Pure and Applied Functional Analysis
Volume 1, Number 2, 2016, 197–215

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

A NEWTON-TYPE METHOD AND OPTIMALITY TEST FOR PROBLEMS WITH BANG-SINGULAR-BANG OPTIMAL CONTROL

URSULA FELGENHAUER

Abstract. In the paper, a second-order method is investigated for control problems where the control has bang-singular-bang structure. To start with, the first-order system is formulated as a variational inequality including a differential equation for the switching function. For this system, a certain inexact Josephy-Newton method is considered. Under strong second-order sufficient conditions obtained via Goh's transformation of control and state-adjoint data, the local solvability is proved for the iterative problems. For so-called semilinear problems, locally the state and adjoint iterates converge superlinearly. The paper is completed by a Riccati test approach for verifying second-order optimality.

1. Introduction

Optimal control functions with bang-bang and singular control arcs typically appear in Mayer problems with control-affine dynamics and constant control bound constraints. Intervals where the control components take exclusively extremal values are called *bang-arcs*, and the jumps between upper and lower bound values are determined by change in signs of so-called switching functions. If some component of the switching function vanishes along a whole interval, instead, the related control component may have optimal values in the inner of its feasible set, and we speak of *singular arcs*.

In recent years, several interesting results on theoretical behavior as well as numerical approximation of bang-singular-bang optimal controls have been achieved. They cover fundamental optimality conditions [4, 9, 24, 25, 28] for control-affine systems with singular control arcs, their structural stability investigation [13, 14, 26], and convergence proofs for the shooting method [5, 29, 30] and Euler type discretizations [3, 15]. As in problems with continuous or purely bang-bang optimal control functions, there is a close relation between stability of solutions under small data perturbation and second-order sufficient optimality conditions. However, in case of partly singular control regimes it is well known that local quadratic growth of the objective can be expected only in a weak sense: indeed, the growth term is quadratic w.r.t. L_2 norm of the control primitive and its boundary values, but not

2010 *Mathematics Subject Classification.* 49K60, 49K15.
Key words and phrases. Bang-singular-bang control, maximum principle, Josephy–Newton method, second-order conditions, Riccati equation.

coercive w.r.t. controls. The derivation is based on the classical Goh transformation [16, 17]; see also [19, 20]. The consequences for stability analysis of first-order necessary optimality conditions given by Pontryagin's maximum principle had been carefully discussed e.g. in [14]. In that paper of 2013, a detailed analysis was given for the linearized variational inequality related to the optimality system and it was pointed out, that linearizations could be useful in designing approximation methods for control problems with partly singular optimal control. Following this line, convergence has been proved particularly for discretizations applied to so-called semilinear systems; cf. [15].

The present paper will discuss the following two questions: (i) is it possible to use a Newton type iteration for the problem class; and (ii) is there a possibility to confirm the second-order coercivity condition (– at least a-posteriorily –) numerically ?
First, in Section 2 we reformulate the variational inequality (VI) related to the maximum principle in a way that includes a differential equation for the switching function: this formulation slightly differs from the formerly used algebraic notation (cf. [14, 15]) but, in the following, it allows to make the influence of perturbations more transparent (compare, e.g., Theorem 3.3 with [14, Theorems 1,2]).
Secondly, an appropriate sufficient second-order optimality condition is provided; see condition (H2). Section 3.1 is devoted to constructing a Josephy–Newton method for (VI). While here the linearization is defined at current iterates (and not at the solution as in [14]), under Goh's transformation some symmetry properties of underlying main operators may be lost. Guided by the idea of approximating switching function data as good as possible, we therefore propose a certain *inexact* Josephy–Newton approach for which, similarly to (H2), some coercivity condition holds (Lemma 3.2). The latter guarantees well-posedness of the iteration subproblems at least in certain neighborhood of the solution (x^*, u^*, p^*) in $L_\infty \times L_1 \times L_\infty$ (Theorem 3.3). Restricting further the problem class to semilinear state equations where the control coefficients are constant, in Section 3.2 a first local convergence result is proved (Theorem 3.6). We'd like to mention that the investigation in Section 3 was inspired by recent papers [11, 12] on Newton's method under metric regularity conditions. However, applicability of such concept to optimal control problems with bang-bang or bang-singular behavior seems to be a highly challenging, and widely open question yet.
In the final Section 4, an optimality test is derived by adapting the Riccati approach as it is known from problems with Lipschitz continuous control solution [6, 21], but also in the context of bang-bang optimal control, e.g., [22]. Former results from [29, 30] have been completed by some new boundary value inequality (Lemma 4.4) inherent for data of the second variation, and by detailed boundary conditions on the matrix test function (Theorem 4.5). An illustration for the test is given in Section 4.2.

Notations. Let $(x, u) : [0, 1] \to \mathbb{R}^n \times \mathbb{R}$ be a state-control pair, and $p : [0, 1] \to \mathbb{R}^n$ the co-state function. We write $L_q(0, 1; \mathbb{R}^k)$ for the Lebesgue space of order q of vector-valued functions on $[0, 1]$ and denote by $W^l_q(0, 1; \mathbb{R}^k)$ the related Sobolev space. The norms are given as $\| \cdot \|_q$ and $\| \cdot \|_{l,q}$, $(1 \le q \le \infty, l \ge 1)$, resp. Further, denote the scalar product in L_2 by $(\cdot, \cdot)$. For vectors in Euclidean space

$\mathbb{R}^n$ use $|\cdot|$ for the norm, and write the scalar products in matrix notation as $(a, b) = a^T b$. Superscript T is generally used for transposition of matrices resp. vectors. The symbol ∇_x denotes (partial) gradients or transposed Jacobians, i.e. $(\nabla_x h)_{ij} = \partial h_j / \partial x_i$ for $h : \mathbb{R}^n \to \mathbb{R}^m$, $m \geq 1$, whereas $\nabla^2_{xx}, \nabla^2_{xu}$ etc. denote related partial Hessian matrices for real-valued functions. For functions $f, g : \mathbb{R}^n \to \mathbb{R}^n$, Lie brackets then are given by $[f, g] = \nabla_x g^T f - \nabla_x f^T g$. The normal cone $N_+(\mu), \mu \in \mathbb{R}$, is empty in case $\mu \not\geq 0$. Otherwise, $N_+(\mu)$ is defined as $N_+(\mu) = \{z \in \mathbb{R} : \; z\,(\mu' - \mu) \leq 0 \quad \forall\, \mu' \geq 0\}$.

2. Statement of the problem. Assumptions

We consider a control-affine problem with scalar-valued control function under bound constraints,

(CP) $$\text{minimize} \quad J(x, u) \;:=\; k(x(1))$$

subject to

$$\dot{x}(t) \;=\; f(x(t)) \;+\; g(x(t))\, u(t) \quad \text{a.e. in } [0,1], \tag{2.1}$$

$$x(0) \;=\; a, \tag{2.2}$$

$$-1 \;\leq\; u(t) \;\leq\; 1, \quad \text{a.e. in } [0,1]\,, \tag{2.3}$$

$$x \in W^1_\infty(0,1;\mathbb{R}^n), \;\; u \in L_\infty(0,1;\mathbb{R}). \tag{2.4}$$

For brevity, the analysis will be restricted to the case of fixed initial and free terminal state values. The data functions in (CP) are assumed to satisfy

(H0) *The functions f, g and k are twice continuously differentiable on $\mathbb{R}^n$ with Lipschitz continuous second-order derivatives; i.e $f, g, k \in C^{2,1}$. Further, there exist constants c, L independent of x such that*

$$|f(x)| \;+\; |g(x)| \;\leq\; c \;+\; L\,|x|.$$

Under assumption (H0), the optimal control problem has a solution (x^*, u^*).

The pair $(x, u) \in W^1_\infty \times L_\infty$ is called admissible if (2.1) – (2.3) are fulfilled. Let us define the Pontryagin function $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ by

$$H(x, u, p) \;=\; p^T f(x) \;+\; p^T g(x)\, u,$$

and the Lagrange functional $L : \; W^1_2 \times L_2 \times W^1_2 \times L_2 \times \mathbb{R}^n \to \mathbb{R}$ with vector-valued adjoint function $p\colon [0,1] \to \mathbb{R}^n$ and multipliers $\mu\colon [0,1] \to \mathbb{R}^2_+$, $\kappa \in \mathbb{R}^n$,

$$\begin{aligned} L(x, u, p, \mu, \kappa) \;\; &= \;\; k(x(1)) + (p, \dot{x} - f(x) - g(x)u) \\ &\qquad + \kappa^T (x(0) - a) \;+\; (\mu_1, u - 1) - (\mu_2, u + 1). \end{aligned}$$

The first variation of L w.r.t. (x, u) at (x^*, u^*) can be written as

$$\begin{aligned} \delta^1 L(x^*, u^*, p, \mu, \kappa)[z, v] \;\; &= \;\; (\nabla k(x^*(1)) + p(1))^T z(1) \;+\; (\kappa - p(0))^T z(0) \\ &\quad -(\dot{p} + \nabla_x H(x^*, u^*, p),\, z) \\ &\qquad -\,(g(x^*)^T p - \mu_1 + \mu_2, v). \end{aligned}$$

In the control variation term, the so-called switching function $\sigma = g(x^*)^T p$ appears and, for stationary points, $\mu_1 = [\sigma]_+, \mu_2 = [\sigma]_-$ coincide with positive resp. negative part of it. Using the control primitive,

$$y(t) := \int_0^t v(s)\,ds, \qquad b := y(1) \tag{2.5}$$

with $y(0)=0$, we may rewrite this part and obtain

$$(g(x^*)^T p - \sigma, v) \;=\; y(1)\cdot\big(g(x^*(1))^T p(1) - \sigma(1)\big) \;-\; (p^T[f,g](x^*) - \dot\sigma, y),$$

where $[f,g] = \nabla g^T f - \nabla f^T g$ denotes the Lie bracket of the vector fields f and g. This observation is used to reformulate first-order necessary conditions on (x^*,u^*) as a variational problem:

(VI) Find functions p, σ, μ_1 and μ_2 in W_2^1 satisfying

$$\dot x - f(x) - g(x)u \;=\; 0, \quad x(0) = a, \tag{2.6}$$

$$\dot p + \nabla_x H(x,u,p) \;=\; 0, \qquad p(1) = -\nabla_x k(x(1)), \tag{2.7}$$

$$\dot\sigma \;=\; p^T[f,g](x), \quad \sigma(1) = p(1)^T g(x(1)), \tag{2.8}$$

$$\sigma \;-\; \mu_1 + \mu_2 \;=\; 0, \tag{2.9}$$

$$u - 1 \;\in\; N_+(\mu_1), \qquad -u - 1 \;\in\; N_+(\mu_2) \tag{2.10}$$

for almost every $t \in [0,1]$.

The set N_+ in (2.10) stands for the normal cone to $\mathbb{R}_+$, so that (2.6) – (2.10) represents a generalized equation. Since (2.10) is equivalent to the maximum condition

$$\sigma(t)\cdot(u(t) - u^*(t)) \;\le\; 0$$

for all feasible u and almost every $t \in [0,1]$, one can equally speak of a variational inequality (VI) determining (x^*,u^*) and (p,σ,μ). Notice that the adjoint equation is linear and has the form $\dot p = -A^T p$ with A given in (2.11) below. Therefore, adjoint and multiplier functions related to any solution of (CP) are uniquely determined by (VI) and will belong to W_∞^1.

Next consider the second variation of L at the stationary solution,

$$\delta^2 L(x^*,u^*,p,\mu,\kappa)[z,v]^2 \;=\; \frac{1}{2} z(1)^T K\, z(1) \;-\; \frac{1}{2}(z, \nabla^2_{xx}H^* z \;+\; 2\,\nabla^2_{xu}H^* v),$$

where $K = \nabla^2 k(x^*(1))$, and $\nabla^2_{xx}H^*, \nabla^2_{xu}H^*$ stand for partial second-order derivative data of H evaluated along (x^*,u^*,p). The variations (z,v) in $\delta^2 L$ will be restricted to the tangent set of the state equation constraint, i.e.,

$$\dot z \;=\; A\,z \;+\; B\,v, \quad z(0) \;=\; 0, \tag{2.11}$$

where $A = \nabla_x (f + gu^*)^T\big|_{x=x^*}$, $B = g(x^*)$. Now apply the so-called Goh transformation to $\delta^2 L$ where we use y defined in (2.5):

Denote $P_1 = \nabla^2_{xu} H^*$ and $b = y(1)$. Then, integration by parts yields

$$\begin{aligned}(z, \nabla^2_{xu}H^* v) \;&=\; b\cdot z(1)^T P_1(1) \;-\; \left(\frac{d}{dt}(P_1^T z), y\right)\\ &=\; b\cdot z(1)^T P_1(1) \;-((\dot P_1^T + P_1^T A)z, y) \;-\; (P_1^T B\dot y, y)\\ &=\; -\frac{1}{2} b^2 B(1)^T P_1(1) \;+\; b\cdot z(1)^T P_1(1)\end{aligned}$$

$$- \int_0^1 z^T (A^T P_1 + \dot{P}_1) y \, dt + \frac{1}{2} \int_0^1 \frac{d}{dt} (B^T P_1) \, y^2 dt.$$

For completing Goh's transformation, substitute $z = \zeta + By$ and get

$$\dot{\zeta} = A\zeta + B_1 y, \;\; \zeta(0) = 0 \quad \text{with } B_1 = AB - \dot{B}. \tag{2.12}$$

Inserting this new function into $\delta^2 L$, the following quadratic form is obtained

$$\begin{aligned} \delta^2 L \;=\; & \frac{1}{2} \zeta(1)^T K \, \zeta(1) + \zeta(1)^T W_1 b + \frac{1}{2} V_1 b^2 \\ & - \frac{1}{2} \int_0^1 (\zeta^T C \, \zeta + 2 \zeta^T M^T y + R \, y^2) \, dt \quad =: \; \Omega(y, b) \end{aligned} \tag{2.13}$$

where $W_1 = K\,B(1) - P_1(1)$, $V_1 = B(1)^T W_1$ (cf. [4]). The matrices C, M and R abbreviate

$$C = \nabla_{xx}^2 H^*, \quad M = B^T C - P_1^T A - \dot{P}_1^T, \quad R = B^T M^T - P_1^T B_1,$$

(or $M = -p^T (\nabla_x [f, g]^*)^T$, $R = -p^T \, [g, [f, g]]^*$; cf. [14]).

The quadratic form $\Omega(y, b)$ has been utilized in [7–9] and [4] to formulate second-order sufficient optimality conditions for (CP). Before recalling this definition, we add a structural assumption on the reference control u^*:

(H1) (strict structural assumption)
The function u^* is of strict *bang-singular-bang structure, i.e., there exist points t_1, t_2 with $0 < t_1 < t_2 < 1$ and a constant $m \in (0, 1)$ such that $u^* \equiv u_1$ a.e. on $[0, t_1)$, $u^* \equiv u_2$ on $(t_2, 1]$, $u_1, u_2 \in \{-1, 1\}$, and $|u^*(t)| < 1 - m$ a.e. on (t_1, t_2). Moreover, $\sigma \neq 0$ on $[0, t_1) \cup (t_2, 1]$.*

(H2) (strong second-order optimality condition)
There exists a constant $m > 0$ such that

$$\Omega(y, b) \;\geq\; m \, \left(\|y\|_2^2 + |b|^2 \right)$$

for all $(y, b) \in L_2 \times \mathbb{R}$ satisfying $y \equiv 0$ on $[0, t_1]$, $y \equiv b$ on $[t_2, 1]$, and ζ, Ω defined by (2.12), (2.13).

(Without loss of generality, for both conditions (H1) and (H2) the same constant m can be used.)

Proposition 2.1 ([4, Theor. 5.5]). *Suppose that there exists $m > 0$ such that* (H2) *is fulfilled. Then (x^*, u^*) is a Pontryagin minimum satisfying γ–*quadratic growth*: $\exists\, m' > 0$ such that, for any sequence (x^k, u^k) of feasible pairs with $v^k = u^k - u^*$ converging to zero in the Pontryagin sense,*

$$J(x^k, u^k) \,-\, J(x^*, u^*) \,\geq\, m' \gamma(y^k, b^k), \qquad \gamma(y, b) = \|y\|_2^2 + |b|^2,$$

holds for big enough k, with (y^k, b^k) related to v^k by (2.5).

3. An inexact Josephy–Newton approach

3.1. The subproblems (LVI$_G^j$).

The inclusion (VI) has the abstract form

$$0 \;\in\; \psi(w) \,+\, \mathcal{F}(w)$$

with $w = (x, u, p, \sigma, \mu)$, $\psi : W_2^1 \times L_2 \times (W_2^1)^3 \to Y \times L_2$ and $\mathcal{F}(w) = \{0_Y\} \times N_+(\mu)$. The space Y stands for $(L_2 \times \mathbb{R}^n)^2 \times L_2 \times \mathbb{R} \times W_2^1$.

Given w^0 near w^*, the standard Josephy–Newton iteration defines $w = w^{j+1},\ j = 0, 1, \dots,$ by

$$-\psi(w^j) \in \psi'(w^j)\cdot(w - w^j) + \mathcal{F}(w). \tag{3.1}$$

In detail, the following linearized problem has to be solved for $d^j = -\psi(w^j)$:

(LVI^j) Find functions $w - w^j = (z, v, q, \rho, \nu - \mu^j)$ solving the following system :

$$\dot z - A^j z - B^j v = d_1^j, \qquad z(0) = d_2^j, \tag{3.2}$$

$$\dot q + (A^j)^T q + C^j z + P_1^j v = d_3^j, \quad q(1) = -K^j z(1) + d_4^j, \tag{3.3}$$

$$\dot\rho + (B_1^j)^T q + M^j z = d_5^j, \tag{3.4}$$

$$\rho(1) = B^j(1)q(1) + (P_1^j(1))^T z(1) + d_6^j, \tag{3.5}$$

$$\rho + \sigma^j - \nu_1 + \nu_2 = 0, \tag{3.6}$$

$$u^j + v - 1 \in N_+(\nu_1), \qquad -u^j - v - 1 \in N_+(\nu_2) \tag{3.7}$$

for almost every $t \in [0, 1],\ j = 1, 2 \dots.$
The coefficient functions are given by $K^j = \nabla^2 k(x^j(1)),\ B^j = g(x^j)$, and

$$A^j = \nabla_x (f + g u^j)^T\big|_{x=x^j}, \quad B_1^j = -[f, g]\,|_{x=x^j}, \quad P_1^j = \nabla^2_{xu} H(x^j, u^j, p^j),$$
$$C^j = \nabla^2_{xx} H(x^j, u^j, p^j), \qquad M^j = -(p^j)^T\,(\nabla_x [f, g])^T\big|_{x=x^j},$$

whereas $d^j = -\psi(w^j)$ stands for the defects in (2.6)–(2.8) at $w = w^j$, i.e.

$$d_1^j = f(x^j) + g(x^j)u^j - \dot x^j, \qquad d_2^j = a - x^j(0)$$
$$d_3^j = -\nabla_x H(x^j, u^j, p^j) - \dot p^j, \qquad d_5^j = (p^j)^T\,[f, g]\,|_{x=x^j} - \dot\sigma^j$$

etc. As in [14, Section 3], one can apply Goh's transformation (2.5) and define $\zeta = z - B^j y,\ \ \eta = q + P_1^j y$. The new functions will satisfy

$$\dot\zeta = A^j\zeta + \bar B_1^j y + d_1^j, \quad \dot\eta = -(A^j)^T\eta - C^j z - (\bar M^j)^T y + d_3^j \tag{3.8}$$

for $\bar B_1^j = A^j B^j - \dot B^j,\ \bar M^j = (B^j)^T C^j - (\dot P_1^j)^T - (P_1^j)^T A^j$. Since $(x^j, u^j, p^j, \sigma^j)$ do not exactly solve (2.6)–(2.8), the matrices differ from B_1^j resp. M^j appearing in (3.4). This may cause some conflict in proving existence and uniqueness of solutions for (LVI^j). Therefore, we consider the following *inexact* version of the Josephy–Newton iteration under Goh's transformation,

(LVI$_G^j$) Find $(z, v, q, \rho, \nu - \mu^j)$ solving the following system :

$$\dot\zeta - A^j\zeta - B_1^j y = d_1^j, \qquad \zeta(0) = d_2^j, \tag{3.9}$$

$$\dot\eta + (A^j)^T\eta + C^j\zeta + (M^j)^T y = d_3^j, \tag{3.10}$$

$$\eta(1) = -K^j\zeta(1) + W^j b + d_4^j, \tag{3.11}$$

$$\dot\rho + (B_1^j)^T\eta + M^j\zeta + R^j y = d_5^j, \tag{3.12}$$

$$\rho(1) = (B^j(1))^T\eta(1) + (P_1^j(1))^T\zeta(1) + d_6^j, \tag{3.13}$$

$$\dot y = v, \qquad y(0) = 0, \tag{3.14}$$

$$\rho + \sigma^j - \nu_1 + \nu_2 = 0, \tag{3.15}$$

$$(3.16) \qquad u^j + v - 1 \in N_+(\nu_1), \qquad -u^j - v - 1 \in N_+(\nu_2)$$

for almost every $t \in [0,1], j = 1,2,\ldots$, where R^j, W^j are given by

$$\begin{aligned} R^j &= M^j B^j - (B_1^j)^T P_1^j = -(p^j)^T \left.[g,[f,g]]\right|_{x=x^j}, \\ W^j &= P_1^j(1) - K^j B^j(1). \end{aligned}$$

If $\xi = (\zeta, y, \eta, \rho, \nu)$ together with $v \in L_2$ solve (LVI$_G^j$), one can find w^{j+1} by use of $z = \zeta + B^j y,\ q = \eta - P_1^j y$ respectively. Hence, in an abstract setting, problem (LVI$_G^j$) replaces (3.1) by

$$(3.17) \qquad d^j := -\psi(w^j) \in T^j(w - w^j) + \mathcal{F}(w)$$

with some linear operator T^j depending only on w^j. The error $(T^j - \psi'(w^j))$ can be estimated as follows:

Lemma 3.1. *Assume that $\{x^j\}$, $\{p^j\}$ are uniformly bounded in L_∞. Then $\Delta^j := (T^j - \psi'(w^j)) \cdot (w - w^j)$ satisfies*

$$\|\Delta_1^j\|_2 + \|\Delta_3^j\|_2 \le c_T \|y\|_\infty \cdot (\|d_1^j\|_2 + \|d_3^j\|_2),$$

for some constant c_T independent of j, and $\Delta_k^j = 0$ for $k \notin \{1,3\}$.

Sketch of the proof. By direct calculation obtain

$$\begin{aligned} \Delta_1^j &= (\bar{B}_1^j - B_1^j) \cdot y = \nabla_x g(x^j)^T d_1^j \cdot y, \\ \Delta_3^j &= (\bar{M}^j - M^j)^T y = \left[\nabla_x^2 \left((p^j)^T g(x^j)\right) d_1^j + \nabla_x g(x^j) d_3^j\right] \cdot y. \end{aligned}$$

The desired estimates follow by assumption (H0). □

For analyzing the inexact linearization (LVI$_G^j$) of (VI), we may equivalently rewrite problem (3.17) as

$$(3.18) \qquad \bar{d}^j := T^j(w^j - w^*) - \psi(w^j) \in T^j(w - w^*) + \mathcal{F}(w)$$

and abbreviate $\bar{w} = w^j - w^*,\ \bar{w}^+ = w^{j+1} - w^*$. The data obtained under Goh's transformation will be denoted by $\bar{\xi}$ and $\bar{\xi}^+$: then $\bar{\xi}^+$ together with $\bar{v}^+ = u^{j+1} - u^*$ solve (3.9)–(3.14) from (LVI$_G^j$) with right-hand side $\bar{d} = \bar{d}^j$ defined above, and satisfy $\rho + \sigma^* = \nu_1 - \nu_2$ together with

$$(3.19) \qquad v + u^* - 1 \in N_+(\nu_1), \qquad -v - u^* - 1 \in N_+(\nu_2).$$

Notice that, for $\bar{d} = 0$, the problem has the reference solution $\bar{v}^+ = 0$ and $\bar{\xi}^+ = (0,0,0,0,\mu^*)$ related to $w^{j+1} = w^*$.

Solution existence and uniqueness for (3.18) in case that w^j is sufficiently close to w^* can be proved in analogy to [14, Lemma 3.1-3.3]. The crucial point is a coercivity property similar to (H2) for the following symmetric quadratic form:

$$\begin{aligned} (3.20) \qquad \Omega^j(y,b) &= \frac{1}{2}\zeta_0(1)^T K^j \zeta_0(1) + \zeta_0(1)^T W_1^j b + \frac{1}{2} V_1^j b^2 \\ &\quad - \frac{1}{2}\int_0^1 (\zeta_0^T C^j \zeta_0 + 2\zeta_0^T (M^j)^T y + R^j y^2)\, dt, \end{aligned}$$

where $W_1^j = K^j B^j(1) - P_1^j(1),\ V_1^j = B^j(1)^T W_1^j$, and $\zeta_0 = S^j v$ is obtained from solving (3.9)–(3.11) for given $v \in L_2$, y from (3.14) and $d^j = 0$.

Lemma 3.2. *Assume* (H0) –(H2) *to hold true. Then there exist constants* $R', m' > 0$, $\delta > 0$ *independent of* j *such that, for all* (x^j, u^j, p^j) *inside*

$$W' := \{w : \|u^j - u^*\|_1 + \|x^j - x^*\|_\infty + \|p^j - p^*\|_\infty \leq R'\},$$

the estimate

$$\Omega^j(y,b) \geq m' \left(\|y\|_2^2 + |b|^2\right)$$

holds for all $(y,b) \in L_2 \times \mathbb{R}$ *with* $y \equiv 0$ *on* $[0, t_1 - \delta]$ *and* $y \equiv b$ *on* $[t_2 + \delta, 1]$.

Proof. First, we find estimates for the differences of vectors and matrices defining Ω resp. Ω^j: under the given assumptions,

$$\|B^j - B\|_\infty + \|B_1^j - B_1\|_\infty + \|K^j - K\| = \mathrm{O}(\|x^j - x^*\|_\infty),$$

and

$$\|M^j - M\|_\infty + \|P_1^j - P_1\|_\infty + \|R^j - R\|_\infty = \mathrm{O}(\|x^j - x^*\|_\infty + \|p^j - p^*\|_\infty).$$

Further, $\|A^j - A\|_1 + \|C^j - C\|_1 = \mathrm{O}(R')$ because these matrices additionally depend on the control functions.
It remains to compare $\zeta_0 = S^j v$ with $\zeta_0^* = Sv$ as it has appeared in (2.13): let Φ^j, Ψ^j be fundamental matrix solutions for

$$\dot{\Psi}^j = A^j \Psi^j, \quad \dot{\Phi}^j = -(A^j)^T \Phi^j, \quad \Psi^j(0) = \Phi^j(0) = I,$$

and define Φ, Ψ analogously for A replacing A^j. Then the estimate for $(A^j - A)$ yields $\|\Phi^j - \Phi\|_\infty + \|\Phi^j - \Phi\|_\infty = \mathrm{O}(R')$. Now one can represent ζ_0 (and similarly ζ_0^*) in the form

$$\zeta_0(t) = \Psi^j(t) \int_0^t (\Phi^j(s))^T B_1(s) y(s)\, ds,$$

so that $\|\zeta_0 - \zeta_0^*\|_\infty = \|y\|_2 \cdot \mathrm{O}(R')$ follows. Finally, $|\Omega^j(y,b) - \Omega(y,b)| = \|y\|_2^2 \cdot \mathrm{O}(R')$, and the assertion follows from (H2) by continuity of data. □

Based on the last Lemma, existence and uniqueness of solutions of (LVI_G^j) for small *rhs* terms $d = \bar{d}^j$ can be proved in analogy to [14]. In a first step, the solutions of (3.9)–(3.14) are represented as

$$\zeta = S^j v + \zeta_d^j, \quad \eta = \tilde{S}^j v + \eta_d^j, \quad \rho = -\Sigma^j v + r_d^j.$$

The operators $S^j, \tilde{S}^j$ and Σ^j solve the system for $d = 0$ whereas $(\zeta_d^j, \eta_d^j, r_d^j)$ are particular solutions of (3.9)–(3.14) for $v = 0$ and given $d = \bar{d}^j$. With the given notations, it follows from the complementarity system (3.19) that $v = \bar{v}^+$ solves the variational inequality

(LVI') $\qquad (\Sigma^j v - (\sigma^* + r_d^j),\, v' - v) \geq 0$

on $V_\delta = \{v' : |u^* + v'| \leq 1 \text{ a.e. on } [0,1], v' = 0 \text{ a.e. on } [0, t_1 - \delta] \cup [t_2 + \delta, 1]\}$.
Repeating the calculation from [14, Section 3.1] (given originally for Cv) for $\Sigma^j v$ now, an integration by parts leads to

$$(\Sigma^j v, v) = \Omega^j(y,b).$$

Thus, under the conditions from Lemma 3.2, the operator Σ^j is strictly monotone and continuous on the set $V_\delta \subseteq L_2$, and problem (LVI') has a unique solution

$v = \bar{v}^+$; cf. [14, Lemma 3.1]. In order to find related multipliers $\nu_1, \nu_2 \in L_\infty$, consider the auxiliary optimization problem

$$\min \frac{1}{2}(\Sigma^j v, v) - (\sigma^* + r_d^j, v) \qquad \text{s.t. } v \in V_\delta \tag{3.21}$$

and repeat the arguments from the proof of [14, Lemma 3.3].

Dropping details of the proofs, in analogy to [14, Theorem 1] obtain

Theorem 3.3. *Assume* (H0) – (H2) *to hold true at $w = w^*$, and let w^j belong to the set W' from Lemma 3.2. Then there exists a constant $\delta' > 0$ such that, for each $\bar{d}^j \in D := (L_2 \times \mathbb{R}^n)^2 \times L_2 \times \mathbb{R}$ with $\|\bar{d}^j\|_D \leq \delta'$, problem* (LVI$_G^j$) *has a unique solution $(\bar{\xi}^+, \bar{v}^+) \in (W_\infty^1)^5 \times L_\infty$. Moreover, $\nu \geq 0$ a.e. on $[0,1]$, and $(\zeta, \eta, \rho, y, b)$ as elements of $(W_2^1)^3 \times L_2 \times \mathbb{R}$ depend Lipschitz continuously on $d = \bar{d}^j \in D$. The related Lipschitz moduli as well as δ' are independent of j.*

Corollary 3.4. *Let $z = x^{j+1} - x^*, q = p^{j+1} - p^*$ be defined by $z = \zeta + B^j y$ and $q = \eta - P_1^j y$ with related components (ζ, y, η) from $\bar{\xi}^+$. Then, (z, q, y, ν) as elements of $(L_2)^3 \times L_\infty$ as well as the boundary values πz, πq and $b = y(1)$ depend Lipschitz continuously on $d = \bar{d}^j \in D$ with Lipschitz moduli independent of j. (The mapping $\pi : W_2^1 \to \mathbb{R}^n$ denotes the boundary trace operator: $\pi\phi = (\phi(0), \phi(1))$.)*

Remark 3.5. The solution of subproblems (LVI$_G^j$) in practical applications will require further use of finite-dimensional approximation techniques. Since the underlying minimum problems (3.21) are optimization problems in Hilbert space setting, one could use different projection techniques like finite element or spline approximations as well as classical discretization schemes (cf. [15]). Such approach where, first, the continuous problems are approximated by second-order schemes and, afterwards, a discretization is applied, has been successfully used earlier for control problems with smooth optimal control function; see e.g. [1], [2].

3.2. **Discussion of convergence properties.** In a first step, consider the special case of a *semilinear* state equation:

(H3) *The function $g = g(x)$ is a constant vector $g \equiv B \in \mathbb{R}^n$.*

This restriction in particular ensures that the *inexact* Josephy–Newton method given above is equivalent to the *exact* one, i.e., in this case (LVI^j) and (LVI$_G^j$) are equivalent, and $\bar{B}_1^j = B_1^j$, $\bar{M}^j = M^j$ (cf. proof of Lemma 3.1). Moreover, the matrices A^j and C^j are independent of u^j now so that Lemma 3.2 and Theorem 3.3 remain to be valid under the assumption that all elements of $\{w^j\}$ satisfy $\|x^j - x^*\|_\infty + \|p^j - p^*\|_\infty \leq R'$.

To begin with, we will estimate the *rhs* terms $d = \bar{d}^j$ from (3.18) in dependence of $\bar{w} = w^j - w^*$. Direct validation yields

$$\begin{aligned}
d_1 &= \dot{\bar{z}} - A^j \bar{z} - B\bar{v} + f^j + Bu^j - \dot{x}^j - f^* - Bu^* + \dot{x}^*, \\
d_2 &= \bar{z}(0) + a - x^j(0) - a + x^*(0) \;=\; 0, \\
d_3 &= \dot{\bar{q}} - (A^j)^T \bar{q} - C^j \bar{z} - \nabla_x H^j - \dot{p}^j + \nabla_x H^* + \dot{p}^*, \\
d_4 &= \bar{q}(1) + K^j \bar{z}(1) - \nabla k^j - p^j(1) + \nabla k^* + p^*(1), \\
d_5 &= \dot{\bar{\rho}} + (B_1^j)^T \bar{q} + M^j \bar{z} + (p^j)^T [f,g]^j - \dot{\sigma}^j - (p^*)^T [f,g]^* + \dot{\sigma}^*,
\end{aligned}$$

$$d_6 \;=\; \bar\rho(1) - B^T\bar q(1) + B^T p^j - \sigma^j(1) - B^T p^* + \sigma^*(1) \;=\; 0.$$

In semilinear case, the matrices A^j, C^j and vectors H_x^j do not depend on control inputs. If w^j satisfy $\|x^j - x^*\|_\infty + \|p^j - p^*\|_\infty \le R'$ uniformly with the constant R' from Lemma 3.2 then the following pointwise estimates hold uniformly with a constant c independent of j, too:

$$\begin{aligned} |d_1| &= |f^j - f^* - A^j\bar z| \;\le\; c\,|\bar z|^2,\\ |d_3| &= |\nabla_x H^* - \nabla_x H^j - C^j\bar z - (A^j)^T\bar q| \;\le\; c\,(|\bar z|^2 + |\bar q|^2),\\ |d_5| &\le c\,(|\bar z|^2 + |\bar q|^2), \qquad |d_4| \;\le\; c\,|\bar z(1)|^2. \end{aligned}$$

From the last estimates one can derive

$$\begin{aligned} \|\bar d^j\|_D &= \|\bar d_1^j\|_2 + \|\bar d_3^j\|_2 + \|\bar d_5^j\|_2 + |\bar d_4^j| \\ &\le 4c\,\max\{\|\bar z\|_\infty, \|\bar q\|_\infty\}\,(\|\bar z\|_2 + \|\bar q\|_2 + |\bar z(1)|). \end{aligned} \tag{3.22}$$

Based on these preliminaries, we prove the following assertion:

Theorem 3.6. *Suppose the assumptions* (H0)–(H3) *to be fulfilled. Let* $w^0 = (x^0, u^0, p^0, \sigma^0, \mu^0) \in W^1_\infty \times L_\infty \times (W^1_\infty)^3$ *be an initial point such that* $\|x^0 - x^*\|_\infty$ *and* $\|p^0 - p^*\|_\infty$ *are sufficiently small. Then the Josephy–Newton iteration provides a sequence* $\{w^j\}$ *which is enclosed in*

$$W'' := \{w:\ \|x - x^*\|_\infty + \|p - p^*\|_\infty \;\le\; R'\}$$

(with constant R' *given in Lemma 3.2). Moreover,* $\{(x^j, p^j, \pi x^j, \pi p^j)\}$ *converges superlinearly to* $(x^*, p^*, \pi x^*, \pi p^*)$ *in* $L_2 \times L_2 \times \mathbb{R}^{2n} \times \mathbb{R}^{2n}$.

Before proving the theorem, we show an additional uniform estimate for $\{(x^j, p^j)\}$. Here and in the following, we use the notation $\theta^j = (\bar z^j, \pi\bar z^j, \bar q^j, \pi\bar q^j)$ for elements of $\Theta = L_2 \times \mathbb{R}^{2n} \times L_2 \times \mathbb{R}^{2n}$ with $\bar z^j = x^j - x^*, \bar q^j = p^j - p^*$. Further define $\bar y^j$ by $\bar y^j(t) = \int_0^t \bar v^j(s)ds$ with $\bar v^j = u^j - u^*$.

Lemma 3.7. *If* $w^j \in W''$ *and* $\|\bar d^j\|_D \le \delta'$ *then there exist constants* M_0, M_1 *such that* $\|\dot x^{j+1} - \dot x^*\|_2 + \|\dot p^{j+1} - \dot x^*\|_2 \le M_0$, *and*

$$\|\bar x^{j+1} - x^*\|_\infty + \|\bar p^{j+1} - p^*\|_\infty \le M_1(\|\bar x^{j+1} - x^*\|_2 + \|\bar p^{j+1} - p^*\|_2)^{1/2}.$$

The constants herein depend on R', δ' *and the data, but not on counters* j.

Proof. Under the given assumptions, (LVI^j) has a solution w^{j+1} satisfying

$$\|\theta^{j+1}\|_\Theta + \|\bar y^{j+1}\|_2 + |\bar b^{j+1}| \;\le\; \bar c\cdot\|\bar d^j\|_D. \tag{3.23}$$

due to Theorem 3.3. By (3.19), each $u^{j+1} = \bar v^{j+1} + u^*$ is bounded by $\|u^{j+1}\|_\infty \le 1$, and $\|u^{j+1} - u^*\|_\infty \le 2$. Further,

$$\begin{aligned} \dot x^{j+1} &= \dot{\bar z}^{j+1} + \dot x^* \;=\; A^j\bar z^{j+1} + B\bar v^{j+1} + \bar d_1^j + \dot x^*,\\ \dot p^{j+1} &= \dot{\bar q}^{j+1} + \dot p^* \;=\; -(A^j)^T\bar q^{j+1} - C^j\bar z^{j+1} + \bar d_3^j + \dot p^*, \end{aligned}$$

so that, for positive constants c_1, c_2 independent of j,

$$\|\dot x^{j+1}\|_2 + \|\dot p^{j+1}\|_2 \le c_1 + c_2(\|\bar z^{j+1}\|_2 + \|\bar q^{j+1}\|_2) + \|\bar d^j\|_D \le c_1 + (c_2\bar c + 1)\delta'.$$

Consequently, $\dot{\bar z}^{j+1}$ and $\dot{\bar q}^{j+1}$ are bounded in L_2 norm by some constant M_0. The second part of the lemma is a direct consequence of [15, Lemma 13]. □

Proof of Theorem 3.6. Assume $\|x^0 - x^*\|_\infty + \|p^0 - p^*\|_\infty \le \epsilon/3 < R'$. First, it will be shown that, by shrinking ϵ if necessary, the following estimates hold for all $j \ge 0$:

(i) $\|\theta^j\|_\Theta \le \epsilon$, (ii) $\|x^j - x^*\|_\infty + \|p^j - p^*\|_\infty \le R'$, (iii) $\|\bar{d}^j\|_D \le \delta'$.

The relations will be proved by induction: for $j = 0$, the definition of θ^0 yields $\|\theta^0\|_\Theta \le \epsilon$, and (ii) is valid by assumption. Then by (3.22) obtain

$$\|\bar{d}^0\|_D \le 4cR' \, \|\theta^0\|_\Theta \le 4cR' \cdot \epsilon \le \delta', \tag{3.24}$$

or (iii), if only $\epsilon \le \delta'/(4cR')$.

Next assume that (i) – (iii) are satisfied for all $j \le k$: Due to (3.23), (3.22) and Lemma 3.7, for $j = 0, \dots, k$,

$$\begin{aligned} \|\theta^{j+1}\|_\Theta + \|\bar{y}^{j+1}\|_2 + |\bar{b}^{j+1}| \;&\le\; \bar{c}\|\bar{d}^j\|_D \le 4c\bar{c}M_1\|\theta^j\|_\Theta^{3/2} \qquad (3.25)\\ &\le\; 4c\bar{c}M_1\, \epsilon^{1/2}\, \|\theta^j\|_\Theta \\ &=: \; \hat{c}\, \|\theta^j\|_\Theta < \|\theta^j\|_\Theta \le \epsilon \qquad (3.26) \end{aligned}$$

in case $\epsilon^{1/2} < (4c\bar{c}M_1)^{-1}$. Using again Lemma 3.7, we see that

$$\begin{aligned} \|x^{k+1} - x^*\|_\infty + \|p^{k+1} - p^*\|_\infty \;&\le\; M_1(\|\bar{z}^{k+1}\|_2 + \|\bar{q}^{k+1}\|_2)^{1/2} \\ &\le\; M_1\|\theta^{k+1}\|_\Theta^{1/2} \le M_1\epsilon^{1/2} \le R' \end{aligned}$$

if ϵ is taken small enough. In order to derive estimate (iii) for $j = k+1$, as in case $j = 0$ refer to (3.22) and repeat the argumentation from (3.24).

Conditions (ii), (iii) ensure that the Josephy-Newton iteration is well defined for all j. From (3.26) (with $\hat{c} < 1$) we see that $(\theta^j, \bar{y}^j, \bar{b}^j)$ tend to zero in $\Theta \times L_2 \times \mathbb{R}$, and (3.25) yields the superlinear convergence property. □

The last theorem is clearly restricted to the special case that the system is semilinear in the sense of Assumption (H3), and there is no convergence rate available for $\{u^j\}$ in L_1 yet. In general case, the analysis should include influences of control components u^j and the inaccuracy terms as characterized in Lemma 3.1. As it is known from general stability investigation [13, 14], the crucial point is an a-posteriori structural stability analysis for the controls: if one could show that, at least locally, u^j had the same principal bang-singular-bang structure as assumed for u^* and further, prove convergence for the junction times localizations, it should be possible to generalize the superlinear convergence result to the iteration (LVI^j_G). Such generalization mainly requires a detailed estimation of second-order derivatives of iterated switching functions; cf. [14, Section 3.3]. The rigorous convergence proof for (LVI^j_G) in case of a general inclusion (VI) remains subject to future work.

4. Quadratic growth verification

4.1. **Riccati equation approach.** In this section, an auxiliary boundary value problem will be formulated for some matrix function $Q : [0,1] \to \mathbb{R}^{n\times n}$ allowing to guarantee that condition (H2) holds true. Before deriving the Riccati type test problem and proving the related theorem, alternative representations for the quadratic form $\Omega = \Omega(y, b)$ will be found, and conclusions from the quadratic growth condition on Ω will be discussed. In particular, a certain "natural" boundary restriction at t_2 (the end point of the singular arc) is obtained.

In the following, some test point $\tau \in [t_2, 1]$ is used for rewriting the quadratic forms. Although, for theoretical purpose, it would be sufficient to consider $\tau = t_2$, the example in Section 4.2 will show the advantage of possibly varying τ depending on problem data; see also Remark 4.2.

Let us introduce the fundamental matrix solution $\Psi = \Psi(t, s)$ for the linearized canonical system, i.e.

$$\frac{\partial}{\partial t}\Psi(t,s) \,-\, A(t)\Psi(t,s) \,=\, 0, \;\; \frac{\partial}{\partial s}\Psi^T(t,s) \,+\, A^T(s)\Psi^T(t,s) \,=\, 0, \;\; \Psi(t,t) \,=\, I.$$

Then, for y satisfying the restrictions from (H2), $\zeta = \zeta(t)$ has the properties

$$\zeta(t) \,=\, \begin{cases} 0 & \text{if} \quad t \le t_1, \\ \int_{t_1}^{t} \Psi(t,s) B_1(s) y(s)\, ds & \text{if} \quad t > t_1, \end{cases} \tag{4.1}$$

so that, with $\zeta_\tau := \zeta(\tau)$,

$$\zeta(t) = \Psi(t,\tau)\zeta_\tau + \, b \cdot \int_\tau^t \Psi(t,s) B_1(s)\, ds \;=:\; \Psi(t,\tau)\zeta_\tau \,+\, \Phi(t,\tau)\, b \tag{4.2}$$

whenever $t \ge \tau \ge t_2$. In particular, for $\tau \in [t_2, 1]$ denote

$$\zeta(1) \;=:\; \Psi_\tau \zeta_\tau \,+\, \Phi_\tau\, b. \tag{4.3}$$

Remark 4.1. By construction, the vector function $\Phi = \Phi(t, \tau)$ satisfies the differential equation

$$\partial\Phi(t,\tau)/\partial t \,=\, A(t)\Phi(t,\tau) \,+\, B_1(t), \quad \Phi(\tau,\tau) = 0.$$

Thus, Φ_τ can be found as the terminal value $\Phi_\tau = \Phi(1, \tau)$ by solving an additional linear initial value problem.

The boundary terms from (2.13) rewrite as follows:

$$\begin{aligned} \omega(y,b) \;&:=\; \frac{1}{2}\zeta(1)^T K\, \zeta(1) \,+\, \zeta(1)^T W_1 b + \frac{1}{2} V_1 b^2 \\ &=\; \frac{1}{2}\zeta_\tau^T (\Psi_\tau^T K \Psi_\tau)\, \zeta_\tau + \zeta_\tau^T \Psi_\tau^T (K\, \Phi_\tau + W_1) \cdot b \\ &\qquad + \frac{1}{2}(\Phi_\tau^T K \Phi_\tau + V_1 + 2\Phi_\tau^T W_1) \cdot b^2 \\ &=\; \frac{1}{2}\zeta_\tau^T K_\tau \zeta_\tau \,+\, \zeta_\tau W_\tau b \,+\, \frac{1}{2} V_\tau b^2. \end{aligned} \tag{4.4}$$

Notice that for $\tau = 1$ we have $\Psi_\tau = I, \Phi_\tau = 0$ so that (K_τ, W_τ, V_τ) tends to (K, W_1, V_1) for $\tau \to 1$.

For any $\tau \in [t_2, 1]$, the integral term in (2.13) similarly transforms to

$$\begin{aligned} \Omega_I(y,b) \;&:=\; \frac{1}{2}\int_0^1 (\zeta^T C\, \zeta \,+\, 2\zeta^T M^T y \,+\, R\, y^2)\, dt \\ &=\; \frac{1}{2}\int_{t_1}^{\tau} (\zeta^T C\, \zeta \,+\, 2\zeta^T M^T y \,+\, R\, y^2)\, dt \end{aligned} \tag{4.5}$$

$$+\frac{1}{2}\int_{\tau}^{1}(\zeta_\tau^T C_\tau \zeta_\tau + 2\zeta_\tau^T M_\tau^T b + R_\tau b^2)\,dt$$

with $R_\tau(t) = R(t) + \Phi(t,\tau)^T(2M^T(t) + C(t)\Phi(t,\tau))$ and appropriately defined C_τ and M_τ. Combining these expressions, $\Omega = \omega - \Omega_I$ now can be written as the sum of a quadratic term depending on boundary values $(\zeta(\tau), y(\tau)) = (\zeta_\tau, b)$, and a certain integral over $[t_1, \tau] \supseteq [t_1, t_2]$:

$$\begin{aligned}\Omega(y,b) \;&=\; \frac{1}{2}\zeta_\tau^T \tilde{K}_\tau \zeta_\tau + \zeta_\tau \tilde{W}_\tau b + \frac{1}{2}\tilde{V}_\tau b^2 \\ &\qquad -\frac{1}{2}\int_{t_1}^{\tau}(\zeta^T C\,\zeta + 2\zeta^T M^T y + R\,y^2)\,dt.\end{aligned} \tag{4.6}$$

The matrices $\tilde{K}_\tau, \tilde{W}_\tau, \tilde{V}_\tau$ herein are defined as

$$\tilde{K}_\tau = K_\tau - \int_\tau^1 C_\tau(t)\,dt,\;\; \tilde{W}_\tau = W_\tau - \int_\tau^1 M_\tau^T(t)\,dt,\;\; \tilde{V}_\tau = V_\tau - \int_\tau^1 R_\tau(t).$$

Remark 4.2. As functions of τ, K_τ, W_τ and $\tilde{V}_\tau$ solve the linear system

$$\begin{aligned} dK_\tau/d\tau &= -A(\tau)^T K_\tau - K_\tau A(\tau), \\ dW_\tau/d\tau &= -A(\tau)^T W_\tau - K_\tau B_1(\tau), \\ d\tilde{V}_\tau/d\tau &= R(\tau) - 2B_1(\tau)^T W_\tau \end{aligned}$$

on $[t_2, 1]$ with terminal values (K, W_1, V_1) at $\tau = 1$. If, in particular, τ can be chosen equal one, no integration will be needed for their construction.

The following two lemmas contain additional information on coefficients for y^2, b^2 occuring in (4.6):

Lemma 4.3. *Under* (H0) – (H2)*, there exists a positive constant* δ *such that*

$$R(t) = -p(t)^T\,[g,[f,g]]\,(x^*(t)) \;\le\; -\frac{m}{2}$$

for all $t \in [t_1 - \delta, t_2 + \delta]$.

(The given fact is well known; for a short proof see [14, p.567-568].)

Lemma 4.4. *Let* t_1, t_2 *be the boundary points of the singular arc according to* (H1)*, and suppose* (H2) *to hold with the constant* $m > 0$*. Further, let* V_τ, R_τ *be given by* (4.4), (4.5)*, and* $\tilde{V}_\tau$ *introduced in* (4.6)*. Then* $\tilde{V}_{t_2} \ge 2m$.

Proof: Denote by y_s the restriction of y to $[t_1, t_2]$. By (H2), we have

$$\Omega(y,b) \ge m\left(\|y\|_2^2 + b^2\right) \ge m\left(\|y_s\|_2^2 + b^2\right). \tag{4.7}$$

On the other hand, Ω can be estimated via (4.6): taking into acount (4.1), for $\zeta = \zeta(t)$ and $t \in [t_1, t_2]$ obtain $|\zeta(t)| \le c(A,B)\|y_s\|_2$. In case $\tau = t_2$, it further follows from (4.6) that

$$\Omega(y,b) = \frac{1}{2}\tilde{V}_{t_2}\cdot b^2 + \mathrm{O}(\|y_s\|_2^2 + \|y_s\|_2 \cdot |b|) \tag{4.8}$$

for all (y,b) satisfying the restrictions from (H2) and $(y_s, b) \to 0$ in $L_2 \times \mathbb{R}$.

In particular, choose $b = \tilde{b} > 0$ but small enough to ensure $\tilde{b} < m(t_2 - t_1)$, and define

$$\tilde{y}_s(t) \;=\; \max\left\{0,\ \tilde{b} - \int_t^{t_2} (1 - u^0(s))\, ds\right\}.$$

Due to (H1), on $[t_1, t_2]$ the integrand satisfies $0 < m \leq 1 - u^0 \leq 2 - m$, i.e., $\tilde{y}_s$ is monotone increasing. Moreover, there exists some $\theta \in (t_1, t_2)$ such that $\tilde{y}_s = 0,\ t \leq \theta$, and $0 < \tilde{y}_s(t) \leq \tilde{b} + m(t - t_2)$ for $t > \theta$. Hence,

$$\|\tilde{y}_s\|_2^2 \leq \int_\theta^{t_2} (\tilde{b} \,+\, m(t - t_2))^2 dt \;\leq\; \tilde{b}^3/(3m).$$

From (4.8) and (H2) now conclude

$$\Omega(\tilde{y}, \tilde{b}) \;=\; \frac{1}{2}\tilde{V}_{t_2} \cdot \tilde{b}^2 \;+\; \mathrm{O}(\tilde{b}^3 \,+\, \tilde{b}^{5/2}) \;\geq\; m\,\tilde{b}^2 \;+\; \mathrm{O}(\tilde{b}^3).$$

Dividing this relation by $\tilde{b}^2/2$ and tending afterwards $\tilde{b}$ to zero, the assertion of the lemma follows. □

The main result of this section now is formulated as

Theorem 4.5. *Suppose* (H0), (H1) *to hold for problem* (CP) *with the reference solution* (x^*, u^*). *Further assume that there exist a point* $\tau \in [t_2, 1]$ *and positive constants* γ_1, γ_2 *such that*
(i) $R = R(t) \leq -\gamma_1$ *for all* $t \in [t_1, \tau]$,
(ii) $\tilde{V}_\tau \geq \gamma_2$.
Consider the following matrix Riccati differential inequality on $[t_1, \tau]$*:*

$$0 \;\succ\; \dot{Q} + A^T Q + QA + C - R^{-1}(M^T + QB_1)(M + B_1^T Q), \tag{4.9}$$

$$0 \;\prec\; Q(\tau) + \tilde{K}_\tau - \tilde{V}_\tau^{-1}\tilde{W}_\tau \tilde{W}_\tau^T. \tag{4.10}$$

If problem (4.9), (4.10) *has a solution* $Q \in W^1_\infty(t_1, \tau; \mathbb{R}^{n\times n})$, *then* $\Omega = \Omega(y, b)$ *satisfies condition* (H2) *with some positive constant* m.

It should be noticed that conditions (i), (ii) above are necessary for the assertion in the sense that they will hold at least for $\tau = t_2$ if (H2) is fulfilled; cf. Lemmas 4.3, 4.4.

Proof of the theorem: Let us start with the integral term in the representation (4.6) of Ω: since ζ solves the linear equation (2.12), for each differentiable, symmetric $n \times n$-matrix function Q it can be completed to

$$\begin{aligned} I_\tau \;\; := \;\; & \int_{t_1}^{\tau} (\zeta^T C\, \zeta \,+\, 2\zeta^T M^T y \,+\, R\, y^2)\, dt \\ & \quad -2\int_{t_1}^{\tau} \zeta^T Q\, (\dot{\zeta} - A\zeta - B_1 y)\, dt \\ = \;\; & -\zeta_\tau^T Q(\tau)\, \zeta_\tau \,+\, 2\int_{t_1}^{\tau} \zeta^T (M^T + QB_1) y\, dt \\ & + \int_{t_1}^{\tau} \Big(\zeta^T(\dot{Q} + QA + A^T Q + C)\zeta \,+\, R\, y^2\Big)\, dt. \end{aligned}$$

Inserting the expression into formula (4.6) leads to

$$\Omega(y,b) \;=\; \frac{1}{2}(\zeta_\tau^T, b)\, Y_\tau \begin{pmatrix} \zeta_\tau \\ b \end{pmatrix} \;-\; \frac{1}{2}\int_{t_1}^{\tau} (\zeta^T, y)\, Z \begin{pmatrix} \zeta \\ y \end{pmatrix} dt$$

with matrix Y_τ and matrix function Z given by

$$Y_\tau \;=\; \begin{pmatrix} \tilde{K}_\tau + Q(\tau) & \tilde{W}_\tau \\ \tilde{W}_\tau^T & \tilde{V}_\tau \end{pmatrix}, \quad Z \;=\; \begin{pmatrix} \dot{Q} + A^T Q + QA + C & M^T + QB_1 \\ M + B_1^T Q & R \end{pmatrix}.$$

Thus, Ω will satisfy (H2) if $Y_\tau \succ 0$ and $Z = Z(t) \prec 0$ for all $t \in [t_1, \tau]$.
In a first step, consider Z: as a bloc matrix, $Z = Z(t)$ has the lower diagonal element $R = R(t) \le -\gamma_1 < 0$. In order to guarantee the (uniform) negative definiteness of $Z = Z(t)$, remember Schur complements' properties:
The matrix Z with diagonal bloc $R \prec 0$ is negative definite if and only if the Schur complement matrix

$$Z/R \;=\; (\dot{Q} + A^T Q + QA + C) \;-\; (M^T + QB_1)\, R^{-1}(M + B_1^T Q)$$

is negative definite.
The criterion coincides with (4.9).
Secondly, we apply an analogous argument to the bloc matrix Y_τ: under assumption (ii), it is positive definite if and only if (4.10) is fulfilled.
Hence the existence of a bounded matrix solution Q to (4.9), (4.10) is sufficient for the growth condition (H2). □

In practice the inequalities (4.9), (4.10) will be replaced by equations first. If the corresponding boundary value problem has a bounded solution then, by continuity, this remains to be true if we add small (negative or positive definite) perturbations on the left as required for applying the theorem.

4.2. **Test example.** The following problem has been addressed in [14]:

$(\mathbf{P}_\delta)$ $\qquad \min \;\; J(x,u) \;:=\; 0.5\left(x_1^2(T) + x_2^2(T)\right)$

w.r.t.

$$\begin{aligned} \dot{x}_1 &= \cos x_3 - \delta u, & x_1(0) &= a, \\ \dot{x}_2 &= \sin x_3, & x_2(0) &= 0, \\ \dot{x}_3 &= u, & x_3(0) &= \pi/2, \\ |u| &\le 1, & &(T > 0,\; a > 2,\; 0 \le \delta \ll 1). \end{aligned}$$

For the data set $a = 3$, $T = 3.2$, $\delta = 0.15$, an optimal control of bang-singular-bang structure

$$u^0(t) = \begin{cases} 0 & \text{if} \;\; t_1 < t < t_2 \\ 1 & \text{otherwise} \end{cases}$$

with $0 < t_1 < t_2 < T$ was determined numerically in [15]. Using afterwards related piecewise analytic expressions for $x = x(t)$, from transversality and switching conditions

$$\begin{aligned} p_1(T) &= -x_1(T), & p_2(T) &= -x_2(T), \\ \sigma(t_1) &= 0, & \dot{\sigma}(t_1) &= 0, \end{aligned}$$

the switching times $t_1 = 2.1482,\ t_2 = 2.6932$ can be confirmed. Remembering the Pontryagin function

$$H(x,u,p) \ = \ p_1 \cos x_3 + p_2 \sin x_3 + (p_3 - \delta\, p_1)\, u,$$

for the switching fuction obtain the expressions

$$\sigma = p_3 - \delta\, p_1,\ \dot\sigma = p_1 \sin x_3 - p_2 \cos x_3,\ \ddot\sigma = -R\, u,$$

where $R = -(p_1 \cos x_3 + p_2 \sin x_3)$. If the switching function behavior at $t = t_1$ is taken into account, from $x_3(t_1) = \pi/2 + t_1$ we get

$$p_1 : p_2 = -\sin t_1 : \cos t_1 = \ x_1(T) : x_2(T).$$

Thus,

$$R(t) = -\frac{p_2}{\cos t_1}\, (\sin x_3(t) \cos t_1 \ - \ \sin t_1 \cos x_3(t)) \ = \ -\frac{p_2}{\cos t_1}\ \sin(x_3(t) - t_1).$$

From $x_3(t) = \pi/2 + t_1$ for $t \in [t_1, t_2]$, and $x_3(t) = \pi/2 + t_1 - t_2 + t$ for $t \in [t_2, T]$ it follows that

$$R \ = \ \begin{cases} -\frac{p_2}{\cos t_1} & \text{if } t \in [t_1, t_2], \\ -\frac{p_2}{\cos t_1}\ \cos(t - t_2) & \text{if } t \in (t_2, T]. \end{cases}$$

In the given example, $\pi/2 < t_1 < \pi,\ T - t_2 < \pi/2,$, and $p_2 \equiv -x_2(T) < 0$. Therefore,

$$R \le -\frac{p_2}{\cos t_1}\ \cos(T - t_2) \ < \ 0 \qquad \text{on } \ [t_1, T]. \tag{4.11}$$

Next we will discuss the Riccati equation related to Theorem 4.5 for the test function $Q = Q(t)$ on $[t_1, T]$,

$$\begin{aligned} \dot Q + F(Q) \ &= \ 0, \qquad Q(1) \ = \ V_1^{-1} \cdot W_1 W_1^T \ - \ K, \\ \text{with} \qquad F(Q) \ &:= \ A^T Q + QA + C - R^{-1}(M^T + QB_1)(M + B_1^T Q). \end{aligned}$$

Abbreviating $s = \sin x_3,\ c = \cos x_3$, the data in ($\mathrm{P}_\delta$) yield

$$A = \begin{pmatrix} 0 & 0 & -s \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix},\ \ B = \begin{pmatrix} -\delta \\ 0 \\ 1 \end{pmatrix},\ \ B_1 = \begin{pmatrix} -s \\ c \\ 0 \end{pmatrix},$$

together with

$$C = H^0_{xx} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & R \end{pmatrix},\ \ P_1 = H^0_{xu} = \ 0.$$

Moreover, M reduces to $M = B^T C = (0, 0, R)$ so that further $R^{-1} B_1 M = A$, $R^{-1} M^T M = C$. Inserting these expressions into $F(Q)$ leads to

$$\dot Q \ = \ R^{-1}\, Q\, B_1 B_1^T Q. \tag{4.12}$$

For the boundary data obtain

$$K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},\quad W_1 = KB = \begin{pmatrix} -\delta \\ 0 \\ 0 \end{pmatrix},\quad V_1 = B^T W_1 = \delta^2 > 0.$$

Together with (4.11), the last inequality ensures that, for $\tau = 1$, conditions (i), (ii) from Theorem 4.5 hold true. The boundary condition on Q counts as

$$Q(1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let us assume that $Q(t) = (q_{ij}(t))$ with $q_{ij} = q(t)$ for $i = j = 2$, and $q_{ij} \equiv 0$ elsewhere. Then, $B_1^T Q$ writes as

$$B_1^T Q = (0, q\,c, 0),$$

and the boundary value problem for Q on $[t_1, T]$ reduces to

$$\dot{q} = \frac{\cos^2 x_3}{R}\, q^2, \qquad q(1) = -1. \tag{4.13}$$

As long as $q \neq 0$ one can write

$$\frac{d}{dt}\left(\frac{1}{q(t)}\right) = -\frac{\cos^2 x_3(t)}{R(t)} > 0.$$

For the solution find

$$\frac{1}{q(t)} = \int_t^T \frac{\cos^2 x_3(s)}{R(s)}\, ds - 1 \leq -1 < 0 \qquad \forall\, t \in [t_1, T]$$

and consequently, $-1 \leq q(t) < 0$. Hence, problem (4.13) has a bounded solution on $[t_1, T]$, and the same is true for (4.9), (4.10). By Theorem 4.5, for the solution (x^0, u^0) of the problem, the growth condition from Proposition 2.1 is valid.

Acknowledgement

The author is grateful to the anonymous referee for important hints and remarks which helped to improve this presentation.

References

[1] W. Alt, *Discretization and mesh-independence of Newton's method for generalized equations*, in: Mathematical Programming with Data Perturbation, A. V. Fiacco (ed.), Lect. Notes Pure Appl. Mathem, vol. 195, Marcel Dekker Inc., New York, 1997, pp. 1–30.

[2] W. Alt, *Mesh-independence of the Lagrange-Newton method for nonlinear optimal control problems and their discretizations*, in: Optimization with Data Perturbation II, D. Ward, D. Klatte and J. Rückmann (eds.), Ann. Oper. Res.,vol. 101, Kluwer Acad. Publ., Dordrecht, 2001, pp. 101–117.

[3] W. Alt and M. Seydenschwanz, *Regularization and discretization of linear-quadratic control problems*, Control & Cybernetics **40**(4) (2011), 903–920.

[4] M. S. Aronna, J. F. Bonnans, A. V. Dmitruk and P. A. Lotito, *Quadratic order conditions for bang-singular extremals*, Numerical Algebra, Control and Optimization **2** (2012), 511–546.

[5] M. S. Aronna, J. F. Bonnans and P. Martinon, *A shooting algorithm for optimal control problems with singular arcs*, J. Optim. Theory Appl. **158** (2013), 419–459.

[6] D. Augustin and H. Maurer, *Computational sensitivity analysis for state constrained optimal control problems*, in: Optimization with Data Perturbation II, D. Ward, D. Klatte and J. Rückmann (eds.), Ann. Oper. Res., vol. 101, Kluwer Acad. Publ., Dordrecht, 2001, pp. 75–99.

[7] A. V. Dmitruk, *Quadratic conditions for a weak minimium for singular regimes in optimal control problems*, Soviet Math. Doklady **18** (1977).

[8] A. V. Dmitruk, *Conditions of Jacobi type for Bolza's problem with inequalities*, Math. Notes **35** (1984), 427–435.

[9] A. V. Dmitruk, *Jacobi type conditions for singular extremals*, Control & Cybernetics **37** (2008), 285–306.

[10] A. V. Dmitruk and A. A. Milyutin, *A condition of Legendre type for optimal control problems, linear in control. Calculus of variations and optimal control*, in: Calculus of Variations and Optimal Control, Haifa, 1998, A. Ioffe, S. Reich and I. Shafrir (eds.), Chapman & Hall/CRC, Boca Raton, FL, 2000, pp. 49–61.

[11] A. L. Dontchev, *Newton's method for generalized equations: a sequential implicit function theorem*, Math. Program., Ser. B **123** (2010), 139–159.

[12] A. L. Dontchev and R. T. Rockafellar, *Convergence of inexact Newton methods for generalized equations*, Math. Program., Ser. B **139** (2013), 115–137.

[13] U. Felgenhauer, *Structural stability investigation of bang-singular-bang optimal controls*, J. Optim. Theory Appl. **152** (2012), 605–631.

[14] U. Felgenhauer, *Stability analysis of variational inequalities for bang-singular-bang controls*, Control & Cybernetics **42** (2013), 557–592.

[15] U. Felgenhauer, *Discretization of semilinear bang-singular-bang control problems*, Comput. Optim. Appl. **64** (2016), 295–326.

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

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

[18] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, Academic Press, New York, 1980.

[19] H. W. Knobloch, *Higher Order Necessary Conditions in Optimal Control Theory*, Lecture Notes Contr. Inf. Sci. **34**, Springer, Berlin, 1981.

[20] A. J. Krener, *The high order maximum principle and its application to singular extremals*, SIAM J. Control Optim. **15** (1977), 256–292.

[21] H. Maurer and H. J. Oberle, *Second order sufficient conditions for optimal control problems with free final time: The Riccati approach*, SIAM J. Control **41** (2002), 380–403.

[22] H. Maurer and N. P. Osmolovskii, *Second order sufficient conditions for time-optimal bang-bang control*, SIAM J. Control **42** (2004), 2239–2263.

[23] A. A. Milyutin and N. P. Osmolovskii, *Calculus of variations and optimal control*, Amer. Mathem. Soc., Providence, Rhode Island, 1998.

[24] L. Poggiolini and G. Stefani, *Sufficient optimality conditions for a bang-singular extremal in the minimum time problem*, Control & Cybernetics **37** (2008), 469–490.

[25] L. Poggiolini and G. Stefani, *Bang-singular-bang extremals: Sufficient optimality conditions*, J. Dyn. Contr. Systems **17** (2011), 469–514.

[26] L. Poggiolini and G. Stefani, *Structural stability for bang-singular-bang extremals in the minimum time problem*, SIAM J. Control Optim. **51** (2013), 3511–3531.

[27] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res. **5** (1980), 43–62.

[28] G. Stefani, *Strong optimality of singular trajectories*, in: Geometric Control and Nonsmooth Analysis, F. Ancona et al. (eds.), Series Adv. Math. Appl. Sci. vol. 76 World Sci. Publ., Hackensack, NJ, 2008, pp. 300–327.

[29] G. Vossen, *Numerische Lösungsmethoden, hinreichende Optimalitätsbedingungen und Sensitivitätsanalyse für optimale bang-bang und singuläre Steuerungen*, PhD Thesis, Faculty of Mathematics and Science, University Münster, Germany, 2005, (195 pp).

[30] G. Vossen, *Switching time optimization for bang-bang and singular controls*, J. Optim. Theory Appl. **144** (2010), 409–429.

Manuscript received 9 December 2015
revised 8 March 2016

U. Felgenhauer
Institut für Angewandte Mathematik und Wissenschaftliches Rechnen, Brandenburgische Technische Universität Cottbus – Senftenberg, Germany
E-mail address: felgenh@b-tu.de