \leq v_t \leq 1, \quad \text{a.e. on } [0,1]. \qquad (37)$$

More precisely, it holds for solutions for which each control component is a concatenation of *bang* and singular arcs, i.e. arcs saturating the corresponding inequality in (37), and arcs in the interior of the constraint. This extension follows from a transformation of the problem to one without bounds, and it is detailed in [2, Section 8] for the totally-affine case.

VI. An example

Consider the following optimal control problem treated in Dmitruk and Shishov [7]:

$$\begin{aligned} &J := -2x_{1,1}x_{2,1} + x_{3,1} \to \min, \\ &\dot{x}_1 = x_2 + u, \\ &\dot{x}_2 = v, \\ &\dot{x}_3 = x_1^2 + x_2^2 + 10x_2v + u^2, \\ &x_{1,0} = 0, \quad x_{2,0} = 0, \quad x_{3,0} = 0. \end{aligned} \qquad (38)$$

Here, Assumption 1 holds since no final constraints are considered. The pre-Hamiltonian function associated with (38) is, omitting arguments,

$$H = p_1(x_2 + u) + p_2 v + p_3(x_1^2 + x_2^2 + 10x_2v + u^2).$$

We can easily deduce that $p_3 \equiv 1$. The equations (10) and (12) for this problem give

$$H_u = p_1 + 2u, \quad \ddot{H}_v = -2v + 2x_1, \qquad (39)$$

and, therefore, Assumption 2 holds true. Agreeing with Proposition 3.1, the control can be eliminated from (39). This yields

$$u = -p_1/2, \quad v = x_1.$$

We can then write the optimality system (OS) related to (38). The state and costate equations are

$$\begin{aligned} &\dot{x}_1 = x_2 - p_1/2, \\ &\dot{x}_2 = x_1, \\ &\dot{x}_3 = x_1^2 + x_2^2 + 10x_2x_1 + p_1^2/4, \\ &\dot{p}_1 = -2x_1, \\ &\dot{p}_2 = -2x_2 - 10x_1 - p_1, \end{aligned} \qquad (40)$$

where we do not include p_3 since it is constantly equal to 1. The boundary conditions are

$$\begin{aligned} &x_{1,0} = 0, \quad x_{2,0} = 0, \quad x_{3,0} = 0, \\ &p_{1,1} = -2x_{2,1}, \quad p_{2,1} = -2x_{1,1}, \\ &H_{v,1} = p_{2,1} + 10x_{2,1} = 0, \\ &\dot{H}_{v,1} = -2x_{2,1} - p_{1,1} = 0. \end{aligned} \qquad (41)$$

Observe that the last line in (41) can be removed since it is implied by the first equation in the second line. Here the shooting function is given by

$$\mathcal{S} : \mathbb{R}^2 \to \mathbb{R}^3, \ (p_{1,0}, p_{2,0}) \mapsto \begin{pmatrix} p_{1,1} + 2x_{2,1} \\ p_{2,1} + 2x_{1,1} \\ p_{2,1} + 10x_{2,1} \end{pmatrix}. \qquad (42)$$

In [7] it was checked that the second order sufficient condition (36) held for the control $(u \equiv 0, v \equiv 0)$. The solution associated with this control has $x_1 = x_1 = x_3 = p_1 = p_2 = 0$. In view of Theorem 4.2, we know that the shooting algorithm converges quadratically for appropriate initial values of $(p_{1,0}, p_{2,0})$.

We solved (38) numerically by applying the Gauss-Newton method to the equation $\mathcal{S}(p_{1,0}, p_{2,0}) = 0$, for $\mathcal{S}$ defined in (42). We used implicit Euler scheme for numerical integration of the differential equation. For arbitrary guesses of $(p_{1,0}, p_{2,0})$, the algorithm converged to $(0,0)$; in all the occasions. The tests were done with Scilab.

VII. Conclusions

We investigated optimal control problems with systems that are affine in some components of the control variable and that have finitely many equality endpoint constraints. For a Mayer problem of this kind of system we proposed a numerical indirect method for approximating a weak solution. For qualified solutions, we proved that the local convergence of the method is guaranteed by a second order sufficient condition for optimality proved before by the author.

We presented an example, in which we showed how to eliminate the control by using the optimality conditions, proposed a shooting formulation and solved it numerically. The tests converged, as it was expected in view of the theoretical result.

ACKNOWLEDGMENT

Part of this work was done under the supervision of Frédéric Bonnans during my Ph.D. study. I acknowledge him for his great guidance.

I also thank Xavier Dupuis for his careful reading, and the three anonymous reviewers for their useful remarks.

This work is supported by the European Union under the 7th Framework Programme FP7-PEOPLE-2010-ITN Grant agreement number 264735-SADCO.

REFERENCES

[1] M.S. Aronna. Singular solutions in optimal control: second order conditions and a shooting algorithm. *INRIA Research Rapport Nr. 7764 or arXiv:1210.7425*, 2011.

[2] M.S. Aronna, J.F. Bonnans, and Martinon P. A shooting algorithm for problems with singular arcs. *J. Optim. Theory Appl.*, Published as 'Online First', 2013.

[3] D.J. Bell and D.H. Jacobson. *Singular Optimal Control Problems*. Academic Press, 1975.

[4] F. Bonnans, J. Laurent-Varin, P. Martinon, and E. Trélat. Numerical study of optimal trajectories with singular arcs for an Ariane 5 launcher. *J. Guidance, Control, and Dynamics*, 32(1):51–55, 2009.

[5] A.E. Bryson, Jr. and Y.C. Ho. *Applied optimal control*. Hemisphere Publishing Corp. Washington, D. C., 1975. Optimization, estimation, and control, Revised printing.

[6] D.I. Cho, P.L. Abad, and M. Parlar. Optimal production and maintenance decisions when a system experience age-dependent deterioration. *Optimal Control Appl. Methods*, 14(3):153–167, 1993.

[7] A. V. Dmitruk and K. K. Shishov. Analysis of a Quadratic Functional with a Partly Singular Legendre Condition. *Moscow University Computational Mathematics and Cybernetics*, 34(1):16–25, 2010.

[8] R. Fletcher. *Practical methods of optimization. Vol. 1*. John Wiley & Sons Ltd., Chichester, 1980. Unconstrained optimization, A Wiley-Interscience Publication.

[9] B.S. Goh. Necessary conditions for singular extremals involving multiple control variables. *J. SIAM Control*, 4:716–731, 1966.

[10] B.S. Goh. *Necessary Conditions for the Singular Extremals in the Calculus of Variations*. University of Canterbury, 1966.

[11] B.S. Goh. The second variation for the singular Bolza problem. *J. SIAM Control*, 4(2):309–325, 1966.

[12] B.S. Goh. Compact forms of the generalized Legendre-Clebsch conditions and the computation of singular control trajectories. In *Proceedings of the American Control Conference*, volume 5, pages 3410–3413, 1995.

[13] B.S. Goh. Optimal singular rocket and aircraft trajectories. In *Control and Decision Conference, 2008. CCDC 2008*, pages 1531 –1536, 2008.

[14] H.J. Kelley, R.E. Kopp, and H.G. Moyer. Singular extremals. In *Topics in Optimization*, pages 63–101. Academic Press, New York, 1967.

[15] D. F. Lawden. *Optimal trajectories for space navigation*. Butterworths, London, 1963.

[16] H. Maurer. Numerical solution of singular control problems using multiple shooting techniques. *J. Optim. Theory Appl.*, 18(2):235–257, 1976.

[17] H. Maurer and W. Gillessen. Application of multiple shooting to the numerical solution of optimal control problems with bounded state variables. *Computing*, 15(2):105–126, 1975.

[18] H.J. Oberle. Numerische Behandlung singulärer Steuerungen mit der Mehrzielmethode am Beispiel der Klimatisierung von Sonnenhäusern. *PhD thesis. Technische Universität München*, 1977.

[19] H.J. Oberle. On the numerical computation of minimum-fuel, Earth-Mars transfer. *J. Optimization Theory Appl.*, 22(3):447–453, 1977.

[20] H.J. Oberle. Numerical computation of singular control functions in trajectory optimization problems. *J. Guidance Control Dynam.*, 13(1):153–159, 1990.

[21] H.J. Oberle and K. Taubert. Existence and multiple solutions of the minimum-fuel orbit transfer problem. *J. Optim. Theory Appl.*, 95(2):243–262, 1997.

[22] L. Pontryagin, V. Boltyanski, R. Gamkrelidze, and E. Michtchenko. *The Mathematical Theory of Optimal Processes*. Wiley Interscience, New York, 1962.

[23] H.M. Robbins. A generalized Legendre-Clebsch condition for the singular case of optimal control. *IBM J. of Research and Development*, 11:361–372, 1967.

[24] E. Trélat. Optimal Control and Applications to Aerospace: Some Results and Challenges. *J. Optim. Theory Appl.*, 154(3):713–758, 2012.