JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 130, No. 3, pp. 441–460, September 2006 (© 2006)
DOI: 10.1007/s10957-006-9121-9

Local Minimum Principle for Optimal Control Problems Subject to Differential-Algebraic Equations of Index Two[1]

M. GERDTS[2]

Communicated by H. J. Pesch

Published Online: 11 November 2006

Abstract. Necessary conditions in terms of a local minimum principle are derived for optimal control problems subject to index-2 differential-algebraic equations, pure state constraints, and mixed control-state constraints. Differential-algebraic equations are composite systems of differential equations and algebraic equations, which arise frequently in practical applications. The local minimum principle is based on the necessary optimality conditions for general infinite optimization problems. The special structure of the optimal control problem under consideration is exploited and allows us to obtain more regular representations for the multipliers involved. An additional Mangasarian-Fromowitz-like constraint qualification for the optimal control problem ensures the regularity of a local minimum. An illustrative example completes the article.

Key Words. Optimal control, necessary conditions, local minimum principle, index two differential-algebraic equations, state constraints.

1. Introduction

Optimal control problems subject to ordinary differential equations have a wide range of applications in different disciplines like engineering sciences, chemical engineering, and economics. Necessary conditions, known as maximum principles or minimum principles, have been investigated intensively since the 1950s. Early proofs of the maximum principle were given by Pontryagin (Ref. 1) and Hestenes (Ref. 2). Necessary conditions with pure state constraints were discussed in e.g. Jacobsen et al. (Ref. 3), Girsanov (Ref. 4), Knobloch (Ref. 5),

[1] The author thanks the referees for careful reading and helpful suggestions and comments.
[2] Professor, Department of Mathematics, University of Hamburg, Hamburg, Germany.

Maurer (Refs. 6, 7), Ioffe and Tihomirov (Ref. 8), and Kreindler (Ref. 9). Neustadt (Ref. 10) and Zeidan (Ref. 11) discussed optimal control problems with mixed control-state constraints. Hartl et al. (Ref. 12) provided a survey on maximum principles for optimal control problems with state constraints including an extensive list of references. Necessary conditions for variational problems (i.e., smooth optimal control problems) were developed in Bryson and Ho (Ref. 13). Second-order necessary conditions and sufficient conditions were stated in Zeidan (Ref. 11). Sufficient conditions were also presented in Maurer (Ref. 14), Malanowski (Ref. 15), Maurer and Pickenhain (Ref. 16), and Malanowski et al. (Ref. 17). Necessary conditions for optimal control problems subject to index-1 DAE systems without state constraints and without mixed control-state constraints can be found in de Pinho and Vinter (Ref. 18). Implicit control systems were discussed in Devdariani and Ledyaev (Ref. 19).

Necessary conditions are not only interesting from a theoretical point of view, but also provide the basis of the so-called indirect approach for solving optimal control problems numerically. In this approach, the minimum principle is exploited and usually leads to a multipoint boundary-value problem, which is solved numerically by e.g. the multiple-shooting method.

We extend the results for optimal control problems with ordinary differential equations to problems involving differential-algebraic equations (DAEs). Differential-algebraic equations are composite systems of differential equations and algebraic equations. Particularly, we discuss DAE systems of the type

$$\dot{x}(t) = f(t, x(t), y(t), u(t)), \quad \text{a.e. in } [t_0, t_f], \tag{1}$$

$$0_{n_y} = g(t, x(t)), \quad \text{a.e. in } [t_0, t_f]. \tag{2}$$

Herein, $x(t) \in \mathbb{R}^{n_x}$ is referred to as differential variable and $y(t) \in \mathbb{R}^{n_y}$ as algebraic variable. Correspondingly, (1) is called differential equation and (2) is called algebraic equation. The variable u is a control variable, which is an external function and allows us to control the system in an appropriate way. In this paper, we restrict the discussion to so-called semiexplicit index-2 DAE systems, which are characterized by the following assumption.

Assumption 1.1. The DAE (1)–(2) has index 2; i.e., the matrix $M := g'_x \cdot f'_y$ is nonsingular a.e. in $[t_0, t_f]$ and M^{-1} is essentially bounded in $[t_0, t_f]$.

Due to the algebraic equation, an initial value $(x(t_0), y(t_0))$ has to satisfy the algebraic constraint (2) as well as the derivatives thereof; i.e., the initial values have to be consistent. More precisely, for semiexplicit index-2 DAE systems, $(x(t_0), y(t_0))$ is called consistent if $x(t_0)$ satisfies (2) at $t = t_0$ and $y(t_0)$ satisfies

$$0_{n_y} = g'_t(t_0, x(t_0)) + g'_x(t_0, x(t_0)) \cdot f(t_0, x(t_0), y(t_0), u(t_0)),$$

which is the derivative w.r.t. time of (2) at $t = t_0$. Semiexplicit index 2 systems occur often in mechanical engineering [cf. e.g. Gear et al. (Ref. 20), Gerdts (Refs. 21, 22)].

We investigate nonlinear optimal control problems of the subsequent form. Let $[t_0, t_f] \subset \mathbb{R}$ be a nonempty and bounded interval with fixed time points $t_0 < t_f$ and let $\mathcal{U} \subseteq \mathbb{R}^{n_u}$ be a closed and convex set with nonempty interior. Let $\varphi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}$, $f_0 : [t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}$, $f : [t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x}$, $g : [t_0, t_f] \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$, $\psi : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_\psi}$, $c : [t_0, t_f] \times \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_c}$, $s : [t_0, t_f] \times \mathbb{R}^{n_x} \to \mathbb{R}^{n_s}$ be mappings. For $n \in \mathbb{N}$, the Banach space $L^\infty([t_0, t_f], \mathbb{R}^n)$ consists of all measurable functions $h : [t_0, t_f] \to \mathbb{R}^n$ with

$$\|h\|_\infty := \operatorname*{ess\,sup}_{t_0 \le t \le t_f} \|h(t)\| < \infty,$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^n$. The Banach space $W^{1,\infty}([t_0, t_f], \mathbb{R}^n)$ consists of all absolutely continuous functions $h : [t_0, t_f] \to \mathbb{R}^n$ with

$$\|h\|_{1,\infty} := \max\{\|h\|_\infty, \|h'\|_\infty\} < \infty,$$

where h' denotes the first derivative of h. The space of functions of bounded variation is denoted by $\mathrm{BV}([t_0, t_f], \mathbb{R}^n)$. The space $\mathrm{NBV}([t_0, t_f], \mathbb{R}^n)$ of normalized functions of bounded variation consists of all functions $\mu \in \mathrm{BV}([t_0, t_f], \mathbb{R}^n)$ which are continuous from the right in (t_0, t_f) and satisfy $\mu(t_0) = 0$. We denote the null element of a general Banach space X by Θ_X or, if no confusion is possible, simply by Θ. We use 0_n if $X = \mathbb{R}^n$ and 0 if $X = \mathbb{R}$. We consider the following problem.

Problem 1.1. Higher-Index DAE Optimal Control Problem. Find an absolutely continuous state variable $x \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x})$, an essentially bounded algebraic variable $y \in L^\infty([t_0, t_f], \mathbb{R}^{n_y})$, and an essentially bounded control variable $u \in L^\infty([t_0, t_f], \mathbb{R}^{n_u})$ such that the objective function

$$F(x, y, u) := \varphi(x(t_0), x(t_f)) + \int_{t_0}^{t_f} f_0(t, x(t), y(t), u(t))\, dt \tag{3}$$

is minimized subject to the semiexplicit differential algebraic equation (DAE) (1)–(2), the boundary conditions

$$\psi(x(t_0), x(t_f)) = 0_{n_\psi}, \tag{4}$$

the mixed control-state constraints

$$c(t, x(t), y(t), u(t)) \le 0_{n_c}, \quad \text{a.e. in } [t_0, t_f], \tag{5}$$

the pure state constraints

$$s(t, x(t)) \leq 0_{n_s}, \quad \text{in } [t_0, t_f], \tag{6}$$

and the set constraints

$$u(t) \in \mathcal{U}, \quad \text{a.e. in } [t_0, t_f]. \tag{7}$$

Some definitions and terminologies are in order. $(x, y, u) \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x}) \times L^{\infty}([t_0, t_f], \mathbb{R}^{n_y}) \times L^{\infty}([t_0, t_f], \mathbb{R}^{n_u})$ is called admissible or feasible for Problem 1.1 if the constraints (1)–(2) and (4)–(7) are fulfilled. An admissible pair $(\hat{x}, \hat{y}, \hat{u})$ is called weak [resp. strong] local minimum of Problem 1.1 if there exists $\varepsilon > 0$ such that $F(\hat{x}, \hat{y}, \hat{u}) \leq F(x, y, u)$ holds for all admissible (x, y, u) with $\|x - \hat{x}\|_{1,\infty} < \varepsilon$, $\|y - \hat{y}\|_{\infty} < \varepsilon$, and $\|u - \hat{u}\|_{\infty} < \varepsilon$ [resp. for all admissible (x, y, u) with $\|x - \hat{x}\|_{\infty} < \varepsilon$]. Notice that strong local minima are also weak local minima. The converse is not true. Strong local minima are minimal w.r.t. a larger class of algebraic variables and controls. Weak local minima are optimal only w.r.t. all the algebraic variables and controls in a L^{∞}-neighborhood.

The paper is organized as follows. In Section 2, necessary conditions and Lagrange multiplier representations known from a previous article are summarized. Based on these conditions, the local minimum principle for Problem 1.1 is stated in Section 3. In addition, some important special cases are discussed. Section 4 summarizes a constraint qualification which ensures the regularity of a local minimum. Finally, an example and some concluding remarks close the paper.

2. Abstract Necessary Conditions and Multiplier Representation

The spaces $X := W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x}) \times L^{\infty}([t_0, t_f], \mathbb{R}^{n_y}) \times L^{\infty}([t_0, t_f], \mathbb{R}^{n_u})$ endowed with

$$\|(x, y, u)\|_X := \max\{\|x\|_{1,\infty}, \|y\|_{\infty}, \|u\|_{\infty}\},$$

$Y := L^{\infty}([t_0, t_f], \mathbb{R}^{n_c}) \times C([t_0, t_f], \mathbb{R}^{n_s})$ endowed with

$$\|(y_1, y_2)\|_Y := \max\{\|y_1\|_{\infty}, \|y_2\|_{\infty}\},$$

and $Z := L^{\infty}([t_0, t_f], \mathbb{R}^{n_x}) \times W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_y}) \times \mathbb{R}^{n_\psi}$ endowed with

$$\|(z_1, z_2, z_3)\|_Z := \max\{\|z_1\|_{\infty}, \|z_2\|_{1,\infty}, \|z_3\|_2\}$$

are Banach spaces. The topological dual spaces are denoted by X^*, Y^*, Z^*. The objective function and the constraints of Problem 1.1 are mappings from X into

$\mathbb{R}$, Y, Z respectively and are defined by

$$F(x, y, u) := \varphi(x(t_0), x(t_f)) + \int_{t_0}^{t_f} f_0(t, x(t), y(t), u(t))\, dt,$$
$$G(x, y, u) := (-c(\cdot, x(\cdot), y(\cdot), u(\cdot)), -s(\cdot, x(\cdot))),$$
$$H(x, y, u) := (f(\cdot, x(\cdot), y(\cdot), u(\cdot)) - \dot{x}(\cdot), g(\cdot, x(\cdot)), -\psi(x(t_0), x(t_f))).$$

If the functions φ, f_0, f, g, ψ, c, s are continuous w.r.t. all the arguments and continuously differentiable w.r.t. to x, y, u, then the function F is Fréchet-differentiable and G, H are continuously Fréchet-differentiable [cf. Kirsch et al. (Ref. 23, pp. 94–95), Mukesh Gerdts (Ref. 24)]. The cone $K := K_1 \times K_2 \subseteq Y$ with

$$K_1 := \{z \in L^\infty([t_0, t_f]), \mathbb{R}^{n_c}) \mid z(t) \geq 0_{n_c}, \text{ a.e. in } [t_0, t_f]\},$$
$$K_2 := \{z \in C([t_0, t_f]), \mathbb{R}^{n_s}) \mid z(t) \geq 0_{n_s}, \text{ in } [t_0, t_f]\},$$

is closed, convex, and has nonempty interior. The positive dual cone of K is defined to be

$$K^+ := \{y^* \in Y^* \mid y^*(y) \geq 0, \text{ for all } y \in K\}.$$

The set

$$S := W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x}) \times L^\infty([t_0, t_f], \mathbb{R}^{n_y}) \times U_{\text{ad}},$$

with

$$U_{\text{ad}} := \{u \in L^\infty([t_0, t_f], \mathbb{R}^{n_u}) \mid u(t) \in \mathcal{U}, \text{ a.e. in } [t_0, t_f]\},$$

is closed, convex, and has nonempty interior. With these definitions, the Problem 1.1 is equivalent to the subsequent infinite optimization problem.

Problem 2.1. Find $(x, y, u) \in X$ such that $F(x, y, u)$ is minimized subject to the constraints $G(x, y, u) \in K$, $H(x, y, u) = \Theta z$, $(x, y, u) \in S$.

The following necessary conditions hold; see Gerdts [Ref. 24, Theorem 3.1, Eqs. (19)–(21)].

Theorem 2.1. Necessary Conditions. Let the following assumptions be fulfilled:

(i) Let φ, f_0, f, ψ, c, s be continuous w.r.t. all the arguments and continuously differentiable w.r.t. x, y, u. Let g be continuously differentiable w.r.t. all the arguments.

(ii) Let $U \subseteq \mathbb{R}^{n_u}$ be a closed and convex set with nonempty interior.

(iii) Let $(\hat{x}, \hat{y}, \hat{u}) \in X$ be a weak local minimum of the optimal control problem.

(iv) Let Assumption 1.1 be valid.

Then, there exist nontrivial multipliers $l_0 \in \mathbb{R}$, $\eta^* = (\eta_1^*, \eta_2^*) \in Y^*$, $\lambda^* = (\lambda_f^*, \lambda_g^*, \sigma) \in Z^*$ with

$$l_0 \geq 0, \; \eta^* \in K^+, \eta^*(G(\hat{x}, \hat{y}, \hat{u}) = 0, \tag{8}$$

$$l_0 F_x'(\hat{x}, \hat{y}, \hat{u})(\delta x) - \eta^*(G_x'(\hat{x}, \hat{y}, \hat{u})(\delta x)) - \lambda^*(H_x'(\hat{x}, \hat{y}, \hat{u})(\delta x)) = 0, \tag{9}$$

$$l_0 F_y'(\hat{x}, \hat{y}, \hat{u})(\delta y) - \eta^*(G_y'(\hat{x}, \hat{y}, \hat{u})(\delta y)) - \lambda^*(H_y'(\hat{x}, \hat{y}, \hat{u})(\delta y)) = 0, \tag{10}$$

$$l_0 F_u'(\hat{x}, \hat{y}, \hat{u})(\delta u) - \eta^*(G_u'(\hat{x}, \hat{y}, \hat{u})(\delta u)) - \lambda^*(H_u'(\hat{x}, \hat{y}, \hat{u})(\delta u)) \geq 0, \tag{11}$$

for all $(\delta x, \delta y, \delta u) \in S - \{(\hat{x}, \hat{y}, \hat{u})\}$.

For notational convenience, throughout this article we use the abbreviations

$$\varphi_{x_0}' := \varphi_{x_0}'(\hat{x}(t_0), \hat{x}(t_f)), \quad f_x'[t] := f_x'(t, \hat{x}(t), \hat{y}(t), \hat{u}(t)),$$

and in a similar way φ_{x_f}', $f_{0,x}'[t]$, $f_{0,y}'[t]$, $f_{0,u}'[t]$, $c_x'[t]$, $c_y'[t]$, $c_u'[t]$, $s_x'[t]$, $f_y'[t]$, $f_u'[t]$, $g_t'[t]$, $g_x'[t]$, ψ_{x_0}', ψ_{x_f}' for all respective derivatives.

Under the assumptions of Theorem 2.1, Theorems 4.1 and 4.2 and equation (18) in Gerdts (Ref. 24) provide explicit representations of the multipliers $\eta_1^*, \eta_2^*, \lambda_f^*, \lambda_g^*$ if either (i) there are no mixed control-state constraints (5), or (ii) it holds that $\mathcal{U} = \mathbb{R}^{n_u}$ and rank $(c_u'[t]) = n_c$ a.e. in $[t_0, t_f]$, and the matrix

$$\hat{M}(t) := g_x'[t] \cdot (f_y'[t] - f_u'[t](c_u'[t])^+ c_y'[t]) \tag{12}$$

is nonsingular with essentially bounded inverse $\hat{M}^{-1}$ a.e. in $[t_0, t_f]$, where $(c_u'[t])^+ := c_u'[t]^\top (c_u'[t] c_u'[t]^\top)^{-1}$ denotes the pseudoinverse of $c_u'[t]$.

In each of the two cases (i) and (ii), there exist $\zeta \in \mathbb{R}^{n_y}$ and functions $\mu(\cdot) \in NBV([t_0, t_f], \mathbb{R}^{n_s})$, $\lambda_f(\cdot) \in BV([t_0, t_f], \mathbb{R}^{n_x})$, $\lambda_g(\cdot) \in L^\infty([t_0, t_f], \mathbb{R}^{n_y})$, with

$$\lambda_f^*(h_1(\cdot)) = -\int_{t_0}^{t_f} (\lambda_f(t)^\top + \lambda_g(t)^\top g_x'[t]) h_1(t) dt, \tag{13}$$

$$\lambda_g^*(h_2(\cdot)) = -\zeta^\top h_2(t_0) - \int_{t_0}^{t_f} \lambda_g(t)^\top \dot{h}_2(t) dt, \tag{14}$$

$$\eta_2^*(h_3(\cdot)) = \sum_{i=1}^{n_s} \int_{t_0}^{t_f} h_{3,i}(t) d\mu_i(t), \tag{15}$$

for every $h_1 \in L^\infty([t_0, t_f], \mathbb{R}^{n_x})$, $h_2 \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_y})$, and $h_3 \in C([t_0, t_f], \mathbb{R}^{n_s})$. The latter integral is a Riemann-Stieltjes integral. In case (ii), in addition, there exists $\eta(\cdot) \in L^\infty([t_0, t_f], \mathbb{R}^{n_c})$ with

$$\eta_1^*(k(\cdot)) = \int_{t_0}^{t_f} \eta(t)^\top k(t) dt, \tag{16}$$

for every $k \in L^\infty([t_0, t_f], \mathbb{R}^{n_c})$.

3. Local Minimum Principles

Theorem 2.1 and the multiplier representations (13)–(16) are the basis for the upcoming local minimum principles for Problem 1.1. The Hamiltonian function is defined by

$$\mathcal{H}(t, x, y, u, \lambda_f, \lambda_g, l_0) := l_0 f_0(t, x, y, u) + \lambda_f^\top f(t, x, y, u) + \lambda_g^\top (g_t'(t, x) + g_x'(t, x) f(t, x, y, u)). \qquad (17)$$

Notice that this Hamiltonian function does not use the algebraic constraint (2), but its time derivative.

Theorem 3.1. Local Minimum Principle for Control Problems without Mixed Control-State Constraints. Let the assumptions of Theorem 2.1 be fulfilled. In addition, let g be twice continuously differentiable w.r.t. all the arguments and let there be no mixed control-state constraints (5) in Problem 1.1. Then, there exist multipliers $l_0 \in \mathbb{R}$, $\lambda_f \in \mathrm{BV}([t_0, t_f], \mathbb{R}^{n_x})$, $\lambda_g \in L^\infty([t_0, t_f], \mathbb{R}^{n_y})$, $\mu \in$ *NBV* $([t_0, t_f], \mathbb{R}^{n_s})$, $\zeta \in \mathbb{R}^{n_y}$, and $\sigma \in \mathbb{R}^{n_\psi}$ such that the following conditions are satisfied:

(i) $l_0 \geq 0$, $(l_0, \zeta, \sigma, \lambda_f, \lambda_g, \mu) \neq \Theta$;

(ii) adjoint equations

$$\lambda_f(t) = \lambda_f(t_f) + \int_t^{t_f} \mathcal{H}_x'(\tau, \hat{x}(\tau), \hat{y}(\tau), \hat{u}(\tau), \lambda_f(\tau), \lambda_g(\tau), l_0)^\top d\tau + \sum_{i=1}^{n_s} \int_t^{t_f} s_{i,x}'(\tau, \hat{x}(\tau))^\top d\mu_i(\tau), \quad \text{in } [t_0, t_f], \qquad (18)$$

$$0_{n_y} = \mathcal{H}'_y(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), l_0)^\top \quad \text{a.e. in} [t_0, t_f]; \qquad (19)$$

(iii) transversality conditions

$$\lambda_f(t_0)^\top = -(l_0 \varphi_{x_0}'(\hat{x}(t_0), \hat{x}(t_f)) + \sigma^\top \psi_{x_0}'(\hat{x}(t_0), \hat{x}(t_f)) + \zeta^\top g_x'(t_0, \hat{x}(t_0))), \qquad (20)$$

$$\lambda_f(t_f)^\top = (l_0 \varphi_{x_f}'(\hat{x}(t_0), \hat{x}(t_f)) + \sigma^\top \psi_{x_f}'(\hat{x}(t_0), \hat{x}(t_f)); \qquad (21)$$

(iv) optimality condition: almost everywhere in $[t_0, t_f]$ for all $u \in \mathcal{U}$, it holds that

$$\mathcal{H}'_u(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), l_0)(u - \hat{u}(t)) \geq 0; \qquad (22)$$

(v) complementarity condition: μ_i is monotonically increasing on $[t_0, t_f]$ and constant on every interval (t_1, t_2) with $t_1 < t_2$ and $s_i(t, \hat{x}(t)) < 0$ for all $t \in (t_1, t_2)$.

Proof. By the assumptions, the assertions of Theorem 2.1 hold. In particular, (9)–(11) hold with the multiplier representations (13)–(15).

(a) Equation (9) is equivalent to

$$(l_0\varphi'_{x_0} + \sigma^\top\psi'_{x_0} + \zeta^\top g'_x[t_0])\delta x(t_0) + (l_0\varphi'_{x_f} + \sigma^\top\psi'_{x_f})\delta x(t_f) + \int_{t_0}^{t_f} l_0 f'_{0,x}[t]\delta x(t)\,dt + \int_{t_0}^{t_f} (\lambda_f(t)^\top + \lambda_g(t)^\top g'_x[t])\,(f'_x[t]\delta x(t) - \delta\dot{x}(t))\,dt + \int_{t_0}^{t_f} \lambda g(t)^\top (d/dt)(g'_x[t]\delta x(t))dt + \sum_{i=1}^{n_s}\int_{t_0}^{t_f} s'_{i,x}[t]\delta x(t)d\mu_i(t) = 0,$$

for all $\delta x \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x})$. Exploitation of

$$(d/dt)(g'_x[t]\delta x(t)) = ((d/dt)g'_x[t])\delta x(t) + g'_x[t]\delta\dot{x}(t)$$

yields

$$(l_0\varphi'_{x_0} + \sigma^\top\psi'_{x_0} + \zeta^\top g'_x[t_0])\delta x(t_0) + (l_0\varphi'_{x_f} + \sigma^\top\psi'_{x_f})\delta x(t_f) + \int_{t_0}^{t_f} l_0 f'_{0,x}[t]\delta x(t)dt + \int_{t_0}^{t_f} \lambda_f(t)^\top (f'_x[t]\delta x(t) - \delta\dot{x}\,(t))dt + \int_{t_0}^{t_f} \lambda_g(t)^\top \left(g'_x[t]f'_x[t] + (d/dt)g'_x[t]\right)\delta x(t)dt + \sum_{i=1}^{n_s}\int_{t_0}^{t_f} s'_{i,x}[t]\delta x(t)d\mu_i(t) = 0,$$

which can be written equivalently as

$$(l_0\varphi'_{x_0} + \sigma^\top\psi'_{x_0} + \zeta^\top g'_x[t_0])\,\delta x(t_0) + (l_0\varphi'_{x_f} + \sigma^\top\psi'_{x_f})\,\delta x(t_f) + \int_{t_0}^{t_f} \mathcal{H}'_x[t]\delta x(t)dt + \sum_{i=1}^{n_s}\int_{t_0}^{t_f} s'_{i,x}[t]\delta x(t)d\mu_i(t) - \int_{t_0}^{t_f} \lambda_f(t)^\top\delta\dot{x}(t)\,dt = 0,$$

for all $\delta x \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x})$. Application of the computation rules for Stieltjes integrals yields

$$(l_0\varphi'_{x_0} + \sigma^\top\psi'_{x_0} + \zeta^\top g'_x[t_0])\,\delta x(t_0) + (l_0\varphi'_{x_f} + \sigma^\top\psi'_{x_f})\,\delta x(t_f) + \int_{t_0}^{t_f} \mathcal{H}'_x[t]\delta x(t)dt + \sum_{i=1}^{n_s}\int_{t_0}^{t_f} s'_{i,x}[t]\,\delta x(t)d\mu_i(t) - \int_{t_0}^{t_f} \lambda(t)^\top d\delta x(t) = 0.$$

Integration by parts of the last term leads to

$$(l_0\varphi'_{x_0} + \sigma^\top\psi'_{x_0} + \zeta^\top g'_x[t_0] + \lambda(t_0)^\top)\delta x(t_0) + (l_0\varphi'_{x_f} + \sigma^\top\psi'_{x_f} - \lambda(t_f)^\top)\,\delta x(t_f) + \int_{t_0}^{t_f} \mathcal{H}'_x[t]\delta x(t)\,dt + \sum_{i=1}^{n_s}\int_{t_0}^{t_f} s'_{i,x}[t]\delta x(t)d\mu_i(t) + \int_{t_0}^{t_f}\delta x(t)^\top d\lambda(t) = 0,$$

for all $\delta x \in W^{1,\infty}([t_0,t_f],\mathbb{R}^{n_x})$. This is equivalent to

$$(l_0\varphi'_{x_0} + \sigma^\top\psi'_{x_0} + \zeta^\top g'_x[t_0] + \lambda(t_0)^\top)\delta x(t_o) + (l_0\varphi'_{x_f} + \sigma^\top\varphi'_{x_f} - \lambda(t_f)^\top)\delta x(t_f) + \int_{t_0}^{t_f}\delta x(t)^\top d\left(\lambda(t) - \int_t^{t_f}\mathcal{H}'_x[\tau]^\top d\tau - \sum_{i=1}^{n_s}\int_t^{t_f} s'_{i,x}[\tau]^\top d\mu_i(\tau)\right) = 0,$$

for all $\delta x \in W^{1,\infty}([t_0,t_f],\mathbb{R}^{n_x})$. This implies (20)–(21) and

$$C = \lambda(t) - \int_t^{t_f}\mathcal{H}'_x[\tau]^\top d\tau - \sum_{i=1}^{n_s}\int_t^{t_f} s'_{i,x}[\tau]^\top d\mu_i(\tau),$$

for some constant vector C. Evaluation of the last equation at $t = t_f$ yields $C = \lambda(t_f)$ and thus (18).

(b) Equation (10) is equivalent to

$$\int_{t_0}^{t_f}\mathcal{H}'_y[t]\delta y(t)\,dt = 0,$$

for all $\delta y \in L^\infty([t_0,t_f],\mathbb{R}^{n_y})$. This implies (19).

(c) Introducing (13)–(14) into (11) leads to the variational inequality

$$\int_{t_0}^{t_f}\mathcal{H}'_u[t](u(t) - \hat{u}(t))dt \geq 0,$$

for all $u \in U_{\text{ad}}$. This implies the optimality condition [cf. Kirsch et al. (Ref. 23, p. 102)].

(d) According to Theorem 2.1 (8), it holds that $\eta_2^* \in K_2^+$; i.e.,

$$\eta_2^*(z) = \sum_{i=1}^{n_s}\int_{t_0}^{t_f} z_i(t)\,d\mu_i(t) \geq 0,$$

for all $z \in K_2 = \{z \in C([t_0,t_f],\mathbb{R}^{n_s})\,|\,z(t) \geq 0_{n_s} \text{ in } [t_0,t_f]\}$. This implies that μ_i is monotonically increasing. Finally, the condition $\eta_2^*(s(\cdot,\hat{x}(\cdot))) = 0$, i.e.,

$$\eta_2^*(s(\cdot,\hat{x}(\cdot))) = \sum_{i=1}^{n_s}\int_{t_0}^{t_f} s_i(t,\hat{x}(t))\,d\mu_i(t) = 0,$$

together with the monotonicity of μ_i, implies that μ_i is constant in intervals with $s_i(t, \hat{x}(t)) < 0$. □

Likewise, Theorem 2.1 and the multiplier representations (13)–(16) yield the following necessary conditions for Problem 1.1 with mixed control-state constraints. The augmented Hamiltonian function is defined by

$$\hat{\mathcal{H}}(t, x, y, u, \lambda_f, \lambda_g, \eta, l_0) \\ := \mathcal{H}(t, x, y, u, \lambda_f, \lambda_g, l_0) + \eta^\top c(t, x, y, u), \tag{23}$$

Theorem 3.2. Local Minimum Principle for Control Problems without Set Constraints. Let the assumptions of Theorem 2.1 be fulfilled. In addition, let g be twice continuously differentiable w.r.t. all the arguments. Furthermore, assume that $\mathcal{U} = \mathbb{R}^{n_u}$ and rank $(c'_u[t]) = n_c$, a.e. in $[t_0, t_f]$, and that the matrix $\hat{M}$ in (12) is nonsingular with essentially bounded inverse $\hat{M}^{-1}$ a.e. in $[t_0, t_f]$. Then, there exist multipliers $l_0 \in \mathbb{R}$, $\lambda_f \in \mathrm{BV}([t_0, t_f], \mathbb{R}^{n_x})$, $\lambda_g \in L^\infty([t_0, t_f], \mathbb{R}^{n_y})$, $\eta \in L^\infty([t_0, t_f], \mathbb{R}^{n_c})$, $\mu \in \mathrm{NBV}([t_0, t_f], \mathbb{R}^{n_s})$, $\zeta \in \mathbb{R}^{n_y}$, and $\sigma \in \mathbb{R}^{n_\psi}$ such that the following conditions are satisfied:

(i) $l_0 \geq 0$, $(l_0, \zeta, \sigma, \lambda_f, \lambda_g, \eta, \mu) \neq \Theta$;

(ii) adjoint equations,

$$\lambda_f(t) = \lambda_f(t_f) \\ + \int_t^{t_f} \hat{\mathcal{H}}'_x(\tau, \hat{x}(\tau), \hat{y}(\tau), \hat{u}(\tau), \lambda_f(\tau), \lambda_g(\tau), \eta(\tau), l_0)^\top d\tau \\ + \sum_{i=1}^{n_s} \int_t^{t_f} s'_{i,x}(\tau, \hat{x}(\tau))^\top d\mu_i(\tau), \quad \text{in } [t_0, t_f], \tag{24}$$

$$0_{n_y} = \hat{\mathcal{H}}'_y(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), \eta(t), l_0)^\top, \\ \text{a.e. in } [t_0, t_f]; \tag{25}$$

(iii) transversality conditions,

$$\lambda_f(t_0)^\top = -(l_0 \varphi'_{x_0}(\hat{x}(t_0), \hat{x}(t_f)) + \sigma^\top \psi'_{x_0}(\hat{x}(t_0), \hat{x}(t_f)) \\ + \zeta^\top g'_x(t_0, \hat{x}(t_0))), \tag{26}$$

$$\lambda_f(t_f)^\top = l_0\, \varphi'_{x_f}(\hat{x}(t_0), \hat{x}(t_f)) + \sigma^\top \psi'_{x_f}(\hat{x}(t_0), \hat{x}(t_f)); \tag{27}$$

(iv) optimality condition: a.e. in $[t_0, t_f]$, it holds that

$$\hat{\mathcal{H}}'_u(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), \eta(t), l_0) = 0_{n_u}; \tag{28}$$

(v) complementarity conditions: almost everywhere in $[t_0, t_f]$, it holds that

$$\eta(t)^\top c(t, \hat{x}(t), \hat{y}(t), \hat{u}(t)) = 0, \quad \eta(t) \geq 0_{n_c};$$

μ_i is monotonically increasing on $[t_0, t_f]$ and constant on every interval (t_1, t_2) with $t_1 < t_2$ and $s_i(t, \hat{x}(t)) < 0$ for all $t \in (t_1, t_2)$.

Proof. Theorem 2.1 and (13)–(16) yield the existence of the functions $\lambda_f, \lambda_g, \eta$ and provide representations of the functionals $\lambda_f^*, \lambda_g^*, \eta_1^*, \eta_2^*$. The assertions follow by repeating the proof of Theorem 3.1.

The following considerations apply to both Theorem 3.1 and Theorem 3.2 and differ only in the Hamilton functions $\mathcal{H}$ and $\hat{\mathcal{H}}$, respectively. Hence, we restrict the discussion to the situation of Theorem 3.1.

The multiplier μ is of bounded variation. Hence, it has at most countably many jump points and μ can be expressed as $\mu = \mu_a + \mu_d + \mu_s$, where μ_a is absolutely continuous, μ_d is a jump function, and μ_s is singular (continuous, nonconstant, $\dot{\mu}_s = 0$ a.e.). Hence, the adjoint equation (18) can be written as

$$\begin{aligned}\lambda_f(t) = \lambda_f(t_f) &\int_t^{t_f} \hat{\mathcal{H}}_x'(\tau, \hat{x}(\tau), \hat{y}(\tau), \hat{u}(\tau), \lambda_f(\tau), \lambda_g(\tau), l_0)^\top \tau \\ &+ \sum_{i=1}^{n_s} \left(\int_t^{t_f} s_{i,x}'(\tau, \hat{x}(\tau))^\top d\mu_{i,a}(\tau) + \int_t^{t_f} s_{i,x}'(\tau, \hat{x}(\tau))^\top d\mu_{i,d}(\tau) \right. \\ &\left. + \int_t^{t_f} s_{i,x}'(\tau, \hat{x}(\tau))^\top d\mu_{i,s}(\tau) \right),\end{aligned}$$

for all $t \in [t_0, t_f]$. Notice, that λ_f is continuous from the right in (t_0, t_f), since μ is normalized. Let $\{t_j\}$, $j \in \mathcal{J}$, be the jump points of μ. Then, at every jump point t_j, it holds that

$$\begin{aligned}&\lim_{\varepsilon \downarrow 0} \left(\int_{t_j}^{t_f} s_{i,x}'(\tau, \hat{x}(\tau))^\top d\mu_{i,d}(\tau) - \int_{t_{j-\varepsilon}}^{t_f} s_{i,x}'(\tau, \hat{x}(\tau))^\top d\mu_{i,d}(\tau) \right) \\ &= -s_{i,x}'(t_j, \hat{x}(t_j))^\top (\mu_{i,d}(t_j) - \mu_{i,d}(t_j-)).\end{aligned}$$

Since μ_a is absolutely continuous and μ_s is continuous, we obtain the jump condition

$$\lambda_f(t_j) - \lambda_f(t_j-) = -\sum_{i=1}^{n_s} s_{i,x}'(t_j, \hat{x}(t_j))^\top (\mu_i(t_j) - \mu_i(t_j-)), \quad j \in \mathcal{J}.$$

In order to derive a differential equation for λ_f, we need the subsequent auxiliary result, which can be proved using the mean-value theorem for Riemann-Stieltjes integrals.

Corollary 3.1. Let $\omega : [t_0, t_f] \to \mathbb{R}$ be continuous and let μ be montonically increasing on $[t_0, t_f]$. Let μ be differentiable at $t \in [t_0, t_f]$. Then, it holds that

$$(d/dt) \int_{t_0}^{t} w(s)\, d\mu(s) = w(t)\dot{\mu}(t).$$

Furthermore, since every function of bounded variation is differentiable almost everywhere, μ and λ_f are differentiable almost everywhere. Thus, we proved the following corollary.

Corollary 3.2. Let the assumptions of Theorem 3.1 be fulfilled. Then, λ_f is differentiable almost everywhere in $[t_0, t_f]$ with

$$\begin{aligned}\dot{\lambda}_f(t) = &-\mathcal{H}'_x(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), l_0)^\top \\ &- \sum_{i=0}^{n_s} s'_{i,x}(t, \hat{x}(t))^\top \dot{\mu}_i(t).\end{aligned} \tag{29}$$

Furthermore, the jump conditions

$$\lambda_f(t_j) - \lambda_f(t_j-) \;=\; -\sum_{i=1}^{n_s} s'_{i,x}(t_j, \hat{x}(t_j))^\top (\mu_i(t_j) - \mu_i(t_j-)). \tag{30}$$

hold at every point $t_j \in (t_0, t_f)$ of discontinuity of the multiplier μ.

A special case arises if no state constraints are present. Then, the adjoint variable λ_f is even absolutely continuous, i.e. $\lambda_f \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x})$ and the adjoint equations (18)–(19) become

$$\dot{\lambda}_f(t) = -\mathcal{H}'_x(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), l_0)^\top, \quad \text{a.e. in } [t_0, t_f], \tag{31}$$

$$0_{n_y} = \mathcal{H}'_y(t, \hat{x}(t), \hat{y}(t), \hat{u}(t), \lambda_f(t), \lambda_g(t), l_0)^\top, \quad \text{a.e. in } [t_0, t_f]. \tag{32}$$

The adjoint equations (31) and (32) form a DAE system of index one for λ_f and λ_g where λ_f is the differential variable and λ_g denotes the algebraic variable. This follows from (32), which is given by

$$0_{n_y} \;=\; l_0(f'_{0,y}[t])^\top + (f'_y[t])^\top \lambda_f(t) + (g'_x[t] \cdot f'_y[t])^\top \lambda_g(t).$$

Since $g'_x[t] \cdot f'_y[t]$ is nonsingular, we obtain

$$\lambda_g(t) \;=\; -(g'_x[t] \cdot f'_y[t])^{-\top} (l_0(f'_{0,y}[t])^\top + (f'_y[t])^\top \lambda_f(t)).$$

Remark 3.1. Notice that the adjoint system (31)–(32) is an index-1 DAE, while the original DAE is of index 2 according to Assumption 1.1.

In this section, we concentrate only on local minimum principles. The term "local" is due to the fact, that the optimality conditions (22) and (28), respectively, can be interpreted as necessary conditions for a local minimum of the Hamilton function and the augmented Hamilton function, respectively. However, there are also global minimum principles. The main important difference between a local and a global minimum principle is that $\mathcal{U}$ can be an arbitrary subset of $\mathbb{R}^{n_u}$ in the global case e.g., a discrete set. In our approach, we had to assume that $\mathcal{U}$ is a convex set with nonempty interior. Proofs for a global minimum [resp. maximum principles] in the context of ordinary differential equations subject to pure state constraints can be found e.g., in Girsanov (Ref. 4) and in Ioffe and Tihomirov (Ref. 8, pp. 147–159 and 241–253).

4. Regularity

We state conditions which ensure that the multiplier l_0 is nonzero and without loss of generality can be normalized to one. Again, we consider Problem 1.1 and the equivalent infinite optimization Problem 2.1 respectively. It is well-known that the following Mangasarian-Fromowitz condition implies $l_0 = 1$ [cf. Corollary 2.1 in Gerdts (Ref. 24)]:

(i) $\mathcal{H}'(\hat{x}, \hat{y}, \hat{u})$ is surjective;
(ii) there exists some $(\delta x, \delta y, \delta u) \in \operatorname{int}(S - \{(\hat{x}, \hat{y}, \hat{u}\})$ with

$$\mathcal{H}'(\hat{x}, \hat{y}, \hat{u})(\delta x, \delta y, \delta u) = \Theta z,$$
$$G(\hat{x}, \hat{y}, \hat{u}) + G'(\hat{x}, \hat{y}, \hat{u})(\delta x, \delta y, \delta u) \in \operatorname{int}(K).$$

Notice, that the assumption int$(K) \neq \emptyset$ in Corollary 2.1 in Gerdts (Ref. 24) is satisfied for Problem 2.1. A sufficient condition for (i) to hold is given by the following lemma, which is an immediate consequence of part (c) of Lemma 3.1 in Gerdts (Ref. 24). It is an extension to DAEs of a lemma that has been used by several authors (e.g. Malanowski) for control problems with mixed control-state constraints or pure state constraints.

Lemma 4.1. Let Assumption 1.1 be valid and let

$$\operatorname{rank}((\psi'_{x_0}\Phi(t_0) + \psi'_{x_f}\Phi(t_{t_f}))\Gamma) = n_\psi, \tag{33}$$

where Φ is the fundamental solution of the homogeneous linear differential equation

$$\dot{\Phi}(t) = A(t)\Phi(t), \quad \Phi(t_0) = I_{n_x}, \quad t \in [t_0, t_f],$$

and the columns of Γ constitute an orthonormal basis of ker $(g'_x[t_0])$ and

$$\mathrm{M(t)} := g'_x[t] \cdot f'_y[t],$$
$$A(t) := f'_{\mathrm{x}}[t] - f'_y[t]M(t)^{-1}Q(t),$$
$$h(t) := h_1(t) - f'_y[t]M(t)^{-1}d(t),$$

$$Q(t) := (d/dt)\, g'_x[t] + g'_x[t] \cdot f'_x[t],$$
$$d(t) := \dot{h}_2(t) + g'_x[t] h_1(t).$$

Then, $H'(\hat{x}, \hat{y}, \hat{u})$ in Problem 2.1 is surjective.

Condition (ii) is satisfied if there exist $\delta x \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{n_x})$, $\delta y \in L^\infty([t_0, t_f], \mathbb{R}^{n_y})$, $\delta u \in \text{int}\,(U_{ad} - \{\hat{u}\})$, and $\varepsilon > 0$ satisfying

$$c[t] + c'_x[t]\delta x(t) + c'_y[t]\delta y(t) + c'_u[t]\delta u(t)$$
$$\leq -\varepsilon \cdot e, \quad \text{a.e. in} \quad [t_0, t_f], \tag{34}$$
$$s[t] + s'_x[t]\delta x(t) < 0_{n_s}, \quad \text{in } [t_0, t_f], \tag{35}$$
$$f'_x[t]\delta x(t) + f'_y[t]\delta y(t) + f'_u[t]\delta u(t) - \delta\dot{x}(t) = 0_{n_x},$$
$$\text{a.e. in } [t_0, t_f], \tag{36}$$
$$g'_x[t]\delta x(t) = 0_{n_y}, \quad \text{in } [t_0, t_f], \tag{37}$$
$$\psi'_{x_0}\delta x(t_0) + \psi'_{x_f}\delta x(t_f) = 0_{n_\psi}, \tag{38}$$

where $e = (1, \ldots, 1)^\top \in \mathbb{R}^{n_c}$. Hence, we conclude that the following theorem holds.

Theorem 4.1. Let the assumptions of Theorems 3.1 or 3.2 and of Lemma 4.1 be fulfilled. Furthermore, let there exist $\delta x \in W^{1,\infty}([t_0, t_f]\mathbb{R}^{n_x})$, $\delta y \in L^\infty([t_0, t_f], \mathbb{R}^{n_y})$, and $\delta u \in \text{int}(U_{ad} - \{\hat{u}\})$ satisfying (34)–(38). Then, it holds $l_0 = 1$ in Theorems 3.1 or 3.2, respectively.

5. Example

We apply the local minimum principle in Theorem 3.1 to the subsequent index-2 DAE optimal control problem without state constraints. The task is to minimize the functional

$$F(x, y, u) = \int_0^3 u(t)^2 dt,$$

subject to the equations of motion of the mathematical pendulum in the Gear-Gupta-Leimkuhler (GGL) formulation [cf. Gear et al. (Ref. 20)], given by

$$\dot{x}_1(t) = x_3(t) - 2x_1(t)y_2(t), \tag{39}$$
$$\dot{x}_2(t) = x_4(t) - 2x_2(t)y_2(t), \tag{40}$$
$$\dot{x}_3(t) = -2x_1(t)y_1(t) + u(t)x_2(t), \tag{41}$$
$$\dot{x}_4(t) = -g - 2x_2(t)y_1(t) - u(t)x_1(t), \tag{42}$$

$$0 = x_1(t)x_3(t) + x_2(t)x_4(t), \tag{43}$$
$$0 = x_1(t)^2 + x_2(t)^2 - 1, \tag{44}$$

and the boundary conditions

$$\psi(x(0), x(3)) := (x_1(0) - 1, x_2(0), x_3(0), x_4(0), x_1(3), x_3(3))^\top = 0_6. \tag{45}$$

Herein, $g = 9.81$ denotes the acceleration due to gravity. The control u is not restricted, i.e., $\mathcal{U} = \mathbb{R}$. With

$$x = (x_1, x_2, x_3, x_4,)^\top, \quad y = (y_1, y_2)^\top, \quad f_0(u) = u^2,$$
$$f(x, y, u) = (x_3 - 2x_1y_2,\ x_4 - 2x_2y_2,\ -2x_1y_1 + ux_2,\ -g - 2x_2y_1 - ux_1)^\top,$$
$$g(x) = (x_1x_3 + x_2x_4,\ x_1^2 + x_2^2 - 1)^\top,$$

the problem has the structure of Problem 1.1. The matrix

$$g'_x(x) \cdot f'_y(x, y, u) = \begin{pmatrix} -2 & 0 \\ 0 & -4 \end{pmatrix}$$

is nonsingular in a local minimum; hence, the DAE has index two and Assumption 1.1 is satisfied. The remaining assumptions of Theorem 3.1 are satisfied as well and necessarily there exist functions $\lambda_f = (\lambda_{f,1}, \lambda_{f,2}, \lambda_{f,3}, \lambda_{f,4})^\top \in W^{1,\infty}([0, 3], \mathbb{R}^4)$, $\lambda_g = (\lambda_{g,1}, \lambda_{g,2})^\top \in L^\infty([0, 3], \mathbb{R}^2)$, and vectors $\zeta = (\zeta_1, \zeta_2)^\top$, $\sigma = (\sigma_1 \ldots, \sigma_6)^\top$ such that the adjoint equations (31)–(32), the transversality conditions (20)–(21), and the optimality condition (22) are satisfied. The Hamilton function (17) is given by

$$\begin{aligned} H(x, y, u, \lambda_f, \lambda_g, l_0) = {} & l_0u^2 + \lambda_{f,1}(x_3 - 2x_1y_2) + \lambda_{f,2}(x_4 - 2x_2y_2) \\ & + \lambda_{f,3}(-2x_1y_1 + ux_2) + \lambda_{f,4}(-g - 2x_2y_1 - ux_1) \end{aligned}$$

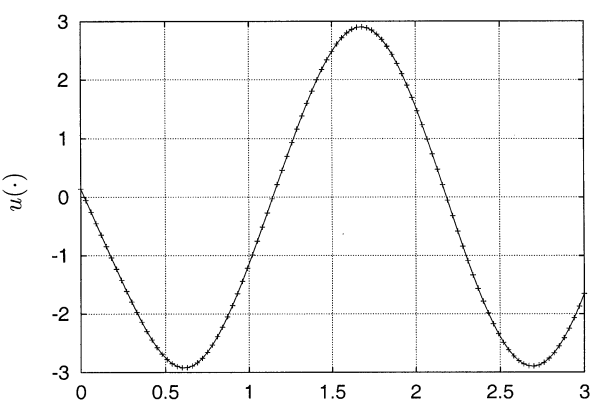

Fig. 1. Numerical solution of BVP resulting from the minimum principle: Control $u(t)$ for $t \in [0, 3]$.

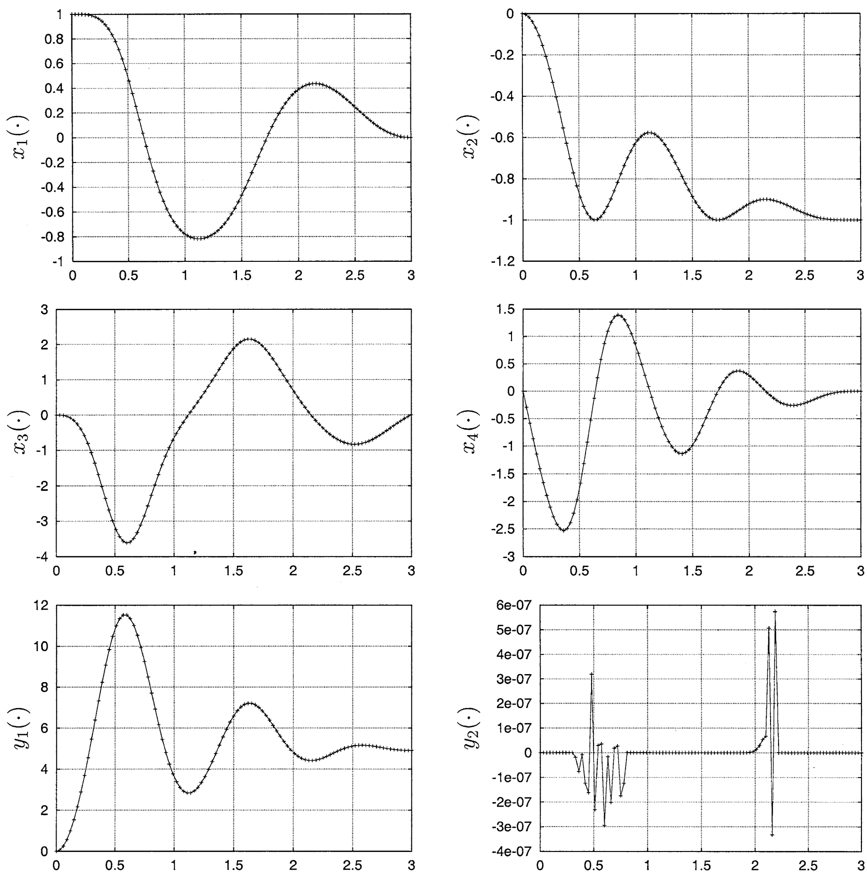

Fig. 2. Numerical solution of BVP resulting from the minimum principle: Differential variable $x(t)$ and algebraic variable $y(t)$ for $t \in [0, 3]$.

$$\begin{aligned}
&+\lambda_{g,1}(-gx_2 - 2y_1(x_1^2 + x_2^2) + x_3^2 + x_4^2 \\
&-2y_2(x_1x_3 + x_2x_4)) + \lambda_{g,2}(2(x_1x_3 + x_2x_4) \\
&-4y_2(x_1^2 + x_2^2)).
\end{aligned}$$

In the sequel, we assume $l_0 = 1$ (actually, the Mangasarian-Fromowitz condition is satisfied). Then, the optimality condition (22) yields

$$0 = 2u + \lambda_{f,3}x_2 - \lambda_{f,4}x_1 \Rightarrow u = (\lambda_{f,4}x_1 - \lambda_{f,3}x_2)/2. \tag{46}$$

The transversality conditions (20)–(21) are given by

$$\lambda_f(0) = (-\sigma_1 - 2\zeta_2, -\sigma_{2,} - \sigma_3 - \zeta_1, -\sigma_4)^\top, \quad \lambda_f(3) = (\sigma_5, 0, \sigma_6, 0)^\top.$$

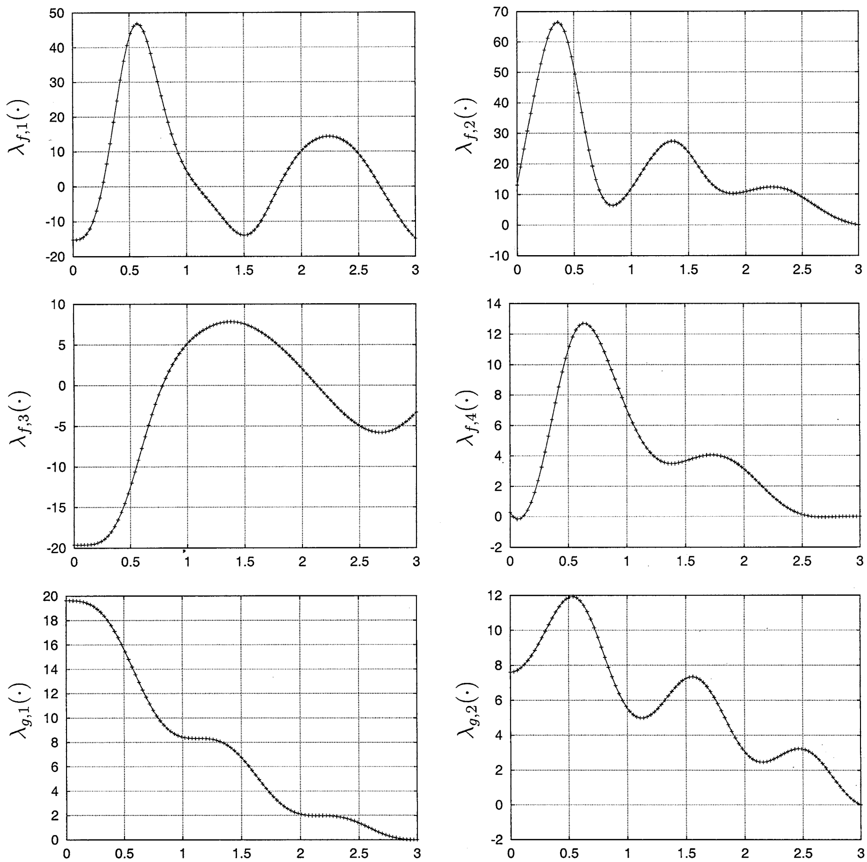

Fig. 3. Numerical solution of BVP resulting from the minimum principle: Adjoint variables $\lambda_f(t)$ and $\lambda_g(t)$ for $t \in [0, 3]$.

The adjoint equations (31)–(32) yield

$$\dot{\lambda}_{f,1} = 2(\lambda_{f,1}y_2 + \lambda_{f,3}y_1) + \lambda_{f,4}u - \lambda_{g,1}(-4y_1x_1 - 2y_2x_3) - \lambda_{g,2}(2x_3 - 8y_2x_1), \tag{47}$$

$$\dot{\lambda}_{f,2} = 2(\lambda_{f,2}y_2 + \lambda_{f,4}y_1) - \lambda_{f,3}u - \lambda_{g,1}(-g - 4y_1x_2 - 2y_2x_4) - \lambda_{g,2}(2y_2x_4 - 8y_2x_2), \tag{48}$$

$$\dot{\lambda}_{f,3} = -\lambda_{f,1} - \lambda_{g,1}(2x_3 - 2x_1y_2) - 2\lambda_{g,2}x_1, \tag{49}$$

$$\dot{\lambda}_{f,4} = -\lambda_{f,2} - \lambda_{g,1}(2x_4 - 2x_2y_2) - 2\lambda_{g,2}x_2, \tag{50}$$

$$0 = -2\big(\lambda_{f,3}x_1 + \lambda_{f,4}x_2 + \lambda_{g,1}\big(x_1^2 + x_2^2\big)\big), \tag{51}$$

$$0 = -2\big(\lambda_{f,1}x_1 + \lambda_{f,2}x_2 + \lambda_{g,1}(x_1x_3 + x_2x_4) + 2\lambda_{g,2}\big(x_1^2 + x_2^2\big)\big). \tag{52}$$

Notice that consistent initial values for $\lambda_{g_1}(0)$ and $\lambda_{g,2}(0)$ can be calculated from (51)–(52) and (43)–(44) by

$$\lambda_{g,1} = -\lambda_{f,3}x_1 - \lambda_{f,4}x_2, \quad \lambda_{g,2}(-\lambda_{f,1}x_1 - \lambda_{f,2}x_2)/2.$$

The differential equations (39)–(44) and (47)–(52), with u replaced by (46) together with the boundary conditions (45) and $\lambda_{f,2}(3) = 0, \quad \lambda_{f,4}(3) = 0$, form a two-point boundary-value problem (BVP). Notice that the DAE system has index-1 constraints (51)–(52) for λ_g as well as index-2 constraints (43)–(44) for y.

Numerically, the BVP is solved by a single shooting method as follows. Let

$$z = (\sigma_1 + 2\zeta_2, \sigma_{2,} \sigma_3 + \zeta_1, \sigma_4)^\top$$

denote the unknown initial values of $-\lambda_f$ and let $x(t;z)$, $y(t;z)$, $\lambda_f(t;z)$, $\lambda_g(t;z)$ denote the solution of the initial value problem given by (39)–(44), (47)–(52), and the initial conditions $x(0) = (1, 0, 0, 0)^\top$ and $\lambda_f(0) = -z$. Then, the BVP is solvable, if the nonlinear equation

$$G(z) := (x_1(3;z),\ x_3(3;z),\ \lambda_{f,2}(3;z), \lambda_{f,4}(3;z)) = (0, 0, 0, 0)$$

is solvable. Numerically, the nonlinear equation is solved by the Newton method. The required Jacobian $G'_z(z)$ is obtained by a sensitivity analysis of the initial-value problem w.r.t. z. Herein, the sensitivity DAE associated with the differential equations is employed. Figs. 1–3 show the numerical solution obtained from the Newton method. Notice, that the initial conditions in (45) and the algebraic equations (43) and (44) contain redundant information. Hence, the multipliers σ and ζ are not unique; e.g., one may set $\zeta = 0_2$. In order to obtain unique multipliers, one can dispense with the first and third initial condition in (45), since these are determined by (43) and (44).

6. Conclusions

The presented local minimum principles for optimal control problems subject to index-2 differential-algebraic equations are not only of theoretical interest but give rise to numerical solution methods for such problems. The so-called indirect approach intends to fulfill the necessary conditions for the optimal control problem numerically and thus produces candidates for an optimal solution. Notice that the evaluation of the local minimum principle will lead to a multipoint boundary-value problem, at least under suitable simplifying assumptions such as e.g. that the structure of active and inactive state constraints is known and that the singular part of the multiplier μ vanishes (cf. Corollary 3.2). Even for so-called direct methods, which are based on a discretization of the optimal control problem [cf. Gerdts (Ref. 21)], the local minimum principle is of great importance in view of the postoptimal approximation of the adjoints. Interestingly, the necessary conditions yield an index-1 DAE for the adjoints, whereas the original DAE has index 2. The example shows that the coupled system of state and adjoint equations has mixed index.

References

1. PONTRYAGIN, L. S., BOLTYANSKIJ, V. G., GAMKRELIDZE, R. V., and MISHCHENKO, E. F., *Mathematische Theorie Optimaler Prozesse*, Oldenbourg, München, Germany, 1964 (in German).
2. HESTENES, M. R., *Calculus of Variations and Optimal Control Theory*, John Wiley and Sons, New York, NY, 1966.
3. JACOBSON, D. H., LELE, M. M., and SPEYER, J. L., *New Necessary Conditions of Optimality for Constrained Problems with State-Variable Inequality Constraints*, Journal of Mathematical Analysis and Applications, Vol. 35, pp. 255–284, 1971.
4. GIRSANOV, I. V., *Lectures on Mathematical Theory of Extremum Problems*, Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, Germany, Vol. 67, 1972.
5. KNOBLOCH, H. W., *Das Pontryaginsche Maximumprinzip für Probleme mit Zustandsbeschränkungen i und ii*, Zeitschrift für Angewandte Mathematik und Mechanik, Vol. 55, pp. 545–556 and 621–634, 1975.
6. MAURER, H., *On Optimal Control Problems with Boundary State Variables and Control Appearing Linearly*, SIAM Journal on Control and Optimization, Vol. 15, pp. 345–362, 1977.
7. MAURER, H., *On the Minimum Principle for Optimal Control Problems with State Constraints*, Schriftenreihe des Rechenzentrums der Universität Münster, Münster, Germany, Vol. 41, 1979.
8. IOFFE, A. D., and TIHOMIROV, V. M., *Theory of Extremal Problems,* Studies in Mathematics and Its Applications, North-Holland Publishing Company, Amsterdam, Holland, Vol. 6, 1979.
9. KREINDLER, E., *Additional Necessary Conditions for Optimal Control with State-Variable Inequality Constraints*, Journal of Optimization Theory and Applications, Vol. 38, pp. 241–250, 1982.
10. NEUSTADT, L. W., *Optimization: A Theory of Necessary Conditions*, Princeton University Press, Princeton, New Jersey, 1976.
11. ZEIDAN, V., *The Riccati Equation for Optimal Control Problems with Mixed State-Control Constraints: Necessity and Sufficiency*, SIAM Journal on Control and Optimization, Vol. 32, pp. 1297–1321, 1994.
12. HARTL, R. F., SETHI, S. P., and VICKSON, G., *A Survey of the Maximum Principles for Optimal Control Problems with State Constraints*, SIAM Review, Vol. 37, pp. 181–218, 1995.
13. BRYSON, A. E., and HO, Y. C., *Applied Optimal Control*, Hemisphere Publishing Corporation, Washington, DC, 1975.
14. MAURER, H., *First and Second-Order Sufficient Optimality Conditions in Mathematical Programming and Optimal Control*, Mathematical Programming Study, Vol. 14, pp. 163–177, 1981.

15. MALANOWSKI, K., *Sufficient Optimality Conditions for Optimal Control Subject to State Constraints*, SIAM Journal on Control and Optimization, Vol. 35, pp. 205–227, 1997.
16. MAURER, H., and PICKENHAIN, S., *Second-Order Sufficient Conditions for Control Problems with Mixed Control-State Constraints*, Journal of Optimization Theory and Applications, Vol. 86, pp. 649–667, 1995.
17. MALANOWSKI, K., MAURER, H., and PICKENHAIN, S., *Second-Order Sufficient Conditions for State-Constrained Optimal Control Problems*, Journal of Optimization Theory and Applications, Vol. 123, pp. 595–617, 2004.
18. DE PINHO, M., and VINTER, R. B., *Necessary Conditions for Optimal Control Problems Involving Nonlinear Differential Algebraic Equations*, Journal of Mathematical Analysis and Applications, Vol. 212, pp. 493–516, 1997.
19. DEVDARIANI, E. N., and LEDYAEV, Y. S., *Maximum Principle for Implicit Control Systems*, Applied Mathematics and Optimization, Vol. 40, pp. 79–103, 1999.
20. GEAR, C. W., LEIMKUHLER, B., and GUPTA, G. K., *Automatic Integration of Euler-Lagrange Equations with Constraints*, Journal of Computational and Applied Mathematics, Vols. 12/13, pp. 77–90, 1985.
21. GERDTS, M., *Direct Shooting Method for the Numerical Solution of Higher Index DAE Optimal Control Problems*, Journal of Optimization Theory and Applications, Vol. 117, pp. 267–294, 2003.
22. GERDTS, M., *Optimal Control and Real-Time Optimization of Mechanical Multibody Systems*, Zeitschrift für Angewandte Mathematik und Mechanik, Vol. 83, pp. 705–719, 2003.
23. KIRSCH, A., WARTH, W., and WERNER, J., *Notwendige Optimalitätsbedingungen und ihre Anwendung*, Lecture Notes in Economics and Mathematical Systems, Springer, Berlin, Germany, Vol. 152, 1978.
24. GERDTS, M., *Representation of Lagrange Multipliers for Optimal Control Problems Subject to Differential-Algebraic Equations of Index Two*, Journal of Optimization Theory and Applications, Vol. 130, 2005.