

# Second-Order Optimality Conditions for Broken Extremals and Bang-Bang Controls: Theory and Applications

Nikolai P. Osmolovskii and Helmut Maurer

**Abstract** We survey the results on no-gap second-order optimality conditions (both necessary and sufficient) in the Calculus of Variations and Optimal Control, that were obtained in the monographs Milyutin and Osmolovskii (Calculus of Variations and Optimal Control. Translations of Mathematical Monographs. American Mathematical Society, Providence, 1998) and Osmolovskii and Maurer (Applications to Regular and Bang-Bang Control: Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control. SIAM Series Design and Control, vol. DC 24. SIAM Publications, Philadelphia, 2012), and discuss their further development. First, we formulate such conditions for broken extremals in the simplest problem of the Calculus of Variations and then, we consider them for discontinuous controls in optimal control problems with endpoint and mixed state-control constraints, considered on a variable time interval. Further, we discuss such conditions for bang-bang controls in optimal control problems, where the control appears linearly in the Pontryagin-Hamilton function with control constraints given in the form of a convex polyhedron. Bang-bang controls induce an optimization problem with respect to the switching times of the control, the so-called Induced Optimization Problem. We show that second-order sufficient condition for the Induced Optimization Problem together with the so-called strict bang-bang property ensures second-order sufficient conditions for the bang-bang control problem. Finally, we discuss optimal control problems with mixed control-state constraints and control appearing linearly. Taking the mixed

---

N.P. Osmolovskii (✉)

University of Technology and Humanities, ul. Malczewskiego 20a, 26-600 Radom, Poland  
Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6,  
01-447 Warszawa, Poland

Moscow State University of Civil Engineering, Jaroslavscoe shosse 26, 129337 Moscow, Russia  
e-mail: [osmolovski@uph.edu.pl](mailto:osmolovski@uph.edu.pl)

H. Maurer

Institut für Numerische und Angewandte Mathematik, Westfälische Wilhelms-Universität  
Münster, Einsteinstr. 62, 48149 Münster, Germany  
e-mail: [maurer@math.uni-muenster.de](mailto:maurer@math.uni-muenster.de)

© Springer International Publishing Switzerland 2016

J.-B. Hiriart-Urruty et al. (eds.), *Advances in Mathematical Modeling, Optimization and Optimal Control*, Springer Optimization and Its Applications 109,  
DOI 10.1007/978-3-319-30785-5\_6

147

constraint as a new control variable we convert such problems to bang-bang control problems. The numerical verification of second-order conditions is illustrated on three examples.

## 1 Introduction

We survey some main results presented in the recent monograph of the authors [36] (SIAM, 2012) and also some results obtained in the earlier monograph of Milyutin and Osmolovskii [28] (AMS, 1998). We discuss further developments of these results and give various applications.

Our main goal is to present and discuss the no-gap second-order necessary and sufficient conditions in control problems with bang-bang controls. In [28], it was shown how, by using quadratic conditions for the general problem of the Calculus of Variations with regular mixed equality constraint  $g(t, x, u) = 0$ , one can obtain quadratic (necessary and sufficient) conditions in optimal control problems in which the control variable enters linearly and the control constraint is given in the form of a convex polyhedron. These features were proved in Milyutin and Osmolovskii [28], who first used the property that the set  $\text{ex } U$  of vertices of a polyhedron  $U$  can be described by a nondegenerate relation  $g(u) = 0$  on an open set  $\mathcal{Q}$  consisting of disjoint open neighborhoods of vertices. This allowed us to develop quadratic necessary conditions for bang-bang controls. Further, in [28] it was shown that a sufficient condition for a minimum on  $\text{ex } U$  guarantees (in the problem in which the control enters linearly) the minimum on its convexification  $U$ . In this way, quadratic sufficient conditions for bang-bang controls were obtained in Osmolovskii and Maurer [36]. This property, which is not discussed in the present paper, constitutes the main link between the second-order optimality conditions for broken extremals in the Calculus of Variations and the second-order optimality conditions for bang-bang controls in optimal control.

The paper is organized as follows. In Sect. 2, we formulate no-gap second-order conditions for broken extremals in the simplest problem of the Calculus of Variations. In Sect. 3, we consider such conditions for discontinuous controls in optimal control problems on a fixed time interval with endpoint constraints of equality and inequality type and mixed state-control constraints of equality type. In Sect. 4, we present an extension of the results of Sect. 3 to problems on a variable time interval. In Sect. 5, we discuss no-gap conditions for bang-bang controls. Bang-bang controls induce an optimization problem with respect to the switching times of the control that we call the *Induced Optimization Problem (IOP)*. We have shown in our monograph [36] that the classical second-order sufficient condition for the IOP, together with the so-called strict bang-bang property, ensures second-order sufficient conditions for the bang-bang control problem. We discuss such conditions in Sect. 6.

In the next two sections, the theoretical results are illustrated by numerical examples. Namely, in Sect. 7, we study the optimal control of the chemotherapy of

HIV, when the control-quadratic objective in [18] of  $L^2$ -type is replaced by a more realistic  $L^1$ -objective. In Sect. 8, we consider time-optimal controls in two models of two-link robots; cf. [36]. Finally, in Sect. 9, we discuss optimal control problems with running mixed control-state constraints and control appearing linearly. Taking the mixed constraint as a new control variable we convert such problem to a bang-bang control problem. We use this transformation to study extremals in the optimal control problem for the Rayleigh equation.

## 2 Second-Order Optimality Conditions for Broken Extremals in the Simplest Problem in the Calculus of Variations

### 2.1 The Simplest Problem in the Calculus of Variations

Let a closed interval  $[t_0, t_f]$ , two points  $a, b \in \mathbb{R}^n$ , an open set  $\mathcal{Q} \subset \mathbb{R}^{2n+1}$ , and a function  $L : \mathcal{Q} \mapsto \mathbb{R}$  of class  $C^2$  be given. The simplest problem of the Calculus of Variations has the form

$$(SP) \quad \text{Minimize } \mathcal{J}(x(\cdot)) := \int_{t_0}^{t_f} L(t, x(t), \dot{x}(t)) dt, \tag{1}$$

$$x(t_0) = a, \quad x(t_f) = b, \quad (t, x(t), \dot{x}(t)) \in \mathcal{Q}. \tag{2}$$

We consider this problem in the space  $W^{1,\infty}$  of Lipschitz continuous functions. The last condition in (2) is assumed to hold almost everywhere. A *weak minimum* is defined as a local minimum in the space  $W^{1,\infty}$ . We say that a function  $x \in W^{1,\infty}([t_0, t_f], \mathbb{R}^{d(x)})$  is *admissible* if  $x$  satisfies (2) and, moreover, there exists a compact set  $\mathcal{C} \subset \mathcal{Q}$  such that  $(t, x(t), \dot{x}(t)) \in \mathcal{C}$  a.e. in  $[t_0, t_f]$ . Set  $u := \dot{x}$  and  $w = (x, u)$ . We call  $u$  the *control*.

Let an admissible function  $x^0(t)$  be an *extremal* in the sense that it satisfies the *Euler equation*

$$\frac{d}{dt} L_{\dot{x}} = L_x. \tag{3}$$

Here and in the sequel, partial derivatives are denoted by subscripts. Set

$$u^0(t) := \dot{x}^0(t), \quad w^0(t) = (x^0(t), u^0(t)).$$

Let

$$\bar{w}(\cdot) = (\bar{x}(\cdot), \bar{u}(\cdot)) \in \mathcal{W}_2 := W^{1,2} \times L^2,$$

where  $W^{1,2}$  is the space of absolutely continuous functions with square integrable derivative and  $L^2$  is the space of square integrable functions. In the space  $\mathscr{W}_2$ , let us define the subspace

$$\mathscr{K} := \{\bar{w} \in \mathscr{W}_2 \mid \frac{d}{dt} \bar{x}(t) = \bar{u}(t) \text{ a.e.}, \bar{x}(t_0) = \bar{x}(t_f) = 0\}$$

and the quadratic form

$$\begin{aligned} \Omega(\bar{w}) &= \int_{t_0}^{t_f} \langle L_{ww}(t, w^0(t)) \bar{w}(t), \bar{w}(t) \rangle dt \\ &= \int_{t_0}^{t_f} (\langle L_{xx} \bar{x}(t), \bar{x}(t) \rangle + 2 \langle L_{xu} \bar{u}(t), \bar{x}(t) \rangle + \langle L_{uu} \bar{u}(t), \bar{u}(t) \rangle) dt. \end{aligned}$$

The following theorem is well known.

- Theorem 1.** (a) *If the extremal  $x^0$  is a weak minimum, then  $\Omega(\bar{w}) \geq 0$  on  $\mathscr{K}$ .*  
 (b) *If  $\Omega(\bar{w})$  is positive definite on  $\mathscr{K}$ , then the extremal  $x^0$  is a (strict) weak minimum.*

As is known, the quadratic conditions in Theorem 1 can be tested via the Jacobi conditions or via bounded solutions to an associated Riccati equation.

For a broken extremal, the quadratic form has to be stated in a different way that allows for the formulation of no-gap necessary and sufficient second-order conditions. We will formulate these conditions and discuss their extensions to different classes of optimal control problems, including bang-bang control problems and problems with mixed constraints and control appearing linearly.

## 2.2 Second-Order Optimality Conditions for Broken Extremals

Let again  $x^0(t)$  be an extremal in the simplest problem (1), (2), and let  $u^0(t) = \dot{x}^0(t)$  be the corresponding control. Assume now that the control  $u^0(t)$  is *piecewise continuous* with one discontinuity point  $t_* \in (t_0, t_f)$ . Hence,  $x^0(t)$  is a *broken extremal* with a corner at  $t_*$ . We say that  $t_*$  is an *L-point* of the function  $u^0(t)$  if there exist  $\varepsilon > 0$  and  $C > 0$  such that  $|u^0(t) - u^0(t_*-)| \leq C|t - t_*|$  for all  $t \in (t_* - \varepsilon, t_*)$  and  $|u^0(t) - u^0(t_*+)| \leq C|t - t_*|$  for all  $t \in (t_*, t_* + \varepsilon)$ . Henceforth, we assume that  $t_*$  is an *L-point* of the function  $u^0(t)$ . The following question naturally arises: which quadratic form corresponds to a broken extremal?

Let us change the definition of a weak local minimum as follows. Set  $\Theta := \{t_*\}$  and define a notion of a  $\Theta$ -weak minimum. Assuming additionally that the control  $u^0(t)$  is left-continuous at  $t_*$ , denote by  $\text{cl } u^0(\cdot)$  the closure of the graph of  $u^0(t)$ . Denote by  $V$  a neighborhood of the compact set  $\text{cl } u^0(\cdot)$ .

**Definition 1.** We say that  $x^0$  is a point of a  $\Theta$ -weak minimum (or an *extended weak minimum*) if there exists a neighborhood  $V$  of the compact set  $\text{cl } u^0(\cdot)$  such that  $\mathscr{J}(x) \geq \mathscr{J}(x^0)$  for all admissible  $x(t)$  such that  $u(t) \in V$  a.e., where  $u(t) = \dot{x}(t)$ .

Clearly, we have the following chain of implications among minima:

$$\text{strong minimum} \implies \Theta\text{-weak minimum} \implies \text{weak minimum.}$$

Let us formulate optimality conditions for a  $\Theta$ -weak minimum. To this end, we introduce the Pontryagin function (Hamiltonian)

$$H(t, x, u, \lambda) = \lambda u + L(t, x, u),$$

where  $\lambda$  is a row vector of the dimension  $n$ . Defining  $\lambda(t) := -L_u(t, x^0(t), u^0(t))$ , we have in view of the Euler equation (3):

$$H_u(t, x^0(t), u^0(t), \lambda(t)) = 0, \quad -\dot{\lambda}(t) = H_x(t, x^0(t), u^0(t), \lambda(t)).$$

Denote by  $[\lambda]$  the jump of the function  $\lambda(t)$  at the point  $t_*$ , i.e.,  $[\lambda] = \lambda^+ - \lambda^-$ , where  $\lambda^- = \lambda(t_*-)$  and  $\lambda^+ = \lambda(t_*+)$ . Let  $[H]$  stand for the jump of the function  $H(t) := H(t, x^0(t), u^0(t), \lambda(t))$  at the same point. The equalities

$$[\lambda] = 0, \quad [H] = 0$$

constitute the *Weierstrass–Erdmann* conditions. They are known as necessary conditions for a strong minimum. However, they are also necessary for the  $\Theta$ -weak minimum. We add one more necessary condition for the  $\Theta$ -weak minimum:

$$D(H) := -L_x^+ \dot{x}^{0-} + L_x^- \dot{x}^{0+} - [L_t] \geq 0,$$

where  $\dot{x}^{0-} = \dot{x}^0(t_*-)$ ,  $L_x^- = L_x(t_*, x^0(t_*-), u^0(t_*-))$ ,  $[L_t] = L_t^+ - L_t^-$ , etc. Clearly,

$$D(H) := \dot{\lambda}^+ \dot{x}^{0-} - \dot{\lambda}^- \dot{x}^{0+} + [\dot{\lambda}_0],$$

where  $\dot{\lambda}_0(t) = -\dot{H}(t)$  (recall that  $\frac{d}{dt}H(t) = H_t(t)$  a.e.). Moreover, it can be shown that  $D(H)$  is equal to the negative derivative of the function

$$\Delta H(t) := \lambda(t)[u^0] + L(t, x^0(t), u^0(t_*+)) - L(t, x^0(t), u^0(t_*-))$$

at  $t_*$ . The existence of the derivative has been proved and, hence, this derivative can also be calculated as  $\frac{d}{dt}(\Delta H)(t_*-)$  or as  $\frac{d}{dt}(\Delta H)(t_*+)$ .

Now, let us formulate second-order optimality conditions for a  $\Theta$ -weak minimum. Denote by  $P_\Theta W^{1,2}$  the Hilbert space of piecewise continuous functions  $\bar{x}(t)$ , absolutely continuous on each of the two intervals  $[t_0, t_*]$  and  $(t_*, t_f]$ , and such that their first derivative is square integrable. Any  $\bar{x} \in P_\Theta W^{1,2}$  can have a nonzero jump  $[\bar{x}] = \bar{x}(t_*+0) - \bar{x}(t_*-0)$  at the point  $t_*$ . Let  $\bar{t}$  be a numerical parameter. Denote by  $Z_2(\Theta)$  the space of triples  $\bar{z} = (\bar{t}, \bar{x}, \bar{u})$  such that  $\bar{t} \in \mathbb{R}$ ,  $\bar{x}(\cdot) \in P_\Theta W^{1,2}$ ,  $\bar{u}(\cdot) \in L^2$ , i.e.,

$$Z_2(\Theta) = \mathbb{R} \times P_\Theta W^{1,2} \times L^2.$$

In this space, define the quadratic form

$$\Omega_{\Theta}(\bar{z}) = D(H)\bar{t}^2 - 2[L_x]\bar{x}_{\text{av}}\bar{t} + \int_{t_0}^{t_f} \langle L_{ww}(t, w^0(t))\bar{w}(t), \bar{w}(t) \rangle dt,$$

where  $[L_x]$  is the jump of the function  $L_x(t, w^0(t))$  at the point  $t_*$ , and

$$\bar{x}_{\text{av}} = \frac{1}{2} \left( \bar{x}(t_*-) + \bar{x}(t_*+) \right).$$

Set

$$\mathcal{K}_{\Theta} = \{ \bar{z} \in Z_2(\Theta) \mid \frac{d}{dt} \bar{x}(t) = \bar{u}(t) \text{ a.e.}, \quad [\bar{x}] + [\dot{x}^0]\bar{t} = 0, \quad \bar{x}(t_0) = \bar{x}(t_f) = 0 \}.$$

**Theorem 2.** (a) If  $x^0$  is a  $\Theta$ -weak minimum, then  $\Omega_{\Theta}(\bar{z}) \geq 0$  on  $\mathcal{K}_{\Theta}$ . (b) If  $\Omega_{\Theta}(\bar{z})$  is positive definite on  $\mathcal{K}_{\Theta}$ , then  $x^0$  is a (strict)  $\Theta$ -weak minimum.

The proof of this theorem is given in [28]. Let us note that in [28], instead of  $\bar{t}$ , we used a numerical parameter  $\bar{\xi}$  such that  $\bar{t} = -\bar{\xi}$ . This remark also applies to the subsequent presentation.

### 3 Second-Order Optimality Conditions for Discontinuous Controls in the General Problem of the Calculus of Variations on a Fixed Time Interval

#### 3.1 The General Problem in the Calculus of Variations on a Fixed Time Interval

Now consider the following optimal control problem in Mayer form on a fixed time interval  $[t_0, t_f]$ . It is required to find a pair of functions  $w(t) = (x(t), u(t))$ ,  $t \in [t_0, t_f]$ , minimizing the functional

$$\min \mathcal{J}(w) := J(x(t_0), x(t_f)) \quad (4)$$

subject to the constraints

$$\left. \begin{aligned} F(x(t_0), x(t_f)) \leq 0, \quad K(x(t_0), x(t_f)) = 0, \quad (x(t_0), x(t_f)) \in \mathcal{P}, \\ \dot{x}(t) = f(t, x(t), u(t)), \quad h(t, x(t), u(t)) = 0, \quad (t, x(t), u(t)) \in \mathcal{Q}, \end{aligned} \right\} \quad (5)$$

where  $\mathcal{P}$  and  $\mathcal{Q}$  are open sets,  $x$ ,  $u$ ,  $F$ ,  $K$ ,  $f$ , and  $h$  are vector-functions.

We assume that  $J$ ,  $F$ , and  $K$  are defined and twice continuously differentiable on  $\mathcal{P}$ , and  $f$  and  $h$  are defined and twice continuously differentiable on  $\mathcal{Q}$ . It is also

assumed that the gradients with respect to the control  $h_{iu}(t, x, u)$ ,  $i = 1, \dots, d(h)$  are *linearly independent* at each point  $(t, x, u) \in \mathcal{Q}$  such that  $h(t, x, u) = 0$  (the *regularity assumption* for the equality constraint  $h(t, x, u) = 0$ ). Here  $h_i$  are the components of the vector function  $h$  and  $d(h)$  is the dimension of this function.

Problem (1), (2) is considered in the space

$$\mathcal{W} := W^{1,1}([t_0, t_f], \mathbb{R}^n) \times L^\infty([t_0, t_f], \mathbb{R}^m),$$

where  $n = d(x)$ ,  $m = d(u)$ . Define a norm in this space as a sum of the norms:

$$\|w\| := \|x\|_{1,1} + \|u\|_\infty = |x(t_0)| + \int_{t_0}^{t_f} |\dot{x}(t)| dt + \text{esssup}_{[t_0, t_f]} |u(t)|.$$

A *weak minimum* is defined as a local minimum in the space  $\mathcal{W}$ . We say that  $w = (x, u)$  is an *admissible pair* if it belongs to  $\mathcal{W}$ , satisfies the constraints of the problem, and, moreover, there exists a compact set  $\mathcal{C} \subset \mathcal{Q}$  such that  $(t, x(t), u(t)) \in \mathcal{C}$  for a.a.  $t \in [t_0, t_f]$ .

It is well known that an optimal control problem with a functional in Bolza form,

$$\min \mathcal{J}(w) := J(x(t_0), x(t_f)) + \int_{t_0}^{t_f} f_0(t, x(t), u(t)) dt, \tag{6}$$

can be converted to Mayer form by introducing the ODE  $\dot{y} = f_0(t, x, u)$ ,  $y(t_0) = 0$ .

### 3.2 First-Order Necessary Conditions

Let  $w^0 = (x^0, u^0)$  be an *admissible pair*. We introduce the *Pontryagin function* (or the *Hamiltonian*)

$$H(t, x, u, \lambda) = \lambda f(t, x, u)$$

and the *augmented Pontryagin function* (or the *augmented Hamiltonian*)

$$H^a(t, x, u, \lambda, v) = H(t, x, u, \lambda) + v h(t, x, u),$$

where  $\lambda$  and  $v$  are row-vectors of the dimensions  $d(x) = n$  and  $d(h)$ , respectively. For brevity we set

$$x_0 = x(t_0), \quad x_f = x(t_f), \quad \eta = (x_0, x_f).$$

Denote by  $\mathbb{R}^{n*}$  the space of  $n$ -dimensional row-vectors. Define the *endpoint Lagrange function*

$$l(\eta, \alpha_0, \alpha, \beta) = \alpha_0 J(\eta) + \alpha F(\eta) + \beta K(\eta),$$

where  $\alpha_0 \in \mathbb{R}$ ,  $\alpha \in (\mathbb{R}^{d(F)})^*$ ,  $\beta \in (\mathbb{R}^{d(K)})^*$ . Introduce a tuple of *Lagrange multipliers*

$$\mu = (\alpha_0, \alpha, \beta, \lambda(\cdot), \nu(\cdot))$$

such that  $\lambda(\cdot) : [t_0, t_f] \rightarrow \mathbb{R}^{n^*}$  is absolutely continuous and  $\nu(\cdot) : [t_0, t_f] \rightarrow (\mathbb{R}^{d(h)})^*$  is measurable and bounded. Denote by  $\Lambda_0$  the set of the tuples  $\mu$  satisfying the following conditions at the point  $w^0$ :

$$\begin{aligned} \alpha_0 \geq 0, \quad \alpha \geq 0, \quad \alpha F(\eta^0) = 0, \quad \alpha_0 + \sum_{i=1}^{d(F)} \alpha_i + \sum_{j=1}^{d(K)} |\beta_j| = 1, \\ \dot{\lambda} = -H_x^a, \quad \lambda(t_0) = -l_{x_0}, \quad \lambda(t_f) = l_{x_f}, \quad H_u^a = 0, \end{aligned}$$

where  $\eta^0 = (x^0(t_0), x^0(t_f))$ , the derivatives  $l_{x_0}$  and  $l_{x_f}$  are at  $(\eta^0, \alpha_0, \alpha, \beta)$  and the derivatives  $H_x^a$ ,  $H_u^a$  are at  $(t, x^0(t), u^0(t), \lambda(t), \nu(t))$ ,  $t \in [t_0, t_f]$ . By  $\alpha_i$  and  $\beta_j$  we denote the components of the row vectors  $\alpha$  and  $\beta$ , respectively.

**Theorem 3.** *If  $w^0$  is a weak local minimum, then  $\Lambda_0$  is nonempty. Moreover,  $\Lambda_0$  is a finite dimensional compact set, and the projector  $(\alpha_0, \alpha, \beta, \lambda(\cdot), \nu(\cdot)) \rightarrow (\alpha_0, \alpha, \beta)$  is injective on  $\Lambda_0$ .*

The condition  $\Lambda_0 \neq \emptyset$  is called the *local Pontryagin minimum principle*, or the *Euler–Lagrange equation*. Let  $M_0$  be the set of all  $\mu = (\alpha_0, \alpha, \beta, \lambda(\cdot), \nu(\cdot)) \in \Lambda_0$  satisfying the *minimum condition* for a.a.  $t \in [t_0, t_f]$ :

$$H(t, x^0(t), u, \lambda(t)) \geq H(t, x^0(t), u^0(t), \lambda(t)) \quad \forall u \in U(t, x^0(t)),$$

where

$$U(t, x) := \{u \in \mathbb{R}^m \mid (t, x, u) \in \mathcal{Q}, h(t, x, u) = 0\}.$$

The condition  $M_0 \neq \emptyset$  is called the (*integral*) *Pontryagin minimum principle*, which is a necessary condition for the so-called Pontryagin minimum.

**Definition 2 (A.A. Milyutin).** The pair  $w^0$  affords a *Pontryagin minimum* if for any compact set  $\mathcal{C} \subset \mathcal{Q}$  there exists  $\varepsilon > 0$  such that  $\mathcal{J}(w) \geq \mathcal{J}(w^0)$  for all admissible pairs  $w(t) = (x(t), u(t))$  satisfying the conditions

$$\max_{[t_0, t_f]} |x(t) - x^0(t)| < \varepsilon, \quad \int_{t_0}^{t_f} |u(t) - u^0(t)| < \varepsilon, \quad (t, x(t), u(t)) \in \mathcal{C} \quad \text{a.e.}$$



### 3.3 Second-Order Necessary Conditions

Set

$$\mathscr{W}_2 := W^{1,2}([t_0, t_f], \mathbb{R}^n) \times L^2([t_0, t_f], \mathbb{R}^m),$$

Let  $\mathscr{K}$  be the set of all  $\bar{w} = (\bar{x}, \bar{u}) \in \mathscr{W}_2$  satisfying the following conditions:

$$\begin{aligned} J'(\eta^0)\bar{\eta} &\leq 0, \quad F'_i(\eta^0)\bar{\eta} \leq 0 \quad \forall i \in I_F(\eta^0), \quad K'(\eta^0)\bar{\eta} = 0, \\ \frac{d}{dt}\bar{x}(t) &= f_w(t, w^0(t))\bar{w}(t), \quad \text{for a.a. } t \in [t_0, t_f], \\ h_w(t, w^0(t))\bar{w}(t) &= 0, \quad \text{for a.a. } t \in [t_0, t_f], \end{aligned}$$

where  $\bar{\eta} = (\bar{x}(t_0), \bar{x}(t_f))$ ,  $I_F(\eta^0) := \{i : F_i(\eta^0) = 0\}$  is the set of active indices. Obviously,  $\mathscr{K}$  is a convex cone in the Hilbert space  $\mathscr{W}_2$ . We call it the *critical cone*.

Let us introduce a quadratic form in  $\mathscr{W}_2$ . For  $\mu \in \Lambda_0$  and  $\bar{w} = (\bar{x}, \bar{u}) \in \mathscr{W}_2$ , we set

$$\Omega(\mu, \bar{w}) = \langle l_{\eta\eta}^\mu(\eta^0)\bar{\eta}, \bar{\eta} \rangle + \int_{t_0}^{t_f} \langle H_{ww}^{a\mu}(t)\bar{w}(t), \bar{w}(t) \rangle dt,$$

where  $l_{\eta\eta}^\mu(\eta^0) = l_{\eta\eta}(\eta^0, \alpha_0, \alpha, \beta)$ ,  $H_{ww}^{a\mu}(t) = H_{ww}^a(t, x^0(t), u^0(t), \lambda(t), v(t))$ , and  $\bar{\eta} = (\bar{x}(t_0), \bar{x}(t_f))$ .

**Theorem 4.** *If  $w^0$  is a weak minimum, then the set  $\Lambda_0$  is nonempty and*

$$\max_{\mu \in \Lambda_0} \Omega(\mu, \bar{w}) \geq 0 \quad \text{for all } \bar{w} \in \mathscr{K}.$$

The necessary condition for a Pontryagin minimum differs from this condition only by replacing the set  $\Lambda_0$  by the set  $M_0$ .

**Theorem 5.** *If  $w^0$  is a Pontryagin minimum, then the set  $M_0$  is nonempty and*

$$\max_{\mu \in M_0} \Omega(\mu, \bar{w}) \geq 0 \quad \text{for all } \bar{w} \in \mathscr{K}.$$

We now assume that the control  $u^0$  is a piecewise continuous function on  $[t_0, t_f]$  with the set of discontinuity points  $\Theta = \{t_1, \dots, t_s\}$ ,  $t_0 < t_1 < \dots < t_s < t_f$ . We also assume that each  $t_k \in \Theta$  is an  $L$ -point of the function  $u^0$  (see the definition in Sect. 2.2). In this case, the regularity assumption for  $h$  implies that, for any  $\mu = (\alpha_0, \alpha, \beta, \lambda(\cdot), v(\cdot)) \in \Lambda_0$ ,  $v(t)$  has the same properties as  $u^0(t)$ : the function  $v(t)$  is piecewise continuous and each of its point of discontinuity is an  $L$ -point which belongs to  $\Theta$ . By virtue of the adjoint equation  $\dot{\lambda} = -H_x^a$ , the same is true for the derivative  $\dot{\lambda}(t)$  of the adjoint variable  $\lambda$ . Now, the second-order necessary conditions can be refined as follows.

For  $\mu \in M_0$ , set

$$D^k(H^{a\mu}) = \dot{\lambda}^{k+} \dot{x}^{0k-} - \dot{\lambda}^{k-} \dot{x}^{0k+} - [H_t^{a\mu}]^k, \quad (7)$$

where  $[H_t^{a\mu}]^k$  is the jump of the derivative  $H_t^a(t, x^0(t), u^0(t), \lambda(t), v(t))$  at the point  $t_k$ , and  $\dot{\lambda}^{k-} := \dot{\lambda}(t_{k-})$ ,  $\dot{\lambda}^{k+} := \dot{\lambda}(t_{k+})$ , etc. Note that  $H_t^a = -\dot{\lambda}_0$ , where  $\lambda_0(t) = -H(t)$ , and hence  $-[H_t^a]^k = [\dot{\lambda}_0]^k$ . Sometimes we omit the superscript  $\mu$  in the notation  $D^k(H^{a\mu})$ .

We can calculate  $D^k(H^a)$  using another method. Namely,  $D^k(H^a)$  can be calculated as the derivative of the “jump of  $H^a$ ” at the point  $t_k$ . Introduce the function

$$\begin{aligned} (\Delta_k H^a)(t) &= (\Delta_k H)(t) + (\Delta_k(vh))(t) \\ &= \lambda(t) (f(t, x^0(t), u^{0k+}) - f(t, x^0(t), u^{0k-})) \\ &\quad + (v^{k+} h(t, x^0(t), u^{0k+}) - v^{k-} h(t, x^0(t), u^{0k-})). \end{aligned}$$

It can be shown that the function  $(\Delta_k H^a)(t)$  is continuously differentiable at the point  $t_k \in \Theta$ , and its derivative at this point coincides with  $-D^k(H^a)$ . Therefore, we can obtain the value of  $D^k(H^a)$  by calculating the left or right limit of the derivatives of the function  $(\Delta_k H^a)(t)$  at the point  $t_k$ :

$$D^k(H^a) = -\frac{d}{dt}(\Delta_k H^a)(t_{k\pm}).$$

For any  $\mu \in M_0$ , it can be shown that  $D^k(H^{a\mu}) \geq 0$ ,  $k = 1, \dots, s$ . Set

$$Z_2(\Theta) := \mathbb{R}^s \times P_\Theta W^{1,2}([t_0, t_f], \mathbb{R}^n) \times L^2([t_0, t_f], \mathbb{R}^m),$$

where  $P_\Theta W^{1,2}([t_0, t_f], \mathbb{R}^n)$  is the Hilbert space of piecewise continuous functions  $x(t)$ , absolutely continuous on each interval of the set  $[t_0, t_f] \setminus \Theta$  such that their first derivatives are square integrable. Define a *quadratic form* in  $Z_2(\Theta)$  as follows:

$$\begin{aligned} \Omega_\Theta(\mu, \bar{z}) &= \sum_{k=1}^s (D^k(H^{a\mu}) \bar{t}_k^2 + 2[\dot{\lambda}]^k \bar{x}_{av}^k \bar{t}_k) \\ &\quad + \langle l_{\bar{\eta}}^\mu(\eta^0) \bar{\eta}, \bar{\eta} \rangle + \int_{t_0}^{t_f} \langle H_{ww}^{a\mu}(t) \bar{w}(t), \bar{w}(t) \rangle dt, \end{aligned}$$

where  $\bar{z} = (\bar{\theta}, \bar{x}, \bar{u})$ ,  $\bar{\theta} = (\bar{t}_1, \dots, \bar{t}_s)$ ,  $\bar{\eta} = (\bar{x}(t_0), \bar{x}(t_f))$ ,  $\bar{x}_{av}^k = \frac{1}{2}(\bar{x}(t_{k-}) + \bar{x}(t_{k+}))$ ,  $\bar{w} = (\bar{x}, \bar{u})$ . Define the *critical cone*  $\mathcal{K}_\Theta$  in the same space by the relations

$$\begin{aligned} J'(\eta^0) \bar{\eta} &\leq 0, \quad F'_i(\eta^0) \bar{\eta} \leq 0 \quad \forall i \in I_F(\eta^0), \quad K'(\eta^0) \bar{\eta} = 0, \\ \frac{d}{dt} \bar{x}(t) &= f_w(t, w^0(t)) \bar{w}(t), \quad \text{for a.a. } t \in [t_0, t_f], \\ [\bar{x}]^k + [\dot{x}^0]^k \bar{t}_k &= 0, \quad k = 1, \dots, s \\ h_w(t, w^0(t)) \bar{w}(t) &= 0 \quad \text{for a.a. } t \in [t_0, t_f]. \end{aligned}$$

**Theorem 6.** *If  $w^0$  is a Pontryagin minimum, then the following Condition  $\mathcal{A}_\Theta$  holds: the set  $M_0$  is nonempty and*

$$\max_{\mu \in M_0} \Omega_\Theta(\mu, \bar{z}) \geq 0 \quad \text{for all } \bar{z} \in \mathcal{K}_\Theta.$$

Let us give another possible representation for the terms  $(D^k(H^{a\mu})\bar{t}_k^2 + 2[\lambda]^k \bar{x}_{av}^k \bar{t}_k)$  of the quadratic form  $\Omega_\Theta(\mu, \bar{z})$  on the critical cone  $\mathcal{K}$ .

**Lemma 1.** *Let  $\mu \in M_0$  and  $z = (\bar{\theta}, \bar{x}, \bar{u}) \in \mathcal{K}_\Theta$ . Then, for any  $k = 1, \dots, s$ , the following formula holds*

$$D^k(\bar{H}^\mu)\bar{t}_k^2 + 2[\lambda]^k \bar{x}_{av}^k \bar{t}_k = [\lambda_0 + \dot{\lambda} \dot{x}^0]^k \bar{t}_k^2 + 2[\lambda \bar{x}]^k \bar{t}_k. \quad (8)$$

*Proof.* Everywhere in this proof we will omit the subscript and superscript  $k$ . Taking into account that

$$D(H^a) = \dot{\lambda}^+ \dot{x}^{0-} - \dot{\lambda}^- \dot{x}^{0+} + [\lambda_0], \quad [\bar{x}] + [\dot{x}^0] \bar{t} = 0,$$

we obtain

$$\begin{aligned} D(H^a) \bar{t}^2 + 2[\lambda] \bar{x}_{av} \bar{t} &= \bar{t}^2 [\lambda_0] + \bar{t}^2 (\dot{\lambda}^+ \dot{x}^{0-} - \dot{\lambda}^- \dot{x}^{0+}) + 2\bar{t} [\lambda] \bar{x}_{av} \\ &= \bar{t}^2 [\lambda_0] + \bar{t}^2 ([\dot{\lambda} \dot{x}^0] - \dot{\lambda}^+ [\dot{x}^0] - \dot{\lambda}^- [\dot{x}^0]) + 2\bar{t} [\lambda] \bar{x}_{av} \\ &= \bar{t}^2 ([\lambda_0] + [\dot{\lambda} \dot{x}^0]) + \dot{\lambda}^+ [\bar{x}] \bar{t} + \dot{\lambda}^- [\bar{x}] \bar{t} + 2\bar{t} [\lambda] \bar{x}_{av} \\ &= [\lambda_0 + \dot{\lambda} \dot{x}^0] \bar{t}^2 + (\dot{\lambda}^+ (\bar{x}^+ - \bar{x}^-) + \dot{\lambda}^- (\bar{x}^+ - \bar{x}^-) + (\dot{\lambda}^+ - \dot{\lambda}^-) (\bar{x}^- + \bar{x}^+)) \bar{t} \\ &= [\lambda_0 + \dot{\lambda} \dot{x}^0] \bar{t}^2 + 2[\lambda \bar{x}] \bar{t}. \end{aligned}$$

### 3.4 Second-Order Sufficient Conditions

Here, we will formulate sufficient optimality conditions, but only in the case of *discontinuous* control  $u^0$ . Let again  $u^0$  be a piecewise continuous function with the set of discontinuity points  $\Theta$  and let each  $t_k \in \Theta$  be an  $L$ -point. A natural strengthening of the necessary condition  $\mathcal{A}$  in Theorem 6 turned out to be sufficient not only for the Pontryagin minimum, but also for the so-called bounded strong minimum. This type of minimum will be defined below.

**Definition 3.** The component  $x_i$  of the state vector  $x$  is called *inessential* if the functions  $f$  and  $h$  do not depend on  $x_i$  and the functions  $J$ ,  $F$ , and  $K$  are affine in  $x_i(t_0)$  and  $x_i(t_f)$ . Let  $\underline{x}$  denote the vector composed by essential components of vector  $x$ .

For instance, the integral functional  $\mathcal{J} = \int_{t_0}^{t_f} f_0(t, x, u) dt$  can be brought to the endpoint form:  $\mathcal{J} = y(t_f) - y(t_0)$ , where  $\dot{y} = f_0(t, x, u)$ . Clearly,  $y$  is unessential component.

**Definition 4.** An admissible pair  $w^0$  affords a *bounded strong minimum* if for any compact set  $\mathcal{C} \subset \mathcal{Q}$  there exists  $\varepsilon > 0$  such that  $\mathcal{J}(w) \geq \mathcal{J}(w^0)$  for all admissible pairs  $w(t) = (x(t), u(t))$  satisfying the conditions  $|x(t_0) - x^0(t_0)| < \varepsilon$ ,  $\max_{[t_0, t_f]} |\dot{x}(t) - \dot{x}^0(t)| < \varepsilon$  and  $(t, x(t), u(t)) \in \mathcal{C}$  a.e. on  $[t_0, t_f]$ .

**Definition 5.** An admissible pair  $w^0$  affords a *strong minimum* if there exists  $\varepsilon > 0$  such that  $\mathcal{J}(w) \geq \mathcal{J}(w^0)$  for all admissible pairs  $w(t) = (x(t), u(t))$  satisfying the conditions  $|x(t_0) - x^0(t_0)| < \varepsilon$  and  $\max_{[t_0, t_f]} |\dot{x}(t) - \dot{x}^0(t)| < \varepsilon$ .

The following assertion follows from the definitions.

**Lemma 2.** *If there exists a compact set  $\mathcal{C} \subset \mathcal{Q}$  such that  $\{(t, x, u) \in \mathcal{Q} : h(t, x, u) = 0\} \subset \mathcal{C}$ , then the bounded strong minimum is equivalent to the strong minimum.*

Let us formulate sufficient conditions for a bounded strong minimum. For  $\mu \in M_0$ , we introduce the following conditions of the *strict minimum principle*:

- (a)  $H(t, x^0(t), u, \lambda(t)) > H(t, x^0(t), u^0(t), \lambda(t))$   
for all  $t \in [t_0, t_f] \setminus \Theta$ ,  $u \neq u^0(t)$ ,  $u \in U(t, x^0(t))$ ,
- (b)  $H(t_k, x^0(t_k), u, \lambda(t_k)) > H^k$   
for all  $t_k \in \Theta$ ,  $u \in U(t_k, x^0(t_k))$ ,  $u \neq u^0(t_k-)$ ,  $u \neq u^0(t_k+)$ , where  
 $H^k := H(t_k, x^0(t_k), u^0(t_k-), \lambda(t_k)) = H(t_k, x^0(t_k), u^0(t_k+), \lambda(t_k))$ .

We denote by  $M_0^+$  the set of all  $\mu \in M_0$  satisfying conditions (a) and (b).

For  $\mu \in M_0$  we also introduce the *strengthened Legendre-Clebsch conditions*:

- (i) for each  $t \in [t_0, t_f] \setminus \Theta$  the quadratic form

$$\langle H_{uu}^a(t, x^0(t), u^0(t), \lambda(t), \nu(t))u, u \rangle$$

is positive definite on the subspace of vectors  $u \in \mathbb{R}^m$  such that

$$h_u(t, x^0(t), u^0(t))u = 0.$$

- (ii) for each  $t_k \in \Theta$ , the quadratic form

$$\langle \bar{H}_{uu}(t_k, x^0(t_k), u^0(t_k-), \lambda(t_k), \nu(t_k-))u, u \rangle$$

is positive definite on the subspace of vectors  $u \in \mathbb{R}^m$  such that

$$h_u(t_k, x^0(t_k), u^0(t_k-))u = 0.$$

- (iii) this condition is symmetric to condition (ii) by replacing  $(t_k-)$  everywhere by  $(t_k+)$ .

Note that for each  $\mu \in M_0$  the non-strengthened Legendre–Clebsch conditions hold, i.e., the same quadratic forms are *nonnegative* on the corresponding subspaces.

We denote by  $\text{Leg}_+(M_0^+)$  the set of all  $\mu \in M_0^+$  satisfying the strengthened Legendre–Clebsch conditions (i)–(iii) and also the conditions

(iv)  $D^k(H^{a\mu}) > 0$  for all  $k = 1, \dots, s$ .

Let us introduce the functional

$$\bar{\gamma}(\bar{z}) = \langle \bar{\theta}, \bar{\theta} \rangle + \langle \bar{x}(t_0), \bar{x}(t_0) \rangle + \int_{t_0}^{t_f} \langle \bar{u}(t), \bar{u}(t) \rangle dt,$$

where  $\bar{z} = (\bar{\theta}, \bar{x}, \bar{u})$  and  $\bar{\theta} = (\bar{t}_1, \dots, \bar{t}_s)$ .

**Theorem 7.** *For the pair  $w^0$ , assume that the following Condition  $\mathcal{B}_\Theta$  holds: the set  $\text{Leg}_+(M_0^+)$  is nonempty and there exist a nonempty compact set  $M \subset \text{Leg}_+(M_0^+)$  and a number  $C > 0$  such that*

$$\max_{\mu \in M} \Omega_\Theta(\mu, \bar{z}) \geq C \bar{\gamma}(\bar{z})$$

for all  $\bar{z} \in \mathcal{H}$ . Then the pair  $w^0$  affords a (strict) bounded strong minimum.

The sufficient condition  $\mathcal{B}_\Theta$  guarantees a certain growth condition for the cost which will be presented below. We define now the concept of the *order function*  $\Gamma(t, u)$ .

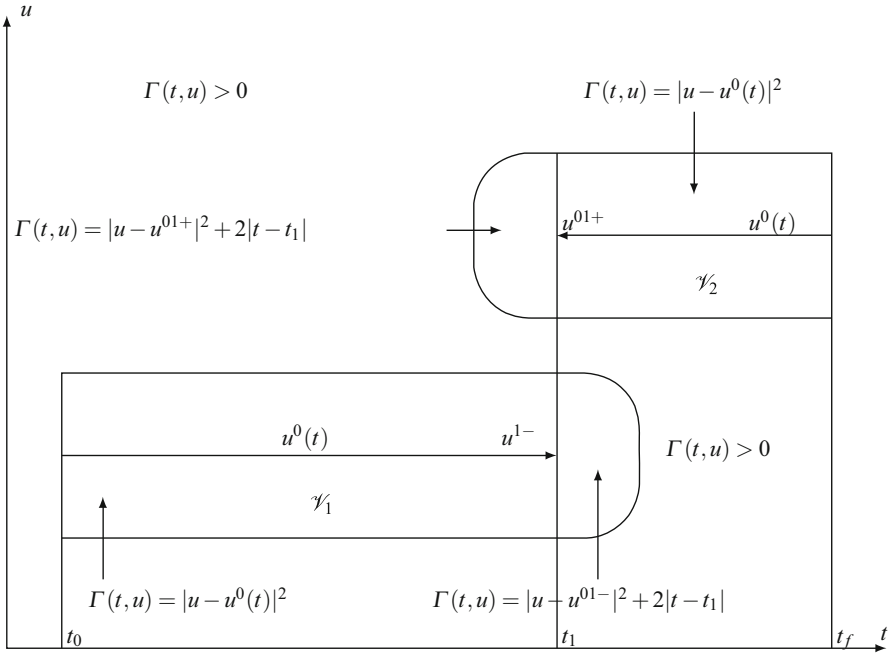
Assuming that the function  $u^0(t)$  is left-continuous, denote by  $\text{cl } u^0(\cdot)$  the closure (in  $\mathbb{R}^{m+1}$ ) of its graph. Denote by  $\text{cl } u^0(t_{k-1}, t_k)$  the closure in  $\mathbb{R}^{m+1}$  of the graph of the restriction of  $u^0(t)$  to the interval  $(t_{k-1}, t_k)$ ,  $k = 1, \dots, s+1$ , where  $t_{s+1} = t_f$ . Then

$$\text{cl } u^0(\cdot) = \bigcup_{k=1}^{s+1} \text{cl } u^0(t_{k-1}, t_k).$$

Denote by  $\mathcal{V}_k$ ,  $k = 1, \dots, s+1$ , a system of non-overlapping neighborhoods of the compact sets  $\text{cl } u^0(t_{k-1}, t_k)$ . Let  $\mathcal{V} = \bigcup_{k=1}^{s+1} \mathcal{V}_k$ .

**Definition 6.** The function  $\Gamma(t, u) : \mathbb{R}^{1+m} \rightarrow \mathbb{R}$  is said to be an *order function* if there exist disjoint neighborhoods  $\mathcal{V}_k$  of the compact sets  $\text{cl } u^0(t_{k-1}, t_k)$  such that the following five conditions hold (Fig. 1):

- (1)  $\Gamma(t, u) = |u - u^0(t)|^2$  if  $(t, u) \in \mathcal{V}_k$ ,  $t \in (t_{k-1}, t_k)$ ,  $k = 1, \dots, s+1$ ;
- (2)  $\Gamma(t, u) = 2|t - t_k| + |u - u^{0k-}|^2$  if  $(t, u) \in \mathcal{V}_k$ ,  $t > t_k$ ,  $k = 1, \dots, s$ ;
- (3)  $\Gamma(t, u) = 2|t - t_k| + |u - u^{0k+}|^2$  if  $(t, u) \in \mathcal{V}_{k+1}$ ,  $t < t_k$ ,  $k = 1, \dots, s$ ;
- (4)  $\Gamma(t, u) > 0$  if  $(t, u) \notin \mathcal{V}$ ;
- (5)  $\Gamma(t, u)$  is Lipschitz continuous on each compact set in  $\mathbb{R}^{1+m}$ .



**Fig. 1** Illustration of the order function  $\Gamma(t, u)$

For  $\delta w(t) = (\delta x(t), \delta u(t))$  in  $\mathscr{W}$  we set

$$\gamma(\delta w) = \|\delta x\|_\infty^2 + \int_{t_0}^{t_f} \Gamma(t, u^0(t) + \delta u(t)) dt.$$

We call  $\gamma$  the *higher order*. This higher order corresponds to a typical minimum in the case of discontinuous control, and the order function  $\Gamma(t, v)$  corresponds to a typical Hamiltonian in this case.

Note that the order  $\int_{t_0}^{t_f} (\Gamma(t, u^0(t) + \delta u(t)) dt$  is much finer (smaller) than the functional  $\int_{t_0}^{t_f} |\delta u(t)|^2 dt$ . On the other hand, it can be proved that, on each compact set (in  $\mathbb{R}^m$ ), the following lower bound holds for  $\int_{t_0}^{t_f} \Gamma(t, u^0(t) + \delta u(t)) dt$  :

$$\left( \int_{t_0}^{t_f} |\delta u(t)| dt \right)^2 \leq C \int_{t_0}^{t_f} \Gamma(t, u^0(t) + \delta u(t)) dt,$$

where  $C > 0$  depends only on the compact set.

Define the *violation function* at the point  $w^0$ :

$$V(\delta w) = (J(\eta^0 + \delta \eta) - J(\eta^0))_+ + \sum_{i=1}^{d(F)} F_i(\eta^0 + \delta \eta)_+ + |K(\eta^0 + \delta \eta)| + \|\delta \dot{x} - \delta f\|_1,$$

where  $\eta^0 = (x^0(t_0), x^0(t_f))$ ,  $\delta\eta = (\delta x(t_0), \delta x(t_f))$ ,  $\delta f = f(t, w^0 + \delta w) - f(t, w^0)$ ,  $\delta w = (\delta x, \delta u)$ ,  $\|\cdot\|_1$  is the norm in the space  $L^1$  of integrable functions, and  $a_+ := \max\{a, 0\}$  for  $a \in \mathbb{R}$ .

**Definition 7.** We say that a *bounded strong  $\gamma$ -sufficiency* holds at the point  $w^0$  if there exists  $C > 0$  such that for any compact set  $\mathcal{C} \subset \mathcal{Q}$  there exists  $\varepsilon > 0$  such that the inequality  $V(\delta w) \geq C\gamma(\delta w)$  holds for all  $\delta w = (\delta x, \delta u) \in \mathcal{W}$  satisfying the conditions

$$\left. \begin{aligned} |\delta x(t_0)| < \varepsilon, \|\delta x\|_\infty < \varepsilon, \\ (t, w^0(t) + \delta w(t)) \in \mathcal{C}, h(t, w^0(t) + \delta w(t)) = 0 \text{ a.e.} \end{aligned} \right\} \quad (9)$$

Obviously, a bounded strong  $\gamma$ -sufficiency implies a (strict) bounded strong minimum. Moreover, if the point  $w^0 + \delta w$  is admissible, then, obviously,  $V(\delta w) = (J(w^0 + \delta w) - J(w^0))_+$ . Therefore, a bounded strong  $\gamma$ -sufficiency implies the following:

*$\gamma$ -growth condition for the cost:* there exists  $C > 0$  such that for any compact set  $\mathcal{C} \subset \mathcal{Q}$  there exists  $\varepsilon > 0$  such that

$$J(w^0 + \delta w) - J(w^0) \geq C\gamma(\delta w)$$

for all  $\delta w = (\delta x, \delta u) \in \mathcal{W}$  satisfying (9) and such that  $(w^0 + \delta w)$  is an admissible pair.

**Theorem 8.** *The sufficient condition  $\mathcal{B}_\Theta$  in Theorem 7 is equivalent to the bounded strong  $\gamma$ -sufficiency.*

Theorems 6–8 were proved in [31]. Generalizations of these theorems for optimal control problem with regular mixed *inequality* state-control constraints were recently published in [33, 34]. An extension of the results of this section to problems on a variable time interval was obtained in [32].

## 4 The General Problem in the Calculus of Variations on a Variable Time Interval

### 4.1 Statement of the Problem

Here, quadratic optimality conditions, both necessary and sufficient, are presented in the following canonical problem on a variable time interval. Let  $\mathcal{T}$  denote a process  $(x(t), u(t) \mid t \in [t_0, t_f])$ , where the state variable  $x(\cdot)$  is a Lipschitz continuous function, and the control variable  $u(\cdot)$  is a bounded measurable function on a time interval  $\Delta = [t_0, t_f]$ . The interval  $\Delta$  is not fixed. For each process  $\mathcal{T}$ , we denote here by

$$\eta = (t_0, x(t_0), t_f, x(t_f))$$

the vector of the endpoints of time-state variable  $(t, x)$ . It is required to find  $\mathcal{T}$  minimizing the functional

$$\min \mathcal{J}(\mathcal{T}) := J(\eta) \quad (10)$$

subject to the constraints

$$F(\eta) \leq 0, \quad K(\eta) = 0, \quad \eta \in \mathcal{P}, \quad (11)$$

$$\dot{x}(t) = f(t, x(t), u(t)), \quad h(t, x(t), u(t)) = 0, \quad (t, x(t), u(t)) \in \mathcal{Q}, \quad (12)$$

where  $\mathcal{P}$  and  $\mathcal{Q}$  are open sets,  $x, u, F, K, f$ , and  $h$  are vector-functions.

We assume that the functions  $J, F$ , and  $K$  are defined and twice continuously differentiable on  $\mathcal{P}$ , and the functions  $f$  and  $h$  are defined and twice continuously differentiable on  $\mathcal{Q}$ . It is also assumed that the gradients with respect to the control  $h_{iu}(t, x, u)$ ,  $i = 1, \dots, d(h)$  are linearly independent at each point  $(t, x, u) \in \mathcal{Q}$  such that  $h(t, x, u) = 0$ . Here  $d(h)$  is a dimension of the vector  $h$ .

## 4.2 First-Order Necessary Conditions

We say that the function  $u(t)$  is *Lipschitz-continuous* if it is piecewise continuous and satisfies the Lipschitz condition on each interval of the continuity. Let

$$\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_f]) \quad (13)$$

be a fixed admissible process such that the control  $u(\cdot)$  is a piecewise Lipschitz-continuous function on the interval  $\Delta$  with the set of discontinuity points

$$\Theta = \{t_1, \dots, t_s\}, \quad \text{where } t_0 < t_1 < \dots < t_s < t_f.$$

In order to make the notations simpler we do not use here such symbols and indices as zero, hat or asterisk to distinguish the process  $\mathcal{T}$  from others.

Let us formulate the first-order necessary condition for optimality of the process  $\mathcal{T}$ . We introduce the Pontryagin function  $H$  (Hamiltonian), the *augmented* Pontryagin function  $H^a$ , and the endpoint Lagrange function  $l$  as in Sect. 3.2, but remember that now  $\eta = (t_0, x_0, t_f, x_f)$ . Also we introduce a tuple of Lagrange multipliers

$$\mu = (\alpha_0, \alpha, \beta, \lambda(\cdot), \lambda_0(\cdot), \nu(\cdot)) \quad (14)$$



such that  $\lambda(\cdot) : \Delta \rightarrow (\mathbb{R}^{d(x)})^*$  and  $\lambda_0(\cdot) : \Delta \rightarrow \mathbb{R}^1$  are piecewise smooth functions, continuously differentiable on each interval of the set  $\Delta \setminus \Theta$ , and  $v(\cdot) : \Delta \rightarrow (\mathbb{R}^{d(h)})^*$  is a piecewise continuous function and Lipschitz continuous on each interval of the set  $\Delta \setminus \Theta$ .

Denote by  $M_0$  the set of the normed tuples  $\mu$  satisfying the conditions of the *minimum principle* for the process  $\mathcal{T}$ :

$$\begin{aligned} &\alpha_0 \geq 0, \alpha \geq 0, \alpha F(\eta) = 0, \alpha_0 + \sum \alpha_i + \sum |\beta_j| = 1, \\ &\dot{\lambda} = -H_x^a, \dot{\lambda}_0 = -H_t^a, H_u^a = 0, t \in \Delta \setminus \Theta, \\ &\lambda(t_0) = -l_{x_0}, \lambda(t_f) = l_{x_f}, \lambda_0(t_0) = -l_{t_0}, \lambda_0(t_f) = l_{t_f}, \\ &\min_{u \in U(t,x(t))} H(t,x(t),u,\lambda(t)) = H(t,x(t),u(t),\lambda(t)), t \in \Delta \setminus \Theta, \\ &H(t,x(t),u(t),\lambda(t)) + \lambda_0(t) = 0, t \in \Delta \setminus \Theta, \end{aligned} \tag{15}$$

where  $U(t,x) = \{u \in \mathbb{R}^{d(u)} \mid h(t,x,u) = 0, (t,x,u) \in \mathcal{Q}\}$ . The derivatives  $l_{x_0}$  and  $l_{x_f}$  are at  $(\eta, \alpha_0, \alpha, \beta)$ , where  $\eta = (t_0, x(t_0), t_f, x(t_f))$ , and the derivatives  $H_x^a, H_u^a$ , and  $H_t^a$  are at  $(t, x(t), u(t), \lambda(t), v(t))$ , where  $t \in \Delta \setminus \Theta$ . (Condition  $H_u^a = 0$  follows from other conditions in this definition, and therefore, could be excluded; yet, we need to use it later.)

Let us give the definition of *Pontryagin minimum* in problem (10)–(12) on a variable interval  $[t_0, t_f]$ .

**Definition 8.** The process  $\mathcal{T}$  affords a *Pontryagin minimum* if for each compact set  $\mathcal{C} \subset \mathcal{Q}$  there exists  $\varepsilon > 0$  such that  $\mathcal{J}(\tilde{\mathcal{T}}) \geq \mathcal{J}(\mathcal{T})$  for all admissible processes  $\tilde{\mathcal{T}} = (\tilde{x}(t), \tilde{u}(t) \mid t \in [\tilde{t}_0, \tilde{t}_f])$  satisfying the conditions

- (a)  $|\tilde{t}_0 - t_0| < \varepsilon, \quad |\tilde{t}_f - t_f| < \varepsilon,$
- (b)  $\max_{\Delta \cap \tilde{\Delta}} |\tilde{x}(t) - x(t)| < \varepsilon, \quad \text{where } \tilde{\Delta} = [\tilde{t}_0, \tilde{t}_f],$
- (c)  $\int_{\tilde{\Delta} \cap \Delta} |\tilde{u}(t) - u(t)| dt < \varepsilon,$
- (d)  $(t, \tilde{x}(t), \tilde{u}(t)) \in \mathcal{C} \quad \text{a.e. on } \tilde{\Delta}.$

The condition  $M_0 \neq \emptyset$  is equivalent to Pontryagin’s minimum principle. It is the first-order necessary condition for Pontryagin minimum for the process  $\mathcal{T}$ . Thus, the following theorem holds.

**Theorem 9.** *If the process  $\mathcal{T}$  affords a Pontryagin minimum, then the set  $M_0$  is nonempty.*

Assume that the set  $M_0$  is nonempty. Using its definition and the full rank condition for the matrix  $h_u$  on the surface  $h = 0$  one can easily prove the following statement:

**Proposition 1.** *The set  $M_0$  is a finite-dimensional compact set, and the mapping  $\mu \mapsto (\alpha_0, \alpha, \beta)$  is injective on  $M_0$ .*

As in Sect. 3, for each  $\mu \in M_0, t_k \in \Theta$ , we define  $D^k(H^{a\mu})$  by relation (7). Then, for each  $\mu \in M_0$  the following inequalities hold:  $D^k(H^{a\mu}) \geq 0, \quad k = 1, \dots, s.$

### 4.3 Second-Order Necessary Conditions

Let us formulate a quadratic necessary condition for a Pontryagin minimum for the process  $\mathcal{T}$  as in (13). First, for this process, we introduce a Hilbert space  $\mathcal{Z}_2(\Theta)$  and the critical cone  $\mathcal{K} \subset \mathcal{Z}_2(\Theta)$ . Again, we denote by  $P_\Theta W^{1,2}(\Delta, \mathbb{R}^{d(x)})$  the Hilbert space of piecewise continuous functions  $\bar{x}(\cdot) : \Delta \rightarrow \mathbb{R}^{d(x)}$ , absolutely continuous on each interval of the set  $\Delta \setminus \Theta$  and such that their first derivative is square integrable. We set

$$\bar{z} = (\bar{t}_0, \bar{t}_f, \bar{\theta}, \bar{x}, \bar{u}),$$

where

$$\bar{t}_0 \in \mathbb{R}^1, \bar{t}_f \in \mathbb{R}^1, \bar{\theta} = (\bar{t}_1, \dots, \bar{t}_s) \in \mathbb{R}^s, \bar{x} \in P_\Theta W^{1,2}(\Delta, \mathbb{R}^{d(x)}), \bar{u} \in L^2(\Delta, \mathbb{R}^{d(u)}).$$

Thus,

$$\bar{z} \in \mathcal{Z}_2(\Theta) := \mathbb{R}^2 \times \mathbb{R}^s \times P_\Theta W^{1,2}(\Delta, \mathbb{R}^{d(x)}) \times L^2(\Delta, \mathbb{R}^{d(u)}).$$

Moreover, for given  $\bar{z}$  we set

$$\bar{w} = (\bar{x}, \bar{u}), \bar{x}_0 = \bar{x}(t_0), \bar{x}_f = \bar{x}(t_f), \quad (16)$$

$$\bar{\bar{x}}_0 = \bar{x}(t_0) + \bar{t}_0 \dot{\bar{x}}(t_0), \bar{\bar{x}}_f = \bar{x}(t_f) + \bar{t}_f \dot{\bar{x}}(t_f), \bar{\bar{\eta}} = (\bar{t}_0, \bar{\bar{x}}_0, \bar{t}_f, \bar{\bar{x}}_f). \quad (17)$$

By  $I_F(\eta) = \{i \in \{1, \dots, d(F)\} \mid F_i(\eta) = 0\}$  we denote the set of active indices of the constraints  $F_i \leq 0$ . Let  $\mathcal{K}_\Theta$  be the set of all  $\bar{z} \in \mathcal{Z}_2(\Theta)$  satisfying the following conditions:

$$\begin{aligned} J'(\eta) \bar{\eta} &\leq 0, \quad F'_i(\eta) \bar{\eta} \leq 0 \quad \forall i \in I_F(\eta), \quad K'(\eta) \bar{\eta} = 0, \\ \frac{d}{dt} \bar{x}(t) &= f_w(t, w(t)) \bar{w}(t), \quad \text{for a.a. } t \in [t_0, t_f], \\ [\bar{x}]^k + [\dot{\bar{x}}]^k \bar{t}_k &= 0, \quad k = 1, \dots, s, \\ h_w(t, w(t)) \bar{w}(t) &= 0, \quad \text{for a.a. } t \in [t_0, t_f]. \end{aligned} \quad (18)$$

Clearly,  $\mathcal{K}_\Theta$  is a convex cone in the Hilbert space  $Z_2(\Theta)$ . We call it the *critical cone*. If the interval  $\Delta$  is fixed, then we set  $\eta := (x_0, x_f) = (x(t_0), x(t_f))$ , and in the definition of  $\mathcal{K}$  we have  $\bar{t}_0 = \bar{t}_f = 0$ ,  $\bar{\bar{x}}_0 = \bar{x}_0$ ,  $\bar{\bar{x}}_f = \bar{x}_f$ , and  $\bar{\bar{\eta}} = \bar{\eta} := (\bar{x}_0, \bar{x}_f)$ .

Let us introduce a quadratic form on  $\mathcal{Z}_2(\Theta)$ . For  $\mu \in M_0$  and  $\bar{z} \in \mathcal{K}_\Theta$ , we set

$$\begin{aligned} 2\Omega_\Theta(\mu, \bar{z}) &= \langle l_{\eta\eta}^\mu \bar{\eta}, \bar{\eta} \rangle + \int_{t_0}^{t_f} \langle H_{ww}^{a\mu} \bar{w}(t), \bar{w}(t) \rangle dt + \sum_{k=1}^s \left( D^k (H^{a\mu}) \bar{t}_k^2 + 2[\dot{\lambda}]^k \bar{x}_{av}^k \bar{t}_k \right) \\ &\quad + \left( \dot{\lambda}(t_0) \dot{\bar{x}}(t_0) + \dot{\lambda}_0(t_0) \right) \bar{t}_0^2 + 2\dot{\lambda}(t_0) \bar{x}(t_0) \bar{t}_0 \\ &\quad - \left( \dot{\lambda}(t_f) \dot{\bar{x}}(t_f) + \dot{\lambda}_0(t_f) \right) \bar{t}_f^2 - 2\dot{\lambda}(t_f) \bar{x}(t_f) \bar{t}_f, \end{aligned} \quad (19)$$

where  $l_{\eta\eta}^\mu = l_{\eta\eta}(\eta, \alpha_0, \alpha, \beta)$ ,  $H_{ww}^{a\mu} = H_{ww}^a(t, x(t), u(t), \lambda(t), v(t))$ . We now formulate the main necessary quadratic condition of Pontryagin minimum in the problem on a variable time interval.

**Theorem 10.** *If the process  $\mathcal{T}$  yields a Pontryagin minimum, then the following Condition  $\mathcal{A}_\Theta$  holds: the set  $M_0$  is nonempty and*

$$\max_{\mu \in M_0} \Omega_\Theta(\mu, \bar{z}) \geq 0 \quad \text{for all } \bar{z} \in \mathcal{H}_\Theta.$$

Using (8), we can represent the quadratic form  $\Omega_\Theta$  on  $\mathcal{H}_\Theta$  as follows:

$$\begin{aligned} 2\Omega_\Theta(\mu, \bar{z}) &= \langle l_{\eta\eta}^\mu \bar{\eta}, \bar{\eta} \rangle + \int_{t_0}^{t_f} \langle H_{ww}^{a\mu} \bar{w}(t), \bar{w}(t) \rangle dt + \sum_{k=1}^s \left( [\dot{\lambda} \dot{x} + \dot{\lambda}_0]^k \bar{t}_k^2 + 2[\dot{\lambda} \bar{x}]^k \bar{t}_k \right) \\ &\quad + \left( \dot{\lambda}(t_0) \dot{x}(t_0) + \dot{\lambda}_0(t_0) \right) \bar{t}_0^2 + 2\dot{\lambda}(t_0) \bar{x}(t_0) \bar{t}_0 \\ &\quad - \left( \dot{\lambda}(t_f) \dot{x}(t_f) + \dot{\lambda}_0(t_f) \right) \bar{t}_f^2 - 2\dot{\lambda}(t_f) \bar{x}(t_f) \bar{t}_f. \end{aligned} \tag{20}$$

### 4.4 Second-Order Sufficient Conditions

Let us give the definition of a bounded strong minimum in problem (10)–(12) on a variable interval  $[t_0, t_f]$ . Let again  $\underline{x}$  denote a vector composed of all essential components of vector  $x$  (cf. Definition 3).

**Definition 9.** The process  $\mathcal{T}$  affords a bounded strong minimum if for each compact set  $\mathcal{C} \subset \mathcal{Q}$  there exists  $\varepsilon > 0$  such that  $\mathcal{J}(\tilde{\mathcal{T}}) \geq \mathcal{J}(\mathcal{T})$  for all admissible processes  $\tilde{\mathcal{T}} = (\tilde{x}(t), \tilde{u}(t) \mid t \in [\tilde{t}_0, \tilde{t}_f])$  satisfying the conditions

- (a)  $|\tilde{t}_0 - t_0| < \varepsilon, \quad |\tilde{t}_f - t_f| < \varepsilon, \quad |\tilde{x}(\tilde{t}_0) - x(t_0)| < \varepsilon,$
- (b)  $\max_{\tilde{\Delta} \cap \Delta} |\tilde{x}(t) - \underline{x}(t)| < \varepsilon, \quad \text{where } \tilde{\Delta} = [\tilde{t}_0, \tilde{t}_f],$
- (c)  $(t, \tilde{x}(t), \tilde{u}(t)) \in \mathcal{C} \quad \text{a.e. on } \tilde{\Delta}.$

The strict bounded strong minimum is defined in a similar way, with the nonstrict inequality  $\mathcal{J}(\tilde{\mathcal{T}}) \geq \mathcal{J}(\mathcal{T})$  replaced by the strict one and the process  $\tilde{\mathcal{T}}$  required to be different from  $\mathcal{T}$ .

Let us formulate a sufficient optimality Condition  $\mathcal{B}_\Theta$ , which is a natural strengthening of the necessary Condition  $\mathcal{A}_\Theta$ . The condition  $\mathcal{B}_\Theta$  is sufficient not only for a Pontryagin minimum, but also for a strict bounded strong minimum.

**Theorem 11.** *For the process  $\mathcal{T}$ , assume that the following Condition  $\mathcal{B}_\Theta$  holds: the set  $\text{Leg}_+(M_0^+)$  is nonempty and there exist a nonempty compact set  $M \subset \text{Leg}_+(M_0^+)$  and a number  $C > 0$  such that*

$$\max_{\mu \in M} \Omega_{\Theta}(\mu, \bar{z}) \geq C\bar{\gamma}(\bar{z}) \quad (21)$$

for all  $\bar{z} \in \mathcal{K}_{\Theta}$ . Then the process  $\mathcal{T}$  affords a strict bounded strong minimum.

Here the set  $\text{Leg}_+(M_0^+)$  has the same definition as in Sect. 3.4.

## 5 Second-Order Optimality Conditions for Bang-Bang Controls

### 5.1 Optimal Control Problems with Control Appearing Linearly

Let again  $\mathcal{T}$  denote a process  $(x(t), u(t) \mid t \in [t_0, t_f])$ , where the time interval  $\Delta = [t_0, t_f]$  is not fixed. As above, we set

$$\eta = (t_0, x(t_0), t_f, x(t_f)).$$

We will refer to the following control problem (22)–(24) as the *basic control problem*:

$$\text{Minimize } \mathcal{J}(\mathcal{T}) := J(\eta) \quad (22)$$

subject to the constraints

$$F(\eta) \leq 0, \quad K(\eta) = 0, \quad \eta \in \mathcal{P}, \quad (23)$$

$$\dot{x}(t) = f(t, x(t)) + g(t, x(t))u(t), \quad u(t) \in U, \quad (t, x(t)) \in \mathcal{Q}, \quad t_0 \leq t \leq t_f. \quad (24)$$

Here  $x \in \mathbb{R}^n$ ,  $u \in \mathbb{R}^m$ ,  $F$ ,  $K$ , and  $f$  are vector functions,  $g$  is  $n \times m$  matrix function with column vector functions  $g_1(t, x, u), \dots, g_m(t, x, u)$ ,  $\mathcal{P} \subset \mathbb{R}^{2n+2}$  and  $\mathcal{Q} \subset \mathbb{R}^{n+1}$  are open sets,  $U \subset \mathbb{R}^m$  is a convex polyhedron. The functions  $J$ ,  $F$ , and  $K$  are assumed to be twice continuously differential on  $\mathcal{P}$ , and the functions  $f$  and  $g$  are twice continuously differential on  $\mathcal{Q}$ .

A process  $\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_1])$  is said to be *admissible* if  $x(\cdot)$  is absolutely continuous,  $u(\cdot)$  is measurable bounded and the pair of functions  $(x(t), u(t))$  on the interval  $\Delta = [t_0, t_1]$  with the end-points  $\eta = (t_0, x(t_0), t_1, x(t_1))$  satisfies the constraints (23), (24).

Let us give the definition of Pontryagin minimum for the basic problem.

**Definition 10.** The process  $\widehat{\mathcal{T}} = (\widehat{x}(t), \widehat{u}(t) \mid t \in [\widehat{t}_0, \widehat{t}_f])$  affords a *Pontryagin minimum* in the basic problem if there exists  $\varepsilon > 0$  such that  $\mathcal{J}(\mathcal{T}) \geq \mathcal{J}(\widehat{\mathcal{T}})$  for all admissible processes  $\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_f])$  satisfying

$$|t_0 - \hat{t}_0| < \varepsilon, \quad |t_1 - \hat{t}_1| < \varepsilon, \quad \max_{\Delta \cap \hat{\Delta}} |x(t) - \hat{x}(t)| < \varepsilon, \quad \int_{\Delta \cap \hat{\Delta}} |u(t) - \hat{u}(t)| dt < \varepsilon,$$

where  $\Delta = [t_0, t_f]$ ,  $\hat{\Delta} = [\hat{t}_0, \hat{t}_f]$ .

Note that, for a fixed time interval  $\Delta$ , a Pontryagin minimum corresponds to an  $L^1$ -local minimum with respect to the control variable.

### 5.2 Necessary Optimality Conditions: The Minimum Principle of Pontryagin et al.

Let  $\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_f])$  be a fixed admissible process such that the control  $u(\cdot)$  is a piecewise constant function on the interval  $\Delta = [t_0, t_f]$ . Denote by

$$\Theta = \{t_1, \dots, t_s\}, \quad t_0 < t_1 < \dots < t_s < t_f,$$

the finite set of all discontinuity points (jump points) of the control  $u(t)$ . Then  $\dot{x}(t)$  is a piecewise continuous function whose discontinuity points belong to  $\Theta$ , and hence  $x(t)$  is a piecewise smooth function on  $\Delta$ .

Let us formulate the *Pontryagin minimum principle*, which is the first-order necessary condition for optimality of the process  $\mathcal{T}$ . The *Pontryagin function* has the form

$$H(t, x, u, \lambda) = \lambda f(t, x) + \lambda g(t, x)u = \lambda f(t, x) + \sum_{i=1}^m \lambda g_i(t, x)u_i, \tag{25}$$

where  $\lambda$  is a row-vector of the dimension  $d(\lambda) = d(x) = n$  while  $x, u, f, F$ , and  $K$  are column-vectors. The factor of the control  $u$  in the Pontryagin function is the *switching vector function*, a row vector of dimension  $d(u) = m$ . Set

$$\begin{aligned} \sigma(t, x, \lambda) &:= H_u(t, x, u, \lambda) &= \lambda g(t, x), \\ \sigma_i(t, x, \lambda) &:= H_{u_i}(t, x, u, \lambda) &= \lambda g_i(t, x), \quad i = 1, \dots, m, \\ \sigma_i(t) &:= \sigma_i(t, x(t), u(t)). \end{aligned} \tag{26}$$

The *endpoint Lagrange function* is

$$l(\alpha_0, \alpha, \beta, \eta) = \alpha_0 J(\eta) + \alpha F(\eta) + \beta K(\eta),$$

where  $\alpha$  and  $\beta$  are row-vectors with  $d(\alpha) = d(F)$ ,  $d(\beta) = d(K)$ , and  $\alpha_0$  is a number. By

$$\mu = (\alpha_0, \alpha, \beta, \lambda(\cdot), \lambda_0(\cdot))$$

we denote a tuple of Lagrange multipliers such that  $\lambda(\cdot) : \Delta \rightarrow \mathbb{R}^{n^*}$  and  $\lambda_0(\cdot) : \Delta \rightarrow \mathbb{R}$  are continuous on  $\Delta$  and continuously differentiable on each interval of the set  $\Delta \setminus \Theta$ .

Let  $M_0$  be the set of the normed collections  $\mu$  satisfying the conditions of Minimum Principle for the process  $\mathcal{T}$ :

$$\alpha_0 \geq 0, \quad \alpha \geq 0, \quad \alpha F(\eta) = 0, \quad \alpha_0 + \sum_{i=1}^{d(F)} \alpha_i + \sum_{j=1}^{d(K)} |\beta_j| = 1, \quad (27)$$

$$\dot{\lambda} = -H_x, \quad \dot{\lambda}_0 = -H_t \quad \forall t \in \Delta \setminus \Theta, \quad (28)$$

$$\lambda(t_0) = -l_{x_0}, \quad \lambda(t_f) = l_{x_f}, \quad \lambda_0(t_0) = -l_{t_0}, \quad \lambda_0(t_f) = l_{t_f}, \quad (29)$$

$$\min_{u \in U} H(t, x(t), u, \lambda(t)) = H(t, x(t), u(t), \lambda(t)) \quad \forall t \in \Delta \setminus \Theta, \quad (30)$$

$$H(t, x(t), u(t), \lambda(t)) + \lambda_0(t) = 0 \quad \forall t \in \Delta \setminus \Theta. \quad (31)$$

The derivatives  $l_{x_0}$  and  $l_{x_f}$  are taken at the point  $(\alpha_0, \alpha, \beta, \eta)$ , and the derivatives  $H_x, H_t$  are evaluated at the point  $(t, x(t), u(t), \lambda(t))$ . We use the simple abbreviation  $(t)$  for indicating all arguments  $(t, x(t), u(t), \lambda(t))$ ,  $t \in \Delta \setminus \Theta$ .

**Theorem 12.** *If the process  $\mathcal{T}$  affords a Pontryagin minimum, then the set  $M_0$  is nonempty. The set  $M_0$  is a finite-dimensional compact set and the projector  $\mu \mapsto (\alpha_0, \alpha, \beta)$  is injective on  $M_0$ .*

In view of this theorem, we can identify each tuple  $\mu \in M_0$  with its projection  $(\alpha_0, \alpha, \beta)$ . In what follows we set  $\mu = (\alpha_0, \alpha, \beta)$ . For each  $\mu \in M_0$  and  $t_k \in \Theta$ , we define again the quantity  $D^k(H^\mu)$ . Set

$$(\Delta_k H)(t) = H(t, x(t), u^{k+}, \lambda(t)) - H(t, x(t), u^{k-}, \lambda(t)) = \sigma(t) [u]^k. \quad (32)$$

For each  $\mu \in M_0$  the following equalities hold:

$$\frac{d}{dt}(\Delta_k H)|_{t=t_k^-} = \frac{d}{dt}(\Delta_k H)|_{t=t_k^+}, \quad k = 1, \dots, s.$$

Consequently, for each  $\mu \in M_0$  the function  $(\Delta_k H)(t)$  has a derivative at the point  $t_k \in \Theta$ . Set

$$D^k(H^\mu) = -\frac{d}{dt}(\Delta_k H)(t_k).$$

Then, for each  $\mu \in M_0$ , the minimum condition (30) implies the inequalities:

$$D^k(H^\mu) \geq 0, \quad k = 1, \dots, s. \quad (33)$$

As we know, the value  $D^k(H^\mu)$  can be written in the form

$$D^k(H^\mu) = -H_x^{k+}H_\lambda^{k-} + H_x^{k-}H_\lambda^{k+} - [H_t]^k = \dot{\lambda}^{k+}\dot{x}^{k-} - \dot{\lambda}^{k-}\dot{x}^{k+} + [\lambda_0]^k,$$

where  $H_x^{k-}$  and  $H_x^{k+}$  are the left-hand and the right-hand values of the function  $H_x(t) := H_x(t, x(t), u(t), \lambda(t))$  at  $t_k$ , respectively,  $[H_t]^k$  is a jump of the function  $H_t(t)$  at  $t_k$ , etc. It also follows from the above representation that we have

$$D^k(H^\mu) = -\dot{\sigma}(t_k \pm)[u]^k, \tag{34}$$

where the values on the right-hand side agree for the derivative  $\dot{\sigma}(t_k+)$  from the right and the derivative  $\dot{\sigma}(t_k-)$  from the left. In the case of a *scalar* control  $u$ , the total derivative  $\sigma_t + \sigma_x \dot{x} + \sigma_\lambda \dot{\lambda}$  does not contain the control variable explicitly and hence the derivative  $\dot{\sigma}(t)$  is *continuous* at  $t_k$ .

**Definition 11.** For a given extremal process  $\mathcal{T} = \{(x(t), u(t)) \mid t \in \Delta\}$  with a piecewise constant control  $u(t)$  we say that  $u(t)$  is a *strict bang-bang control* if there exists  $\mu = (\alpha_0, \alpha, \beta, \lambda, \lambda_0) \in M_0$  such that

$$\text{Arg min}_{u' \in U} \sigma(t)u' = [u(t-), u(t+)], \quad t \in [t_0, t_f] \tag{35}$$

where  $[u(t-), u(t+)]$  denotes the line segment spanned by the vectors  $u(t-)$  and  $u(t+)$  in  $\mathbb{R}^{d(u)}$  and  $\sigma(t) := \sigma(t, x(t), \lambda(t)) = \lambda(t)g(t, x(t))$ .

Note that  $[u(t-), u(t+)]$  is a singleton  $\{u(t)\}$  at each continuity point of the control  $u(t)$  with  $u(t)$  being a vertex of the polyhedron  $U$ . Only at the points  $t_k \in \Theta$  does the line segment  $[u^{k-}, u^{k+}]$  coincide with an edge of the polyhedron.

It is instructive to evaluate the condition (35) in greater detail when the control set is the *hypercube*

$$U = \prod_{i=1}^{d(u)} [u_{i,\min}, u_{i,\max}], \quad u_{i,\min} < u_{i,\max} \quad (i = 1, \dots, d(u)). \tag{36}$$

Let  $S_i = \{t_{k,i}, k = 1, \dots, k_i\}$ ,  $k_i \geq 0$ , be the set of switching times of the  $i$ -th control component  $u_i(t)$  and let  $\sigma_i(t) = \lambda(t)g_i(t, x(t))$  be switching function for  $u_i$ ,  $i = 1, \dots, d(u)$ . Then the set of all switching times is given by

$$\Theta = \{t_1, \dots, t_s\} = \bigcup_{i=1}^{d(u)} S_i,$$

and the condition (35) for a strict bang-bang control requires that

- (a)  $\sigma_i(t) \neq 0 \quad \forall t \notin S_i \quad (i = 1, \dots, d(u))$ ,
- (b) there is no simultaneous switching of control components  $u_i(t)$ , (37)  
i.e.,  $S_i \cap S_j = \emptyset \quad \forall i \neq j$ .

Hence, the  $i$ -th control component is determined by the control law

$$u_i(t) = \begin{cases} u_{i,\min}, & \text{if } \sigma_i(t) > 0 \\ u_{i,\max}, & \text{if } \sigma_i(t) < 0 \end{cases} \quad \forall t \in \Delta \setminus S_i. \quad (38)$$

*Remark 1.* There exist examples, where condition (a) in (37) is slightly violated as  $\sigma_i(t_f) = 0$  holds for certain control components  $u_i$ ; cf. the Rayleigh problem in Sect. 9.2 and the collision avoidance problem in Maurer et al. [27]. In this case, we require in addition that  $\dot{\sigma}_i(t_f) \neq 0$  holds to compensate for the condition  $\sigma_i(t_f) = 0$ . This property is fulfilled for the Rayleigh problem in Sect. 9.2 and the control problem in [27].

### 5.3 Second-Order Necessary Optimality Conditions

Here, we formulate quadratic necessary optimality conditions for a Pontryagin minimum for a given bang-bang control. (Their strengthening yields quadratic sufficient conditions for a strong minimum.) These quadratic conditions are based on the properties of a quadratic form on the critical cone.

Let again  $\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_f])$  be a fixed admissible process such that the control  $u(\cdot)$  is a piecewise constant function on the interval  $\Delta = [t_0, t_f]$ , and let  $\Theta = \{t_1, \dots, t_s\}$ ,  $t_0 < t_1 < \dots < t_s < t_f$ , be the set of discontinuity points of the control  $u(t)$ . For the process  $\mathcal{T}$ , we introduce the space  $\mathcal{Z}(\Theta)$  and the *critical cone*  $\mathcal{K}_\Theta \subset \mathcal{Z}(\Theta)$  as follows. Denote by  $P_\Theta C^1(\Delta, \mathbb{R}^{d(x)})$  the space of piecewise continuous functions  $\bar{x}(\cdot) : \Delta \rightarrow \mathbb{R}^{d(x)}$  that are continuously differentiable on each interval of the set  $\Delta \setminus \Theta$ . For each  $\bar{x} \in P_\Theta C^1(\Delta, \mathbb{R}^{d(x)})$  and for  $t_k \in \Theta$  we set  $\bar{x}^{k-} = \bar{x}(t_k-)$ ,  $\bar{x}^{k+} = \bar{x}(t_k+)$  and  $[\bar{x}]^k = \bar{x}^{k+} - \bar{x}^{k-}$ . Now set

$$\bar{z} = (\bar{t}_0, \bar{t}_f, \bar{\theta}, \bar{x}),$$

where  $\bar{t}_0, \bar{t}_f \in \mathbb{R}^1$ ,  $\bar{\theta} = (\bar{t}_1, \dots, \bar{t}_s) \in \mathbb{R}^s$ ,  $\bar{x} \in P_\Theta C^1(\Delta, \mathbb{R}^{d(x)})$ . Thus,

$$\bar{z} \in \mathcal{Z}(\Theta) := \mathbb{R}^2 \times \mathbb{R}^s \times P_\Theta C^1(\Delta, \mathbb{R}^{d(x)}).$$

For each  $\bar{z}$  we set

$$\bar{\bar{x}}_0 = \bar{x}(t_0) + \bar{t}_0 \dot{\bar{x}}(t_0), \quad \bar{\bar{x}}_f = \bar{x}(t_f) + \bar{t}_f \dot{\bar{x}}(t_f), \quad \bar{\bar{\eta}} = (\bar{t}_0, \bar{\bar{x}}_0, \bar{t}_f, \bar{\bar{x}}_f). \quad (39)$$

The vector  $\bar{\bar{\eta}}$  is considered as a column vector. Note that  $\bar{t}_0 = 0$ , respectively,  $\bar{t}_f = 0$  for *fixed* initial time  $t_0$ , respectively, final time  $t_f$ . Let

$$I_F(\eta) = \{i \in \{1, \dots, d(F)\} \mid F_i(\eta) = 0\}$$



be the set of indices of all active endpoint inequalities  $F_i \leq 0$  at the point  $\eta = (t_0, x(t_0), t_f, x(t_f))$ . Denote by  $\mathcal{K}_\Theta$  the set of all  $\bar{z} \in \mathcal{L}(\Theta)$  satisfying the following conditions:

$$J'(\eta)\bar{\eta} \leq 0, \quad F'_i(\eta)\bar{\eta} \leq 0 \forall i \in I_F(\eta), \quad K'(\eta)\bar{\eta} = 0, \quad (40)$$

$$\frac{d}{dt}\bar{x}(t) = (f_x(t, x(t)) + g_x(t, x(t))u(t))\bar{x}(t), \quad (41)$$

$$[\bar{x}]^k + [\dot{\bar{x}}]^k \bar{t}_k = 0, \quad k = 1, \dots, s. \quad (42)$$

It is obvious that  $\mathcal{K}_\Theta$  is a convex, finite-dimensional, and finite-faced cone in the space  $\mathcal{L}(\Theta)$ . We call it *the critical cone*. Each element  $\bar{z} \in \mathcal{K}_\Theta$  is uniquely defined by the numbers  $\bar{t}_0, \bar{t}_f$ , the vector  $\bar{\theta}$  and the initial value  $\bar{x}(t_0)$  of the function  $\bar{x}(t)$ . Two important properties of the critical cone are formulated in the next two propositions.

**Proposition 2.** *For any  $\mu \in M_0$  and  $\bar{z} \in \mathcal{K}_\Theta$ , we have*

$$\alpha_0 J'(\eta)\bar{\eta} = 0, \quad \alpha_i F'_i(\eta)\bar{\eta} = 0 \quad \forall i \in I_F(\eta).$$

**Proposition 3.** *Suppose that there exist  $\mu \in M_0$  with  $\alpha_0 > 0$ . Then adding the equalities  $\alpha_i F'_i(\eta)\bar{\eta} = 0 \forall i \in I_F(\eta)$  to the system (40)–(42) defining  $\mathcal{K}_\Theta$ , one can omit the inequality  $J'(\eta)\bar{\eta} \leq 0$  in that system without affecting  $\mathcal{K}_\Theta$ .*

Thus,  $\mathcal{K}_\Theta$  is defined by conditions (41), (42) and by the condition  $\bar{\eta} \in \mathcal{K}_\Theta^e$ , where  $\mathcal{K}_\Theta^e$  is the cone in  $\mathbb{R}^{2d(x)+2}$  given by (40). But if there exists  $\mu \in M_0$  with  $\alpha_0 > 0$ , then we can put

$$\mathcal{K}_\Theta^e = \{\bar{\eta} \in \mathbb{R}^{d(x)+2} \mid F'_i(\eta)\bar{\eta} \leq 0, \alpha_i F'_i(\eta)\bar{\eta} = 0 \forall i \in I_F(\eta), K'(\eta)\bar{\eta} = 0\}. \quad (43)$$

If, in addition,  $\alpha_i > 0$  holds for all  $i \in I_F(\eta)$ , then  $\mathcal{K}_\Theta^e$  is a subspace in  $\mathbb{R}^{d(x)+2}$ .

Let us introduce a quadratic form on the critical cone  $\mathcal{K}_\Theta$  defined by the conditions (40)–(42). For each  $\mu \in M_0$  and  $\bar{z} \in \mathcal{K}_\Theta$  we set

$$\begin{aligned} 2\Omega_\Theta(\mu, \bar{z}) &= \langle l_{\eta\eta}^\mu \bar{\eta}, \bar{\eta} \rangle + \int_{t_0}^{t_f} \langle H_{xx}^\mu \bar{x}(t), \bar{x}(t) \rangle dt + \sum_{k=1}^s \left( D^k(H^\mu) \bar{t}_k^2 + 2[\lambda]^{k, \text{av}} \bar{x}_{\text{av}}^k \bar{t}_k \right) \\ &\quad + \left( \dot{\lambda}(t_0) \dot{\bar{x}}(t_0) + \dot{\lambda}_0(t_0) \right) \bar{t}_0^2 + 2\dot{\lambda}(t_0) \bar{x}(t_0) \bar{t}_0 \\ &\quad - \left( \dot{\lambda}(t_f) \dot{\bar{x}}(t_f) + \dot{\lambda}_0(t_f) \right) \bar{t}_f^2 - 2\dot{\lambda}(t_f) \bar{x}(t_f) \bar{t}_f, \end{aligned} \quad (44)$$

where  $l_{\eta\eta}^\mu = l_{\eta\eta}(\eta, \alpha_0, \alpha, \beta)$ ,  $H_{xx}^\mu = H_{xx}(t, x(t), u(t), \lambda(t))$  and  $\bar{\eta}$  was defined in (39). Note that for a problem on a fixed time interval  $[t_0, t_f]$  we have  $\bar{t}_0 = \bar{t}_f = 0$ . The following theorem gives the main second-order necessary condition of optimality.

**Theorem 13.** *If the process  $\mathcal{T}$  affords a Pontryagin minimum, then the following Condition  $\mathcal{A}_\Theta$  holds: the set  $M_0$  is nonempty and  $\max_{\mu \in M_0} \Omega_\Theta(\mu, \bar{z}) \geq 0$  for all  $\bar{z} \in \mathcal{K}_\Theta$ .*

Using (8), we can also represent the quadratic form  $\Omega_\Theta$  as follows:

$$\begin{aligned} 2\Omega_\Theta(\mu, \bar{z}) &= \langle t_{\eta\eta}^\mu \bar{\eta}, \bar{\eta} \rangle + \int_{t_0}^{t_f} \langle H_{xx}^\mu \bar{x}(t), \bar{x}(t) \rangle dt + \sum_{k=1}^s \left( [\dot{\lambda} \dot{x} + \dot{\lambda}_0]^k \bar{t}_k^2 + 2[\dot{\lambda} \bar{x}]^k \bar{t}_k \right) \\ &\quad + \left( \dot{\lambda}(t_0) \dot{x}(t_0) + \dot{\lambda}_0(t_0) \right) \bar{t}_0^2 + 2\dot{\lambda}(t_0) \bar{x}(t_0) \bar{t}_0 \\ &\quad - \left( \dot{\lambda}(t_f) \dot{x}(t_f) + \dot{\lambda}_0(t_f) \right) \bar{t}_f^2 - 2\dot{\lambda}(t_f) \bar{x}(t_f) \bar{t}_f. \end{aligned} \quad (45)$$

#### 5.4 Second-Order Sufficient Optimality Conditions (SSC)

The state variable  $x_i$  is called *inessential* if the function  $f$  does not depend on  $x_i$  and the functions  $F, J, K$  are affine in  $x_{i0} := x_i(t_0)$  and  $x_{if} := x_i(t_f)$ . Let  $\underline{x}$  denote the vector of all essential components of state vector  $x$ . Let us define a strong minimum in the basic problem.

**Definition 12.** The process  $\mathcal{T}$  affords a *strong minimum* if there exists  $\varepsilon > 0$  such that  $\mathcal{J}(\tilde{\mathcal{T}}) \geq \mathcal{J}(\mathcal{T})$  for all admissible processes  $\tilde{\mathcal{T}} = (\tilde{x}(t), \tilde{u}(t) \mid t \in [\tilde{t}_0, \tilde{t}_f])$  satisfying the conditions

- (a)  $|\tilde{t}_0 - t_0| < \varepsilon, \quad |\tilde{t}_f - t_f| < \varepsilon, \quad |\tilde{x}(\tilde{t}_0) - x(t_0)| < \varepsilon,$
- (b)  $\max_{\tilde{\Delta} \cap \Delta} |\tilde{x}(t) - \underline{x}(t)| < \varepsilon, \quad \text{where } \tilde{\Delta} = [\tilde{t}_0, \tilde{t}_f],$

The *strict* strong minimum is defined in a similar way, with the non-strict inequality  $\mathcal{J}(\tilde{\mathcal{T}}) \geq \mathcal{J}(\mathcal{T})$  replaced by the strict one and the process  $\tilde{\mathcal{T}}$  required to be different from  $\mathcal{T}$ .

A natural strengthening of the necessary Condition  $\mathcal{A}_\Theta$  of Theorem 13 turns out to be a sufficient optimality condition not only for a Pontryagin minimum, but also for a strong minimum.

**Theorem 14.** *Let the following Condition  $\mathcal{B}_\Theta$  be fulfilled for the process  $\mathcal{T}$ :*

- (a)  $u(t)$  is a strict bang-bang control (i.e., there exists  $\mu \in M_0$  such that condition (35) holds),
- (b) there exists  $\mu \in M_0$  such that  $D^k(H^\mu) > 0, k = 1, \dots, s,$
- (c)  $\max_{\mu \in M_0} \Omega_\Theta(\mu, \bar{z}) > 0$  for all  $\bar{z} \in \mathcal{K}_\Theta \setminus \{0\}$ .

Then  $\mathcal{T}$  is a strict strong minimum.

Note that the condition (c) is automatically fulfilled if  $\mathcal{K}_\Theta = \{0\}$ , which gives a *first-order sufficient condition* for a strong minimum in the problem. Also note that the condition (c) is automatically fulfilled if there exists  $\mu \in M_0$  such that

$$\Omega_\Theta(\mu, \bar{z}) > 0 \text{ for all } \bar{z} \in \mathcal{K}_\Theta \setminus \{0\}. \tag{46}$$

Sufficient conditions for inequality (46) were obtained in [21] and [22] (see also [36], Section 6.3). Clearly, there is *no gap* between the necessary condition  $\mathcal{A}_\Theta$  of Theorem 13 and the sufficient condition  $\mathcal{B}_\Theta$  of Theorem 14.

## 6 Induced Optimization Problem for Bang-Bang Controls and the Verification of SSC

We continue our discussion of bang-bang controls. Second-order sufficient optimality conditions for bang-bang controls had been derived in the literature in two different forms. The first form was discussed in the last section. The second one is due to Agrachev et al. [1], who first reduce the bang-bang control problem to a finite-dimensional optimization problem and then show that the well-known sufficient optimality conditions for this optimization problem supplemented by the strict bang-bang property furnish sufficient conditions for the bang-bang control problem. The bang-bang control problem, considered in this section, is more general than that in [1]. Following [35], we claim the equivalence of both forms of second-order conditions for this problem.

### 6.1 Formulation of the Induced Optimization Problem and Necessary Optimality Conditions

Let  $\widehat{\mathcal{F}} = (\widehat{x}(t), \widehat{u}(t) \mid t \in [\widehat{t}_0, \widehat{t}_f])$  be an admissible process for the basic control problem (22)–(24). We denote by  $\text{ex } U$  the set of vertices of the polyhedron  $U$ . Assume that  $\widehat{u}(t)$  is a bang-bang control in  $\widehat{\Delta} = [\widehat{t}_0, \widehat{t}_f]$  taking values in the set  $\text{ex } U$ ,

$$\widehat{u}(t) = u^k \in \text{ex } U \quad \text{for } t \in (\widehat{t}_{k-1}, \widehat{t}_k), \quad k = 1, \dots, s+1,$$

where  $\widehat{t}_{s+1} = \widehat{t}_f$ . Thus,  $\widehat{\Theta} = \{\widehat{t}_1, \dots, \widehat{t}_s\}$  is the set of switching points of the control  $\widehat{u}(\cdot)$  with  $\widehat{t}_k < \widehat{t}_{k+1}$  for  $k = 0, 1, \dots, s$ . Assume now that the set  $M_0$  of multipliers is nonempty for the process  $\widehat{\mathcal{F}}$ . Put

$$\widehat{x}(\widehat{t}_0) = \widehat{x}_0, \quad \widehat{\theta} = (\widehat{t}_1, \dots, \widehat{t}_s), \quad \widehat{\zeta} = (\widehat{t}_0, \widehat{t}_f, \widehat{x}_0, \widehat{\theta}). \tag{47}$$

Then  $\widehat{\theta} \in \mathbb{R}^s$ ,  $\widehat{\zeta} \in \mathbb{R}^2 \times \mathbb{R}^n \times \mathbb{R}^s$ , where  $n = d(x)$ .

Take a small neighborhood  $\mathcal{V}$  of the point  $\widehat{\zeta}$  in  $\mathbb{R}^2 \times \mathbb{R}^n \times \mathbb{R}^s$ , and let

$$\zeta = (t_0, t_f, x_0, \theta) \in \mathcal{V},$$

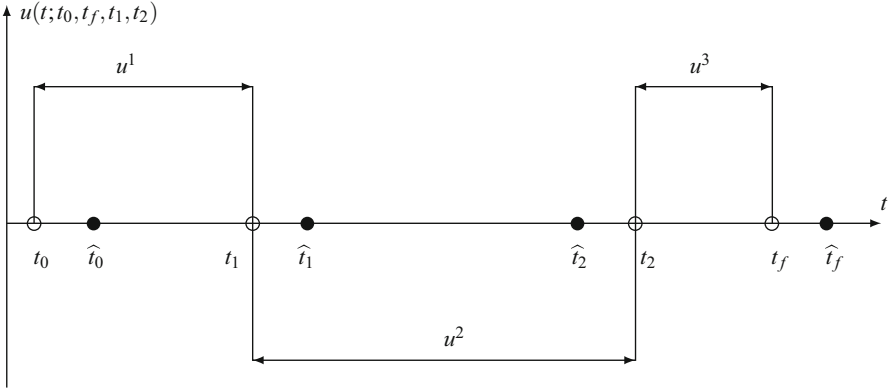


Fig. 2 Bang-bang control with two switches

where  $\theta = (t_1, \dots, t_s)$  satisfies  $t_0 < t_1 < t_2 < \dots < t_s < t_f$ . Define the function  $u(t; t_0, t_f, \theta)$  by the condition (Fig. 2)

$$u(t; t_0, t_f, \theta) = u^k \quad \text{for } t \in (t_{k-1}, t_k), \quad k = 1, \dots, s + 1, \quad (48)$$

where  $t_{s+1} = t_f$ . The values  $u(t_k; t_0, t_f, \theta)$ ,  $k = 1, \dots, s$ , may be chosen in  $U$  arbitrarily. For definiteness, define them by the condition of continuity of the control from the left:  $u(t_k; t_0, t_f, \theta) = u(t_k^-; t_0, t_f, \theta)$ ,  $k = 1, \dots, s$ .

Let  $x(t; t_0, t_f, x_0, \theta)$  be the solution of the initial value problem (IVP)

$$\dot{x} = f(t, x) + g(t, x)u(t; t_0, t_f, \theta), \quad t \in [t_0, t_f], \quad x(t_0) = x_0. \quad (49)$$

For each  $\zeta \in \mathcal{V}$  this solution exists if the neighborhood  $\mathcal{V}$  of the point  $\hat{\zeta}$  is sufficiently small. We obviously have

$$x(t; \hat{t}_0, \hat{t}_f, \hat{x}_0, \hat{\theta}) = \hat{x}(t), \quad t \in \hat{\Delta}, \quad u(t; \hat{t}_0, \hat{t}_f, \hat{\theta}) = \hat{u}(t), \quad t \in \hat{\Delta} \setminus \hat{\Theta}.$$

Consider now the following finite-dimensional optimization problem in the space  $\mathbb{R}^2 \times \mathbb{R}^n \times \mathbb{R}^s$  of the variables  $\zeta = (t_0, t_f, x_0, \theta)$ :

$$\begin{aligned} \mathcal{J}(\zeta) &:= J(t_0, x_0, t_f, x(t_f; t_0, t_f, x_0, \theta)) \rightarrow \min, \\ \mathcal{F}(\zeta) &:= F(t_0, x_0, t_f, x(t_f; t_0, t_f, x_0, \theta)) \leq 0, \\ \mathcal{G}(\zeta) &:= K(t_0, x_0, t_f, x(t_f; t_0, t_f, x_0, \theta)) = 0. \end{aligned} \quad (50)$$

We call (50) the *Induced Optimization Problem (IOP)* or simply *Induced Problem* which represents an extension of the IOP introduced in [1]. The following assertion is almost obvious.

**Theorem 15.** *Let the process  $\widehat{\mathcal{F}}$  be a Pontryagin local minimum for the basic control problem (22)–(24). Then the point  $\hat{\zeta}$  is a local minimum of the IOP (50), and hence it satisfies first and second-order necessary conditions for this problem.*

### 6.2 Second-Order Optimality Conditions for Bang-Bang Controls in Terms of the Induced Optimization Problem

We shall clarify a relationship between the second-order conditions for the Induced Optimization Problem (50) at the point  $\hat{\zeta}$  and those in the basic bang-bang control problem (22)–(24) for the process  $\widehat{\mathcal{F}}$ . It turns out that there is a one-to-one correspondence between Lagrange multipliers in these problems and a one-to-one correspondence between elements of the critical cones. Moreover, for corresponding Lagrange multipliers, the quadratic forms in these problems take equal values on the corresponding elements of the critical cones. This allows to express the necessary and sufficient quadratic optimality conditions for a bang-bang control, formulated in Theorems 13 and 14, in terms of the IOP (50). Thus we are able to establish the equivalence between our quadratic sufficient conditions and those due to Agrachev et al. [1].

Let  $\widehat{\mathcal{F}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_f])$  be an admissible process in the basic problem with the properties assumed in Sect. 5.2 and let  $\hat{\zeta} = (\hat{t}_0, \hat{t}_f, \hat{x}_0, \hat{\theta})$  be the corresponding admissible point in the IOP. The Lagrange function for the Induced Optimization Problem (50) is

$$\mathcal{L}(\mu, \zeta) = \mathcal{L}(\mu, t_0, t_f, x_0, \theta) = \alpha_0 \mathcal{J}(\zeta) + \alpha \mathcal{F}(\zeta) + \beta \mathcal{G}(\zeta), \tag{51}$$

where  $\mu = (\alpha_0, \alpha, \beta)$ ,  $\zeta = (t_0, t_f, x_0, \theta)$ ,  $\theta = (t_1, \dots, t_s)$ . We denote by  $\mathcal{H}_0$  the critical cone at the point  $\hat{\zeta}$  in the IOP. Thus,  $\mathcal{H}_0$  is the set of collections  $\bar{\zeta} = (\bar{t}_0, \bar{t}_f, \bar{x}_0, \bar{\theta})$  such that

$$\mathcal{J}'(\hat{\zeta})\bar{\zeta} \leq 0, \quad \mathcal{F}'_i(\hat{\zeta})\bar{\zeta} \leq 0, \quad i \in I, \quad \mathcal{G}'(\hat{\zeta})\bar{\zeta} = 0, \tag{52}$$

where  $I = \{i \mid \mathcal{F}_i(\hat{\zeta}) = 0\}$  is the set of indices of the inequality constraints active at the point  $\hat{\zeta}$ . For  $\mu \in M_0$  the quadratic form, of the induced optimization problem, is equal to  $\langle \mathcal{L}_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta} \rangle$ .

Let us formulate now second-order optimality conditions for the basic control problem in terms of the IOP.

**Theorem 16 (Second-Order Necessary Conditions).** *If the process  $\widehat{\mathcal{F}}$  affords a Pontryagin minimum in the basic problem, then the following Condition  $\mathcal{A}_0$  holds: the set  $M_0$  is nonempty and*

$$\max_{\mu \in M_0} \langle \mathcal{L}_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta} \rangle \geq 0 \quad \text{for all } \bar{\zeta} \in \mathcal{H}_0.$$

**Theorem 17 (Second-Order Sufficient Conditions).** *Let the following Condition  $\mathcal{B}_0$  be fulfilled for an admissible process  $\widehat{\mathcal{F}}$  in the basic problem:*

- (a)  $\hat{u}(t)$  is a strict bang-bang control with finitely many switching times  $\hat{t}_k$ ,  $k = 1, \dots, s$  (hence, the set  $M_0$  is nonempty and condition (35) holds for some  $\mu \in M_0$ ),
- (b) there exists  $\mu \in M_0$  such that  $D^k(H^\mu) > 0$ ,  $k = 1, \dots, s$ ,
- (c)  $\max_{\mu \in M_0} \langle \mathcal{L}_{\zeta\zeta}(\mu, \hat{\zeta}) \bar{\zeta}, \bar{\zeta} \rangle > 0$  for all  $\bar{\zeta} \in \mathcal{K}_0 \setminus \{0\}$ .

Then  $\widehat{\mathcal{F}}$  is a strict strong minimum in the basic problem.

Theorem 17 is a generalization of sufficient optimality conditions for bang-bang controls obtained in Agrachev et al. [1]. Detailed proofs of Theorems 16 and 17 are given in [35] and in our book [36].

*Remark 2.* Noble and Schättler [29] and Schättler and Ledzewicz [38] develop sufficient optimality conditions using methods of geometric optimal control. There is some evidence that their sufficient conditions are closely related to the SSC in our work [35, 36]. However, a formal proof of the equivalence of both types of sufficient conditions has not yet been worked out.

### 6.3 Numerical Methods for Solving the Induced Optimization Problem

The arc-parametrization method developed in [16, 26] provides an efficient method for solving the IOP. To better explain this method, for simplicity let us consider the basic control problem with fixed initial time  $t_0 = 0$  and fixed initial condition  $x_0(0) = x_0$ , and without inequality constraints  $\mathcal{F}(\zeta) \leq 0$ . For this problem, we slightly change the notation and replace the resulting optimization vector  $\zeta = (t_f, t_1, \dots, t_s)$  by the vector  $z = (t_1, \dots, t_s, t_{s+1})$ ,  $t_{s+1} = t_f$ . Instead of directly optimizing the switching times  $t_k$ ,  $k = 1, \dots, s$ , we determine the *arc lengths* (arc durations)

$$\xi_k := t_k - t_{k-1}, \quad k = 1, \dots, s, s+1, \quad (53)$$

of bang-bang arcs. Hence, the optimization variable  $z = (t_1, \dots, t_s, t_{s+1})^*$  is replaced by the optimization variable

$$\xi := (\xi_1, \dots, \xi_s, \xi_{s+1})^* \in \mathbb{R}^{s+1}, \quad \xi_k := t_k - t_{k-1}. \quad (54)$$

The variables  $z$  and  $\xi$  are related by a linear transformation involving the regular  $(s+1) \times (s+1)$ -matrix  $R$ :

$$\xi = Rz, \quad z = R^{-1}\xi, \quad R = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix}. \tag{55}$$

In the arc-parametrization method, the time interval  $[t_{k-1}, t_k]$  is mapped to the fixed interval

$$I_k := \left[ \frac{k-1}{s+1}, \frac{k}{s+1} \right], \quad k = 1, \dots, s+1, \tag{56}$$

by the linear transformation

$$t = a_k + b_k \tau, \quad \tau \in I_k, \tag{57}$$

where

$$a_k = t_{k-1} - (k-1)\xi_k, \quad b_k = (s+1)\xi_k. \tag{58}$$

Identifying

$$x(\tau) \cong x(a_k + b_k \tau) = x(t) \tag{59}$$

in the relevant intervals, we obtain the ODE system

$$\dot{x}(\tau) = (s+1)\xi_k (f(a_k + b_k \tau, x(\tau)) + g(a_k + b_k \tau, x(\tau))u^k) \quad \text{for } \tau \in I_k. \tag{60}$$

By concatenating the solutions in the intervals  $I_k$  we get the continuous solution  $x(t) = x(t; \xi)$  in the normalized interval  $[0, 1]$ . When expressed via the new optimization variable  $\xi$ , the Induced Optimization Problem (IOP) in (50) is equivalent to the following optimization problem ( $\widetilde{\text{IOP}}$ ) with  $t_f = \sum_{k=1}^{s+1} \xi_k$ :

$$\begin{aligned} &\text{Minimize } \widetilde{\mathcal{J}}(\xi) := J(t_f; x(1, \xi)) \\ &\text{subject to } \widetilde{\mathcal{G}}(\xi) := K(t_f; x(1, \xi)) = 0. \end{aligned} \tag{61}$$

The Lagrangian function is given by

$$\widetilde{\mathcal{L}}(\mu, \xi) = \alpha_0 \widetilde{\mathcal{J}}(\xi) + \beta \widetilde{\mathcal{G}}(\xi), \quad \mu = (\alpha_0, \beta). \tag{62}$$

Using the linear transformation (55), it can easily be seen that the SSCs for the Induced Optimization Problems (IOP) and ( $\widetilde{\text{IOP}}$ ) are equivalent; cf. similar

arguments in [26]. To solve the  $(\widetilde{\text{IOP}})$ , we use a suitable adaptation of the control package NUDOCSS in Büskens [4, 5]. Then we can take advantage of the fact that NUDOCSS also provides the Jacobian of the terminal constraints and the Hessian of the Lagrangian which are needed in the check of the second-order condition in Theorem 17.

In practice, we shall verify the positive definiteness condition (c) in Theorem 17 in a stronger form. We assume that the multiplier can be chosen as  $\mu = (1, \beta)$  and that the following *regularity condition* holds:

$$\text{rank } \widetilde{\mathcal{G}}_{\xi}(\hat{\xi}) = d(K).$$

Let  $N$  be the  $n_{\xi} \times (n_{\xi} - d(K))$  matrix,  $n_{\xi} = n + s + 1$  (where  $n = d(x)$ ), with full column rank  $n_{\xi} - d(K)$ , whose columns span the kernel of  $\widetilde{\mathcal{G}}_{\xi}(\hat{\xi})$ . Then condition (c) in Theorem 17,

$$\langle \widetilde{\mathcal{L}}_{\xi\xi}(\hat{\xi}, \beta) \bar{\xi}, \bar{\xi} \rangle > 0 \quad \forall \bar{\xi} \neq 0, \quad \widetilde{\mathcal{G}}_{\xi}(\hat{\xi}) \bar{\xi} = 0, \quad (63)$$

is equivalent to the condition that the *projected Hessian* is positive definite [6],

$$N^* \widetilde{\mathcal{L}}_{\xi\xi}(\hat{\xi}, \beta) N > 0. \quad (64)$$

## 7 Numerical Example with Fixed Final Time: Optimal Control of the Chemotherapy of HIV

The treatment of patients infected with the human immunodeficiency virus (HIV) is still of great concern today (Kirschner et al. [18]). The problem of determining optimal chemotherapies has been treated in Kirschner et al. [18] in the framework of optimal control theory. The optimal control model is based on a simple dynamic model in Perelson et al. [37] which simulates the interaction of the immune system with HIV. Kirschner et al. [18] use a control quadratic cost functional of  $L^2$ -type. It has been argued in Schättler et al. [39] that in a biological context it is more appropriate to consider cost functionals of  $L^1$ -type which are linear in the control variable. Therefore, in this section, we are studying an objective of  $L^1$ -type, where the quadratic control is replaced by a linear control. The state and control variables have the following meaning:

- $T(t)$ : concentration of uninfected  $\text{CD4}^+$  T cells,
- $T^*(t)$ : concentration of latently infected  $\text{CD4}^+$  T cells,
- $T^{**}(t)$ : concentration of actively infected  $\text{CD4}^+$  T cells,
- $V(t)$ : concentration of free infectious virus particles,
- $u(t)$ : control, rate of chemotherapy.



The treatment starts at  $t_0 = 0$  and terminates at the fixed final time  $t_f = 500$  (days). Thus, the control process is considered in the interval  $[0, t_f]$ . The dynamics of the populations are (omitting the time argument):

$$\begin{aligned}
 dT/dt &= \frac{s}{1+V} - \mu_T T + rT \left(1 - \frac{T+T^*+T^{**}}{T_{\max}}\right) - k_1 VT, & T(0) &= T_0, \\
 dT^*/dt &= k_1 VT - \mu_T T^* - k_2 T^*, & T^*(0) &= T_0^*, \\
 dT^{**}/dt &= k_2 T^* - \mu_b T^{**}, & T^{**}(0) &= T_0^{**} \\
 dV/dt &= (1-u)N\mu_b T^{**} - k_1 VT - \mu_V V, & V(0) &= V_0.
 \end{aligned}
 \tag{65}$$

The control constraint is given by

$$0 \leq u(t) \leq 1 \quad \forall t \in [0, t_f], \tag{66}$$

where  $u(t) = 1$  represents *maximal* chemotherapy, while  $u(t) = 0$  means that no chemotherapy is administered. Note that Kirschner et al. [18] consider the control variable  $v = 1 - u$ . It is convenient to write the ODE (65) as the control affine system (24),

$$\dot{x} = f(x) + g(x)u, \quad x(0) = x_0, \quad x = (T, T^*, T^{**}, V) \in \mathbb{R}^4, \tag{67}$$

with obvious definitions of the vector functions  $f(x)$  and  $g(x)$ . As in [18] we consider two sets of initial conditions which depend on the time at which the treatment starts after the infection. The following initial conditions are interpolated from [18] and have already been used in the [15].

**Initial conditions after 800 days:**

$$T_0 = 982.8, \quad T_0^* = 0.05155, \quad T_0^{**} = 0.0006175, \quad V_0 = 0.07306. \tag{68}$$

**Initial conditions after 1000 days:**

$$T_0 = 904.1, \quad T_0^* = 0.3447, \quad T_0^{**} = 0.004167, \quad V_0 = 0.4939. \tag{69}$$

The parameter and constants are taken from [18] and are listed in Table 1.

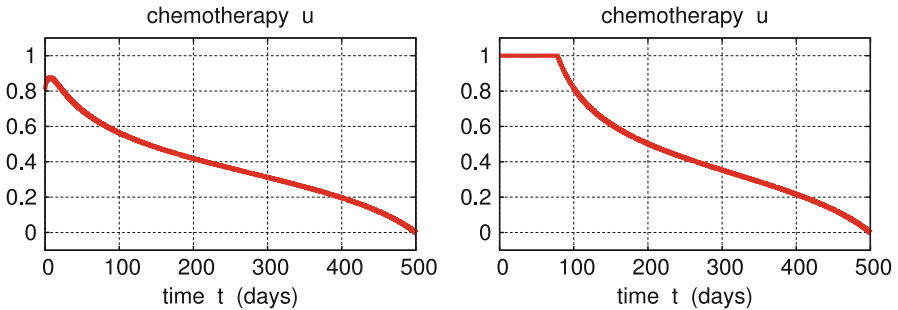
Kirschner et al. [18] consider the following objective of  $L^2$ -type which is quadratic in the control variable:

$$\text{Minimize } J(x, u) = \int_0^{t_f} (-T(t) + B u(t)^2) dt \quad (B = 50). \tag{70}$$

Recall that the state variable is defined as  $x := (T, T^*, T^{**}, V) \in \mathbb{R}^4$ . The optimal control that minimizes (70) subject to the constraints (65)–(69) is a *continuous*

**Table 1** Parameters and constants

Parameters and constants		Values
$\mu_T$	: death rate of uninfected CD4 <sup>+</sup> T cell population	0.02 d <sup>-1</sup>
$\mu_{T^*}$	: death rate of latently infected CD4 <sup>+</sup> T cell population	0.02 d <sup>-1</sup>
$\mu_b$	: death rate of actively infected CD4 <sup>+</sup> T cell population	0.24 d <sup>-1</sup>
$\mu_V$	: death rate of free virus	2.4 d <sup>-1</sup>
$k_1$	: rate CD4 <sup>+</sup> T cells becomes infected by free virus	$2.4 \times 10^{-5} \text{ mm}^3 \text{ d}^{-1}$
$k_2$	: rate T* cells convert to actively infected	$3 \times 10^{-3} \text{ mm}^3 \text{ d}^{-1}$
$r$	: rate of growth for the CD4 <sup>+</sup> T cell population	0.03 d <sup>-1</sup>
$N$	: number of free virus produced by T** cells	1200
$T_{\max}$	: maximum CD4 <sup>+</sup> T cell population level	$1.5 \times 10^3 \text{ mm}^{-3}$
$s$	: source term for uninfected CD4 <sup>+</sup> T cells, where $s$ is the parameter in the source term	$10 \text{ d}^{-1} \text{ mm}^{-3}$ $s/(1 + V)$



**Fig. 3** Optimal control for the  $L^2$  functional (70). (Left) begin of treatment after 800 days: initial conditions (68). (Right) begin of treatment after 1000 days: initial conditions (69)

function, since the associated Hamiltonian  $H(x, \lambda, u)$  has a unique minimum with respect to  $u$  and the strict Legendre-Clebsch condition  $H_{uu} = 2B > 0$  holds. For the two sets of initial conditions (68) and (69), the numerical discretization and NLP approach using AMPL [12] and IPOPT [42] yield the optimal controls shown in Fig. 3 which were also obtained in Hannemann [15].

Hannemann [15] showed that second-order sufficient conditions (SSC) are satisfied for the controls displayed in Fig. 3, since the associated matrix Riccati equation has a bounded solution. Note that Riccati equations are discussed in [19, 24] and in our book [36], Chap. 4.

Instead of the  $L^2$  functional (70) we consider now the functional of  $L^1$ -type:

$$\text{Minimize } J_1(x, u) = \int_0^{t_f} (-T(t) + Bu(t)) dt \quad (B = 50). \tag{71}$$

The Hamiltonian or Pontryagin function for this control problem is given by

$$H(x, \lambda, u) = -T + Bu + \lambda(f(x) + g(x)u), \tag{72}$$

where  $\lambda = (\lambda_T, \lambda_{T^*}, \lambda_{T^{**}}, \lambda_V) \in \mathbb{R}^4$  denotes the adjoint variable. The adjoint equation and transversality condition are given by

$$\dot{\lambda} = -H_x(x, \lambda, u), \quad \lambda(t_f) = (0, 0, 0, 0),$$

since the terminal state  $x(t_f)$  is free and the objective (71) does not contain a Mayer term. We do not write out the adjoint equation  $\dot{\lambda} = -H_x(x, \lambda, u)$  explicitly, since this equation is not needed in the sequel. The adjoint variables can be computed from the Lagrange multipliers of the associated Induced Optimization Problem (IOP). The *switching function* is given by

$$\sigma(x, \lambda) = H_u(x, \lambda, u) = B - \lambda_V N \mu_b T^{**}, \quad \sigma(t) = \sigma(x(t), \lambda(t)). \tag{73}$$

The minimization of the Hamiltonian with respect to  $u$  yields the switching condition

$$u(t) = \begin{cases} 1, & \text{if } \sigma(t) < 0 \\ 0, & \text{if } \sigma(t) > 0 \end{cases}. \tag{74}$$

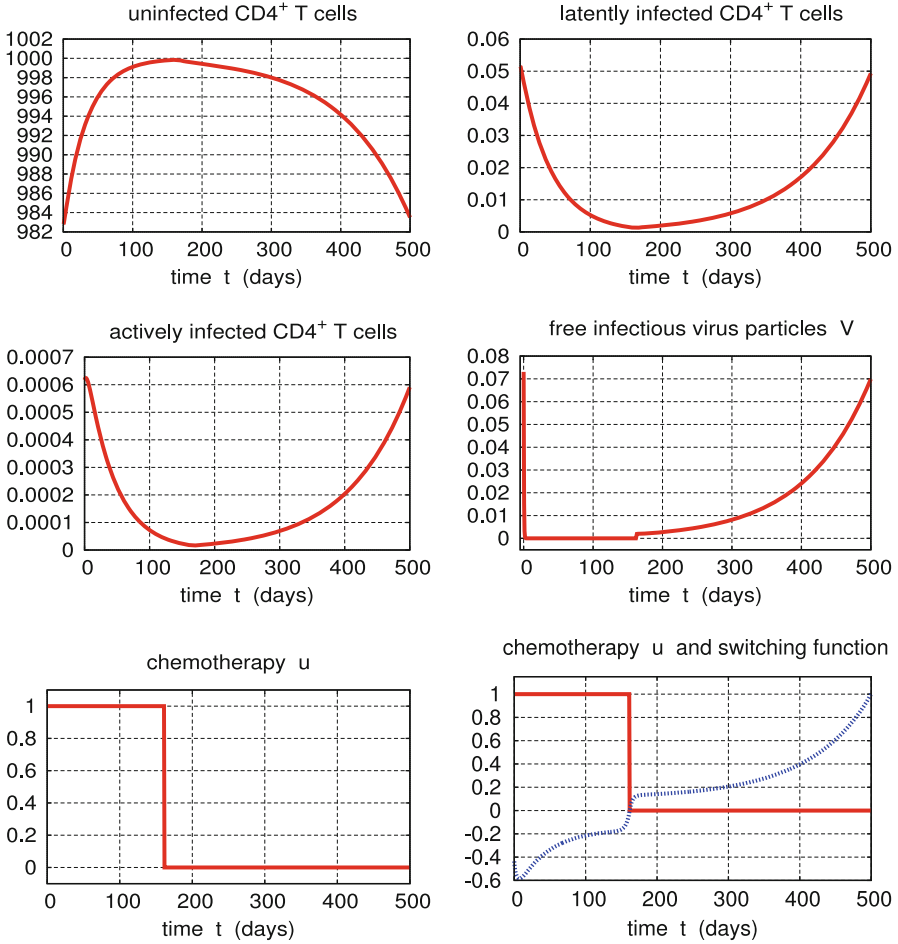
The control has a *singular arc* in an interval  $[t_1, t_2] \subset [0, T]$  if  $\sigma(t) = 0$  holds on  $[t_1, t_2]$ . However, we do not discuss singular controls further because for the data in Table 1 we never found singular arcs. Indeed, the optimal control for the  $L^1$ -functional (71) is the following bang-bang control with only one switch at  $t_1$  :

$$u(t) = \begin{cases} 1 & \text{for } 0 \leq t < t_1 \\ 0 & \text{for } t_1 \leq t \leq t_f \end{cases} \tag{75}$$

The terminal arc  $u(t) = 0$  results from the terminal value  $\sigma(t_f) = B > 0$  of the switching function and the minimum condition (74). Hence, the IOP has only the scalar optimization variable  $t_1$  and thus the objective (71) reduces to a function  $J_1(t_1) = J_1(x, u)$ . The arc-parametrization method [26, 36] and the code NUDOCSS [4] yield the following numerical results, where state variables are listed with 8 digits and adjoint variables with 6 digits.

$$\begin{aligned} J_1 &= -489810.5, & t_1 &= 161.6957, & T(t_f) &= 983.4926, \\ T^*(t_f) &= 0.04934668, & T^{**}(t_f) &= 0.0005910497, & V(t_f) &= 0.06993300, \\ \lambda_T(0) &= -0.125173, & \lambda_{T^*}(0) &= -1.51988, & \lambda_{T^{**}}(0) &= -2.94704, \\ \lambda_V(0) &= -0.449700. \end{aligned} \tag{76}$$

The state and control variables and the switching function are displayed in Fig. 4. To verify that the second-order sufficient conditions (SSC) are satisfied for the computed extremal solution, we have to check the conditions of Theorem 17. The strict bang-bang property is satisfied, since we infer from Fig. 4 (bottom, right) that the switching function satisfies



**Fig. 4** Optimal solution for initial conditions (68): treatment starts after 800 days. *Top row: (left) uninfected  $CD4^+$ T cells, (right) latently infected  $CD4^+T^*$  cells. Middle row: (left) actively infected  $CD4^+T^{**}$  cells, (right) infectious virus particles  $V$ . Bottom row: (left) bang-bang control  $u$ , (right) bang-bang control  $u$  and (scaled) switching function  $\sigma$  in (73) satisfying the switching condition (74)*

$$\sigma(t) < 0 \quad \text{for } 0 \leq t < t_1, \quad \dot{\sigma}(t_1) > 0, \quad \sigma(t) > 0 \quad \text{for } t_1 < t \leq t_f = 500. \quad (77)$$

To verify the positive definiteness in condition (63), we note that the Hessian is simply the second derivative of the objective  $J_1(t_1)$  evaluated at the optimal switching time  $t_1 = 161.695711$  for which we find

$$\frac{d^2 J_1}{dt_1^2} = 1.5469 > 0.$$

Hence, the extremal solution (76) displayed in Fig. 4 provides a strict strong minimum.

Now we try to improve the optimal terminal value  $T(t_f) = 983.493$  of the uninfected  $CDC4^+T$  cells. For that purpose we prescribe a higher terminal value and minimize the functional  $J_1(x, u)$  subject to the boundary condition

$$T(t_f) = 995. \tag{78}$$

The arc-parametrization method [26, 36] and the control package NUDOCCCS [4] furnish the results

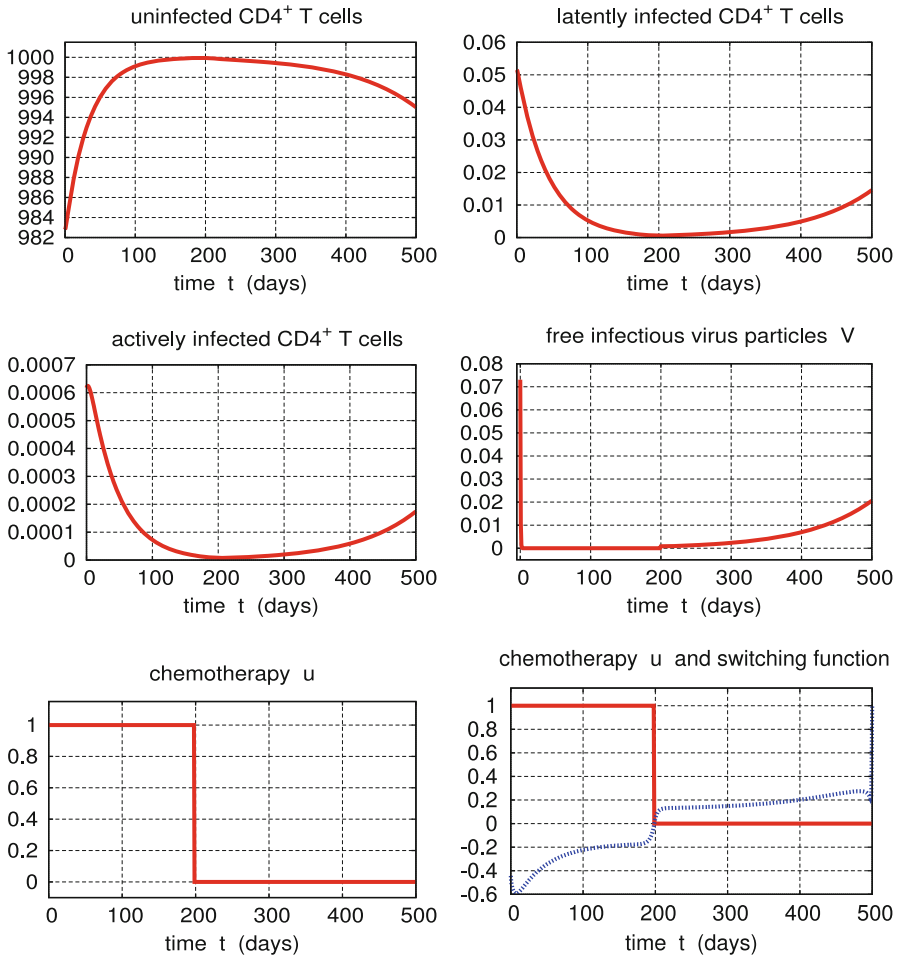
$$\begin{aligned} J_1 &= -489044.529, & t_1 &= 198.566451, & T(t_f) &= 995.0, \\ T^*(t_f) &= 0.014576433, & T^{**}(t_f) &= 0.00017436211, & V(t_f) &= 0.020625442, \\ \lambda_T(0) &= -33.7027, & \lambda_T(t_f) &= -211.377, & \lambda_{T^*}(0) &= 28078.4 \\ \lambda_{T^{**}}(0) &= 2.76312, & \lambda_V(0) &= 405.843. \end{aligned} \tag{79}$$

Figure 5 displays the state and control variables and the switching function. Figure 5 (bottom, right) shows that the strict bang-bang property (77) is satisfied. Condition (c) in Theorem 17 holds because the critical cone  $\mathcal{K}_0 = \{0\}$  contains of zero element. Therefore, the extremal displayed in Fig. 5 provides a strict strong minimum.

Finally, we study the optimal solution for the initial values (69), when the treatment starts after 1000 days and, again, the boundary condition  $T(t_f) = 995$  is prescribed. The arc-parametrization method in [26, 36] and the control package NUDOCCCS yield the results

$$\begin{aligned} J_1 &= -483480.9, & t_1 &= 254.5443, & T(t_f) &= 995.0, \\ T^*(t_f) &= 0.01457694, & T^{**}(t_f) &= 0.0001743682, & V(t_f) &= 0.02062616, \\ \lambda_T(0) &= -35.629, & \lambda_T(t_f) &= -214.021, & \lambda_{T^*}(0) &= 4209.98, \\ \lambda_{T^{**}}(0) &= 2.69187, & \lambda_V(0) &= 136.835. \end{aligned} \tag{80}$$

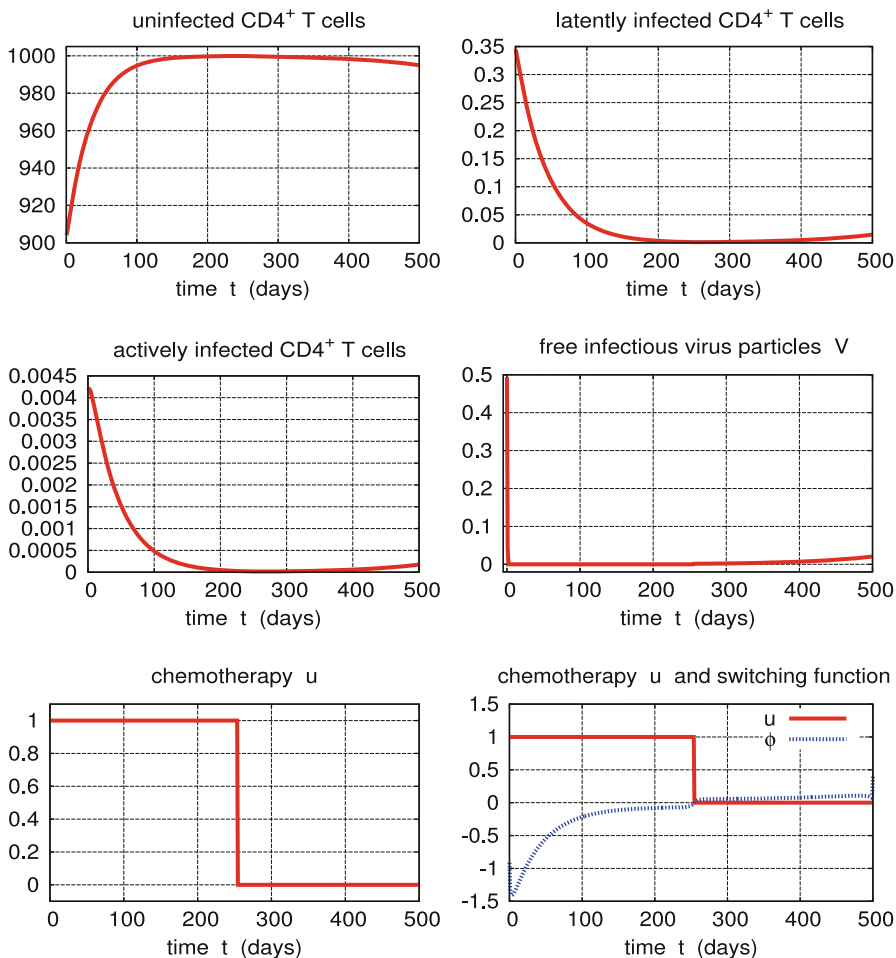
Figure 6 depicts the state and control variables and the switching function. The SSC in Theorem 17 are satisfied, since the *strict bang-bang property* (77) holds and condition (63) holds in view of  $\mathcal{K}_0 = \{0\}$ . Therefore, the extremal (80) provides a strict strong minimum.



**Fig. 5** Optimal solution for initial conditions (68): treatment starts after 800 days and terminal condition  $T(t_f) = 995$ . *Top row: (left) uninfected  $CD4^+T$  cells, (right) latently infected  $CD4^+T^*$ . Middle row: (left) actively infected  $CD4^+T^{**}$  cells, (right) infectious virus particles  $V$ . Bottom row: (left) bang-bang control  $u$ , (right) bang-bang control  $u$  and scaled switching function  $\sigma$  in (73) satisfying the switching condition (74)*

## 8 Numerical Example with Free Final Time: Time-Optimal Control of Two-Link Robots

In this section, we review the results in our book [36] on the optimal control of two-link robots which has been addressed in various articles; cf., e.g. [9, 13, 14, 30]. In these papers, optimal control policies are determined solely on the basis of first

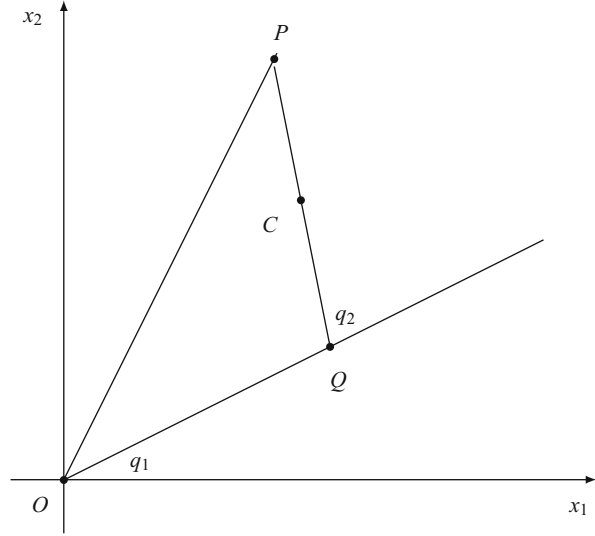


**Fig. 6** Optimal solution for initial conditions (69): treatment starts after 1000 days and terminal condition  $T(t_f) = 995$ . *Top row:* (left) uninfected  $CD4^+T$  cells, (right) latently infected  $CD4^+T^*$ . *Middle row:* (left) actively infected  $CD4^+T^{**}$  cells, (right) infectious virus particles  $V$ . *Bottom row:* (left) bang-bang control  $u$ , (right) bang-bang control  $u$  and scaled switching function  $\sigma$  in (73) satisfying the switching condition (74)

order necessary conditions, since sufficient conditions were not available. In this section we show that SSC hold for both types of robots considered in [9, 14, 30].

First, we study the robot model considered in Chernousko et al. [9]. Göllmann [14] has shown that the optimal control candidate presented in [9] is not optimal, since the sign conditions of the switching functions do not comply with the Minimum Principle. Figure 7 displays the two-link robot schematically. The state variables are the angles  $q_1$  and  $q_2$ . The parameters  $I_1$  and  $I_2$  are the moments of inertia of the upper arm  $\overline{OQ}$  and the lower arm  $\overline{QP}$  with respect to the points  $O$  and

**Fig. 7** Schematical representation of a two-link robot



$Q$ , resp. Further, let  $m_2$  be the mass of the lower arm,  $L_1 = |\overline{OQ}|$  the length of the upper arm, and  $L_2 = |\overline{QC}|$  the distance between the second link  $Q$  and the center of gravity  $C$  of the lower arm. With the abbreviations

$$\begin{aligned} A &= I_1 + m_2 L_1^2 + I_2 + 2m_2 L_1 L_2 \cos q_2, & B &= I_2 + m_2 L_1 L_2 \cos q_2, \\ R_1 &= u_1 + m_2 L_1 L_2 (2\dot{q}_1 + \dot{q}_2) \dot{q}_2 \sin q_2, & R_2 &= u_2 - m_2 L_1 L_2 \dot{q}_1^2 \sin q_2, \\ D &= I_2, & \Delta &= AD - B^2, \end{aligned} \quad (81)$$

the dynamics of the two-link robot can be described by the ODE system

$$\begin{aligned} \dot{q}_1 &= \omega_1, & \dot{\omega}_1 &= \frac{1}{\Delta} (DR_1 - BR_2), \\ \dot{q}_2 &= \omega_2, & \dot{\omega}_2 &= \frac{1}{\Delta} (AR_2 - BR_1), \end{aligned} \quad (82)$$

where  $\omega_1$  and  $\omega_2$  are the angular velocities. The torques  $u_1$  and  $u_2$  in the two links represent the two control variables. Therefore, the state variable and control variable are given by

$$x = (q_1, q_2, \omega_1, \omega_2) \in \mathbb{R}^4, \quad u = (u_1, u_2) \in \mathbb{R}^2.$$

The control problem consists in steering the robot from a given initial position to a terminal position in minimal final time  $t_f$ ,

$$\begin{aligned} q_1(0) &= 0, & q_2(0) &= 0, & \omega_1(0) &= 0, & \omega_2(0) &= 0, \\ q_1(t_f) &= -0.44, & q_2(t_f) &= 1.83, & \omega_1(t_f) &= 0, & \omega_2(t_f) &= 0. \end{aligned} \quad (83)$$



The control components are bounded by

$$|u_1(t)| \leq 2, \quad |u_2(t)| \leq 1, \quad t \in [0, t_f]. \tag{84}$$

The Pontryagin function (Hamiltonian) is

$$H = \lambda_1 \omega_1 + \lambda_2 \omega_2 + \frac{\lambda_3}{\Delta} (DR_1(u_1) - BR_2(u_2)) + \frac{\lambda_4}{\Delta} (AR_2(u_2) - BR_1(u_1)). \tag{85}$$

The adjoint equations are rather complicated and are not given here explicitly. The switching functions are

$$\sigma_1(x, \lambda) = H_{u_1} = \frac{\lambda_3}{\Delta} D - \frac{\lambda_4}{\Delta} B, \quad \sigma_2(x, \lambda) = H_{u_2} = \frac{\lambda_4}{\Delta} A - \frac{\lambda_3}{\Delta} B. \tag{86}$$

For the parameter values

$$L_1 = 1, \quad L = 0.5, \quad m_2 = 10, \quad I_1 = I_2 = \frac{10}{3},$$

Göllmann [14] has found the following control structure with four bang-bang arcs,

$$u(t) = (u_1(t), u_2(t)) = \left\{ \begin{array}{l} (-2, 1), \quad 0 \leq t < t_1 \\ (2, 1), \quad t_1 \leq t < t_2 \\ (2, -1), \quad t_2 \leq t < t_3 \\ (-2, -1), \quad t_3 \leq t \leq t_f \end{array} \right\}, \quad 0 < t_1 < t_2 < t_3 < t_f. \tag{87}$$

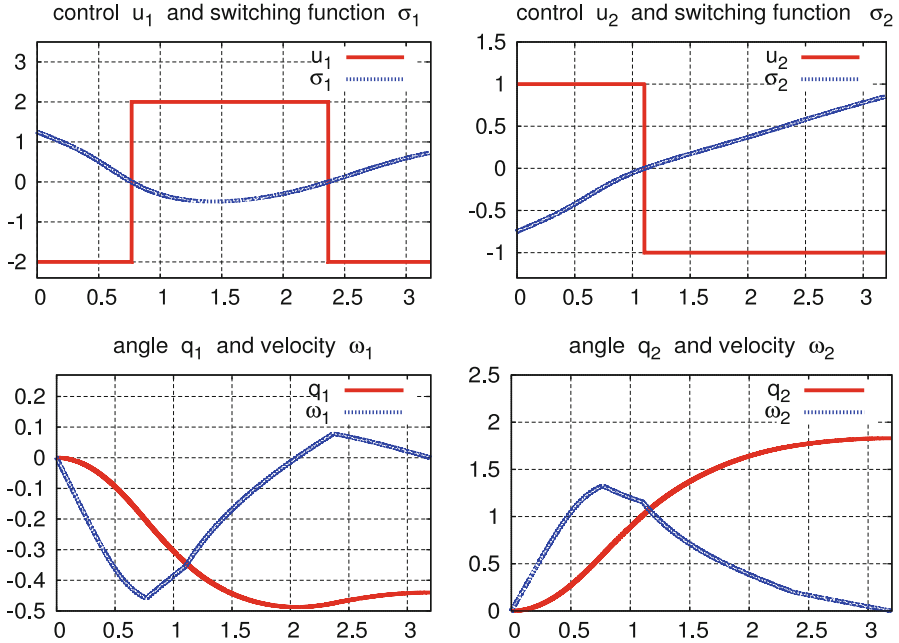
This control structure differs substantially from the one in Chernousko et al. [9] which violates the switching conditions. Obviously, the bang-bang control (87) satisfies the assumption that only one control components switches at a time. Since the initial point  $(q_1(0), q_2(0), \omega_1(0), \omega_2(0))$  is specified, the optimization variable in the IOP (61) is

$$\xi = (\xi_1, \xi_2, \xi_3, \xi_4), \quad \xi_1 = t_1, \quad \xi_2 = t_2 - t_1, \quad \xi_3 = t_3 - t_2, \quad \xi_4 = t_f - t_3.$$

Using the code NUODOCCS we compute the following arc durations and switching times

$$\begin{aligned} t_1 &= 0.7677893, \quad \xi_2 = 0.3358820, \quad t_2 = 1.1036713, \\ \xi_3 &= 1.2626739, \quad t_3 = 2.3663452, \quad \xi_4 = 0.8307667, \\ t_f &= 3.1971119. \end{aligned} \tag{88}$$

Numerical values for the adjoint functions are also provided by the code NUODOCCS, e.g., the initial values



**Fig. 8** Control of the two-link robot (81)–(84). *Top row: (left) control  $u_1$  and scaled switching function  $\sigma_1$ , (right) control  $u_2$  and scaled switching function  $\sigma_2$ . Bottom row: (left) angle  $q_1$  and velocity  $\omega_1$ , (right) angle  $q_2$  and velocity  $\omega_2$*

$$\begin{aligned} \lambda_1(0) &= -1.56972, \quad \lambda_2(0) = -0.917955, \\ \lambda_3(0) &= -2.90537, \quad \lambda_4(0) = -1.45440. \end{aligned} \quad (89)$$

Figure 8 shows that the switching functions  $\sigma_1$  and  $\sigma_2$  comply with the minimum condition and that the strict bang-bang property (35) and the inequalities  $D^k(H) > 0$ ,  $k = 1, 2, 3$  are satisfied:

$$\begin{aligned} \sigma_1(t) &\neq 0 \quad \text{for } t \neq t_1, t_3, \quad \sigma_2(t) \neq 0 \quad \text{for } t \neq t_2, \\ \dot{\sigma}_1(t_1) &< 0, \quad \dot{\sigma}_1(t_3) > 0, \quad \dot{\sigma}_2(t_2) > 0. \end{aligned}$$

For the terminal conditions (83) we obtain the Jacobian

$$\tilde{\mathcal{G}}_{\xi}(\hat{\xi}) = \begin{pmatrix} -0.75104 & 0.035106 & 0.25890 & 0 \\ 3.7612 & 1.8493 & -0.20417 & 0 \\ -0.32635 & 0.077005 & 0.21272 & -0.10782 \\ 1.2685 & 0.44545 & -0.48745 & -0.23363 \end{pmatrix}.$$

This square matrix has full rank in view of

$$\det \tilde{\mathcal{G}}_{\xi}(\hat{\xi}) = 0.076652 \neq 0,$$

which means that the positive definiteness condition (63) trivially holds. Thus we have verified *first-order* sufficient conditions showing that the extremal solution given by (87)–(89) provides a strict strong minimum.

In the model treated above, some parameters like the mass of the upper arm and the mass of a load at the end of the lower arm appear implicitly in the system equations. The mass  $m_1$  of the upper arm is included in the moment of inertia  $I_2$  and the mass  $M$  of a load in the point  $P$  can be added to the mass  $m_2$ , where the point  $C$  and therefore the length  $L$  have to be adjusted. The length  $L_2$  of the lower arm is incorporated in the parameter  $L$ .

The *second robot model* that we are going to discuss is taken from Geering et al. [13] and Oberle [30]. Here, every physical parameter enters the system equation explicitly. The dynamic system is as follows:

$$\begin{aligned} \dot{q}_1 &= \omega_1, & \dot{\omega}_1 &= \frac{1}{\Delta}(AI_{22} - BI_{12} \cos q_2), \\ \dot{q}_2 &= \omega_2 - \omega_1, & \dot{\omega}_2 &= \frac{1}{\Delta}(BI_{11} - AI_{12} \cos q_2), \end{aligned} \tag{90}$$

where we have used the abbreviations

$$\begin{aligned} A &= I_{12}\omega_2^2 \sin q_2 + u_1 - u_2, & B &= -I_{12}\omega_1^2 \sin q_2 + u_2, \\ \Delta &= I_{11}I_{22} - I_{12}^2 \cos^2 q_2, & I_{11} &= I_1 + (m_2 + M)L_1^2, \\ I_{12} &= m_2LL_1 + ML_1L_2, & I_{22} &= I_2 + I_3 + ML_2^2. \end{aligned} \tag{91}$$

Here,  $I_3$  denotes the moment of inertia of the load with respect to the point  $P$  and  $\omega_2$  is now the angular velocity of the angle  $q_1 + q_2$ . For simplicity, we set  $I_3 = 0$ . Again, the torques  $u_1$  and  $u_2$  in the two links are used as control variables by which the robot is steered from a given initial position to a non-fixed end position in minimal final time  $t_f$ ,

$$\begin{aligned} q_1(0) &= 0, & \sqrt{(x_1(t_f) - x_1(0))^2 + (x_2(t_f) - x_2(0))^2} &= r, \\ q_2(0) &= 0, & q_2(t_f) &= 0, \\ \omega_1(0) &= 0, & \omega_1(t_f) &= 0, \\ \omega_2(0) &= 0, & \omega_2(t_f) &= 0, \end{aligned} \tag{92}$$

where  $(x_1(t), x_2(t))$  are the Cartesian coordinates of the point  $P$ ,

$$\begin{aligned} x_1(t) &= L_1 \cos q_1(t) + L_2 \cos(q_1(t) + q_2(t)), \\ x_2(t) &= L_1 \sin q_1(t) + L_2 \sin(q_1(t) + q_2(t)). \end{aligned} \tag{93}$$

The initial point  $(x_1(0), x_2(0)) = (2, 0)$  is fixed. Both control components are bounded,

$$|u_1(t)| \leq 1, \quad |u_2(t)| \leq 1, \quad t \in [0, t_f]. \quad (94)$$

The Hamilton–Pontryagin function is given by

$$H = \lambda_1 \omega_1 + \lambda_2 (\omega_2 - \omega_1) + \frac{\lambda_3}{\Delta} (A(u_1, u_2) I_{22} - B(u_2) I_{12} \cos q_2) \\ + \frac{\lambda_4}{\Delta} (B(u_2) I_{11} - A(u_1, u_2) I_{12} \cos q_2). \quad (95)$$

The switching functions are computed as

$$\sigma_1(x, \lambda) = H_{u_1} = \frac{1}{\Delta} (\lambda_3 I_{22} - \lambda_4 I_{12} \cos q_2), \\ \sigma_2(x, \lambda) = H_{u_2} = \frac{1}{\Delta} (\lambda_3 (-I_{22} - I_{12} \cos q_2) + \lambda_4 (I_{11} + I_{12} \cos q_2)). \quad (96)$$

For the parameter values

$$L_1 = L_2 = 1, \quad L = 0.5, \quad m_1 = m_2 = M = 1, \quad I_1 = I_2 = \frac{1}{3}, \quad I_3 = 0, \quad r = 3,$$

we will show that the optimal control has the following structure with five bang-bang arcs with  $0 = t_0 < t_1 < t_2 < t_3 < t_4 < t_5 = t_f$  (Fig. 9):

$$u(t) = (u_1(t), u_2(t)) = \begin{cases} (-1, 1) & \text{for } 0 \leq t < t_1 \\ (-1, -1) & \text{for } t_1 \leq t < t_2 \\ (1, -1) & \text{for } t_2 \leq t < t_3 \\ (1, 1) & \text{for } t_3 \leq t < t_4 \\ (-1, 1) & \text{for } t_4 \leq t \leq t_f \end{cases}. \quad (97)$$

Since the initial point  $(q_1(0), q_2(0), \omega_1(0), \omega_2(0))$  is specified, the optimization variable in the optimization problem (50), resp., (61) is

$$z = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_5), \quad \xi_k = t_k - t_{k-1}, \quad k = 1, \dots, 5.$$

The code NUDOCSS yields the arc durations and switching times

$$t_1 = 0.546174, \quad \xi_2 = 1.21351, \quad t_2 = 1.75968, \\ \xi_3 = 1.03867, \quad t_3 = 2.79835, \quad \xi_4 = 0.906039, \\ t_4 = 3.70439, \quad \xi_5 = 0.185023, \quad t_f = 3.889409, \quad (98)$$

as well as the initial values of the adjoint variables,

$$\lambda_1(0) = 0.184172, \quad \lambda_2(0) = -0.011125, \\ \lambda_3(0) = 1.482636, \quad \lambda_4(0) = 0.997367. \quad (99)$$

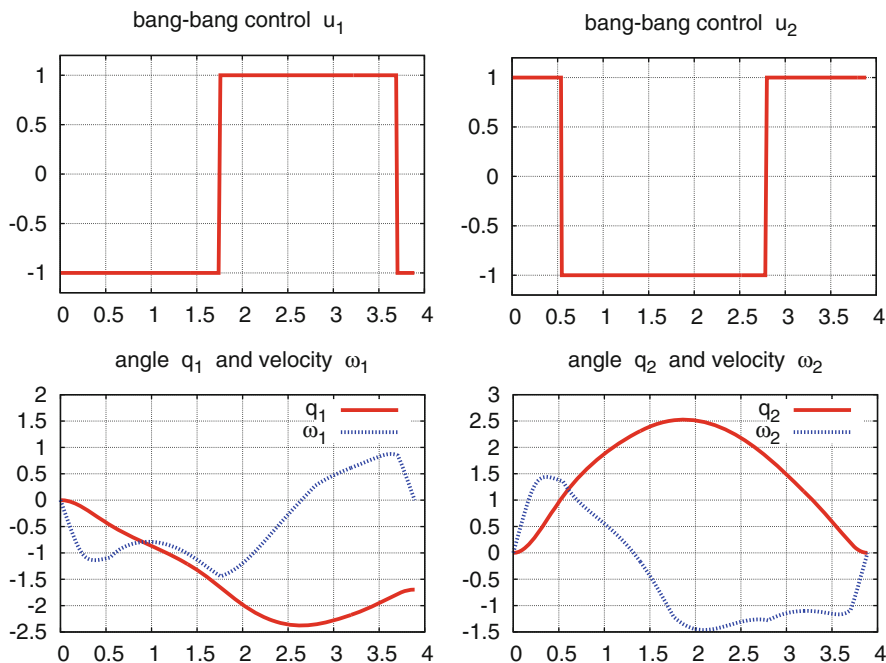
The strict bang-bang property (35) and the inequalities  $D^k(H) > 0, k = 1, 2, 3,$  hold in view of

$$\begin{aligned} \sigma_1(t) &\neq 0 \quad \text{for } t \neq t_2, t_4, & \dot{\sigma}_1(t_2) &< 0, & \dot{\sigma}_1(t_4) &> 0, \\ \sigma_2(t) &\neq 0 \quad \text{for } t \neq t_1, t_3, & \dot{\sigma}_2(t_1) &> 0, & \dot{\sigma}_2(t_3) &< 0. \end{aligned}$$

For the terminal conditions in (92), the Jacobian in the optimization problem is computed as the  $(4 \times 5)$ -matrix

$$\tilde{\mathcal{G}}_{\xi}(\xi) = \begin{pmatrix} -10.858 & -12.746 & -5.8833 & -1.1500 & 0 \\ 0.19928 & -2.7105 & -1.4506 & -1.9148 & -4.83871 \\ -0.62256 & 3.3142 & 2.3155 & 2.9435 & 6.1936 \\ 9.3609 & 3.0393 & 0.48446 & 0.040581 & 0 \end{pmatrix}$$

which has full rank. The Hessian of the Lagrangian is given by



**Fig. 9** Control of the two-link robot (90)–(94). *Top row:* (left) control  $u_1$ , (right) control  $u_2$ . *Bottom row:* (left) angle  $q_1$  and velocity  $\omega_1$ , (right) angle  $q_2$  and velocity  $\omega_2$

$$\widetilde{\mathcal{L}}_{\xi\xi}(\hat{\xi}, \beta) = \begin{pmatrix} 71.142 & 90.761 & 42.130 & 8.4989 & -0.051822 \\ 90.761 & 112.54 & 51.313 & 10.769 & 0.14985 \\ 42.130 & 51.313 & 23.963 & 5.1240 & 0.13860 \\ 8.4989 & 10.769 & 5.1240 & 1.4999 & 0.17078 \\ -0.051822 & 0.14985 & 0.13860 & 0.17078 & 0.29736 \end{pmatrix}.$$

This yields the projected Hessian (64) as the positive number

$$N^* \widetilde{\mathcal{L}}_{\xi\xi}(\hat{\xi}, \beta)N = 0.326929.$$

Hence, all conditions in Theorem 17 are satisfied and thus the extremal (97)–(99) yields a strict strong minimum.

It is interesting to note that there exists a *second local minimum* with the same terminal time  $t_f = 3.88941$ . Though the control has also five bang-bang arcs, the control structure is substantially different from that in (97),

$$u(t) = (u_1(t), u_2(t)) = \begin{cases} (1, -1), & 0 \leq t < t_1 \\ (-1, -1), & t_1 \leq t < t_2 \\ (-1, 1), & t_2 \leq t < t_3 \\ (1, 1), & t_3 \leq t < t_4 \\ (1, -1), & t_4 \leq t \leq t_f \end{cases}, \quad (100)$$

where  $0 < t_1 < t_2 < t_3 < t_4 < t_5 = t_f$ . NUODOCCS determines the switching times

$$\begin{aligned} t_1 &= 0.1850163, \quad t_2 = 1.091075, \quad t_3 = 2.129721, \\ t_4 &= 3.343237, \quad t_f = 3.889409, \end{aligned} \quad (101)$$

for which the strict bang-bang property (35) holds and  $D^k(H) > 0$  for  $k = 1, 2, 3, 4$ . Moreover, computations show that  $\text{rank}(\widetilde{\mathcal{G}}_{\xi}(\hat{\xi})) = 4$  and that the projected Hessian of the Lagrangian (64) is the positive number

$$N^* \widetilde{\mathcal{L}}_{\xi\xi}(\hat{\xi}, \beta)N = 0.326929.$$

It is remarkable that this value is identical with the value of the projected Hessian for the first local minimum. Therefore, also for the second solution we have verified that all conditions in Theorem 17 hold, and thus the extremal (100), (101) is a strict strong minimum. The phenomenon of multiple local solutions all with the same minimal time  $t_f$  has also been observed by Betts [3], Example 6.8 (Reorientation of a rigid body).

## 9 Optimal Control Problems with Mixed Control-State Constraints and Control Appearing Linearly

To the best of our knowledge, second-order sufficient optimality conditions (SSC) for optimal control problems with mixed control-state constraints have only been studied for the class of *regular controls*, where the *strict Legendre–Clebsch* condition holds. Such control problems have not yet been considered, when the control variable appears linearly in the system dynamics and in the mixed control-state constraint. For a two-sided control-state constraint we will show that the constraining function itself can be taken as a new control variable, whereby the original control problem is transformed into a classical control problem with an affine control variable subject to simple control bounds. Hence, optimal controls for the transformed control problem are concatenations of bang-bang and singular arcs. The material in this section is based on our paper [23].

### 9.1 Statement of the Problem and Transformed Control Problem

For simplicity, we consider an optimal control problem with fixed initial time  $t_0 = 0$ , fixed initial conditions and terminal equality constraints, and with a scalar control. Let  $x \in \mathbb{R}^n$  denote the state variable and  $u \in \mathbb{R}$  be the control variable. The terminal time  $t_f > 0$  is either fixed or free. The dynamic equation and boundary conditions are

$$\dot{x} = f(t, x) + g(t, x)u, \quad x(0) = x_0, \quad K(x(t_f)) = 0. \tag{102}$$

We consider a two-sided mixed control-state constraint which is affine in the control variable:

$$\alpha \leq a(x(t)) + b(x(t))u(t) \leq \beta \quad \text{for a.e. } t \in [0, t_f]. \tag{103}$$

The optimal control problem consists in finding a control  $u \in L^\infty([0, t_f], \mathbb{R})$  that *minimizes* the objective functional in Mayer form

$$\mathcal{J}(x, u) = J(x(t_f)). \tag{104}$$

The functions  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $a, b : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $J : \mathbb{R}^n \rightarrow \mathbb{R}$  and  $K : \mathbb{R}^n \rightarrow \mathbb{R}^{d(K)}$  ( $0 \leq d(K) \leq n$ ) are assumed to be twice continuously differentiable. We remind the reader that a Bolza functional of the form

$$\mathcal{J}(x, u) = J(x(t_f)) + \int_0^{t_f} (f_0(t, x) + g_0(t, x)u) dt \tag{105}$$

can be reduced to Mayer form by introducing the additional state variable  $y$  that solves the initial value problem  $\dot{y} = f_0(t, x) + g_0(t, x)u$ ,  $y(0) = 0$  and minimizing the functional  $J(x(t_f)) + y(t_f)$ .

The following *regularity assumption* will be assumed to hold for feasible trajectories:

$$b(t, x(t)) \neq 0 \quad \text{for } t \in [0, t_f]. \quad (106)$$

This assumption allows us to introduce a *new control variable*  $v$  that is related to the control variable  $u$  as follows:

$$v := a(x) + b(x)u, \quad \text{i.e.,} \quad u = (v - a(x))/b(x). \quad (107)$$

The *transformed optimal control problem* consists in minimizing the objective (104) subject to the transformed dynamics

$$\dot{x} = \bar{f}(t, x) + \bar{g}(t, x)v, \quad x(0) = x_0, \quad K(x(t_f)) = 0, \quad (108)$$

where the transformed functions  $\bar{f}, \bar{g}$  are defined by

$$\bar{f}(t, x) = f(t, x) - g(t, x)a(x)/b(x), \quad \bar{g}(t, x) = g(t, x)/b(x). \quad (109)$$

The mixed control-state constraint (103) then is equivalent to the simple control constraint

$$\alpha \leq v(t) \leq \beta. \quad (110)$$

Thus, we can apply the second-order conditions developed in Sects. 5 and 6 to the transformed problem.

## 9.2 Numerical Example: Optimal Control of the Rayleigh Equation

The Rayleigh equation describes oscillations of the electric current, resp., voltage in an electric circuit. The optimal control of the Rayleigh equation for a control-quadratic objective has been studied in Maurer and Augustin [20], Osmolovskii and Maurer [36], and Chen and Gerdtts [8], where both simple control bounds and a mixed control-state constraint were investigated.

Let  $x_1$  denote the electric current and  $x_2$  the voltage. The control  $u$  represents the voltage at the generator which steers the following dynamic equations:

$$\begin{aligned} \dot{x}_1 &= x_2, & x_1(0) &= -5, \\ \dot{x}_2 &= -x_1 + x_2(1.4 - 0.14x_2^2) + u, & x_2(0) &= -5. \end{aligned} \quad (111)$$



We consider the mixed control-state constraint

$$\alpha \leq u + x_1 \leq \beta \quad (\alpha = -5, \beta = 0). \tag{112}$$

Various other bounds  $\alpha$  and  $\beta$  have been studied in Maurer and Omolovskii [23]. The mixed constraint is a slight modification of the one considered in [8, 20]. The objective is to minimize the quadratic functional

$$\mathcal{J}(x, u) = \int_0^{t_f} (x_1(t)^2 + x_2(t)^2) dt. \tag{113}$$

First, we consider this control problem with fixed terminal time  $t_f = 4.5$ . Later, we shall prescribe the terminal condition

$$x_1(t_f) = 0 \tag{114}$$

and solve the control problem with free terminal time  $t_f$ .

According to (107), the new control variable  $v$  is given by

$$v = u + x_1, \quad u = v - x_1. \tag{115}$$

The Pontryagin function (Hamiltonian) with respect to the control  $v$  becomes

$$H(x, \lambda, v) = x_1^2 + x_2^2 + \lambda_1 x_2 + \lambda_2 (-x_1 + x_2(1.4 - 0.14x_2^2) + v - x_1). \tag{116}$$

The adjoint equations are

$$\begin{aligned} \dot{\lambda}_1 &= -H_{x_1} = -2x_1 + 2\lambda_2, \\ \dot{\lambda}_2 &= -H_{x_2} = -2x_2 - \lambda_1 + \lambda_2(0.42x_2^2 - 1.4). \end{aligned} \tag{117}$$

For the first control problem with free endpoint  $x(t_f)$  and fixed final time  $t_f = 4.5$ , we get the transversality condition

$$\lambda_1(t_f) = 0, \quad \lambda_2(t_f) = 0, \tag{118}$$

while the second control problem with terminal constraint  $x_1(t_f) = 0$  and free terminal time  $t_f$  gives the transversality conditions

$$H(t_f) = 0, \quad \lambda_2(t_f) = 0. \tag{119}$$

The switching function  $\sigma = H_v = \lambda_2$  determines the optimal control according to

$$v(t) = \begin{cases} \beta & , \text{ if } \lambda_2(t) < 0, \\ \alpha & , \text{ if } \lambda_2(t) > 0, \\ \text{singular} & , \text{ if } \lambda_2(t) = 0 \quad \forall t \in I_s \subset [0, t_f]. \end{cases} \tag{120}$$

We do not discuss singular controls further, because for the chosen bounds  $\alpha = -5$  and  $\beta = 0$  in the mixed constraint (112) we obtain bang-bang controls. Singular controls for smaller values of  $\alpha$  are discussed in [23].

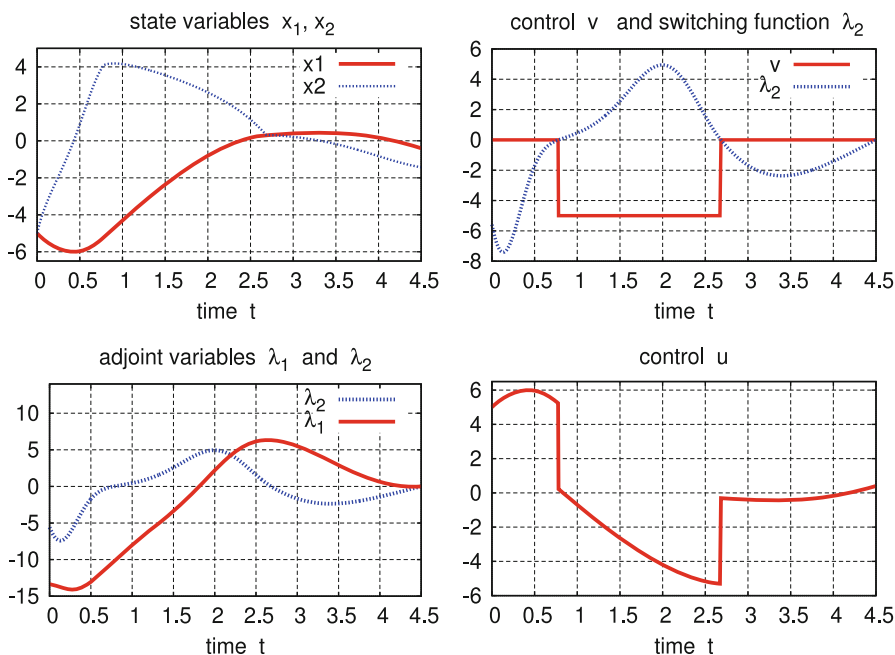
In the first control problem with free terminal state  $x(t_f)$  and fixed terminal time  $t_f$ , we obtain a bang-bang control with three bang-bang arcs:

$$v(t) = \left\{ \begin{array}{ll} 0 & , \text{ if } 0 \leq t \leq t_1 \\ -5 & , \text{ if } t_1 < t \leq t_2 \\ 0 & , \text{ if } t_2 < t \leq t_f \end{array} \right\}. \tag{121}$$

The code NUDOCSS [4] yields the following results:

$$\begin{aligned} \mathcal{J}(x, u) &= 62.165171, & t_1 &= 0.77996717, & t_2 &= 2.6835574, \\ x_1(t_f) &= -0.40342897, & x_2(t_f) &= -1.4332277, \\ \lambda_1(0) &= -13.364385, & \lambda_2(0) &= -5.591549. \end{aligned}$$

The corresponding extremal solution is shown in Fig. 10.



**Fig. 10** Objective (113) and constraint  $-5 \leq v = u + x_1 \leq 0$ . Top row: (left) state variables  $x_1, x_2$ , (right) transformed control  $v$  and switching function  $\sigma = \lambda_2$ . Bottom row: (left) adjoint variables  $\lambda_1, \lambda_2$ , (right) control  $u$

Next, we compute the Hessian of the Lagrangian for the IOP:

$$\mathcal{L}_{zz} = \begin{pmatrix} 573.237 & 458.854 \\ 458.854 & 377.399 \end{pmatrix}.$$

Obviously, this matrix is positive-definite. Moreover, Fig. 10 (top row, right) shows that the switching function  $\sigma(t) = \lambda_2(t)$  satisfies the strict bang-bang property (37); cf. also Remark 5.1:

$$\begin{aligned} \lambda_2(t) < 0 & \quad \forall 0 \leq t < t_1, & \lambda_2(t_1) = 0, & \dot{\lambda}_2(t_1) > 0, \\ \lambda_2(t) > 0 & \quad \forall t_1 < t < t_2, & \lambda_2(t_2) = 0, & \dot{\lambda}_2(t_2) < 0, \\ \lambda_2(t) < 0 & \quad \forall t_2 < t < t_f, & \lambda_2(t_f) = 0, & \dot{\lambda}_2(t_f) > 0. \end{aligned}$$

Hence, the extremal shown in Fig. 10 satisfies the SSC in Theorem 17 and thus is a strict strong minimum.

Now we study the solution, when the terminal condition  $x_1(t_f) = 0$  is imposed and the terminal time  $t_f$  is free. In this case we obtain a bang-bang control with only one switch:

$$v(t) = \begin{cases} 0 & , \text{ if } 0 \leq t \leq t_1 \\ -5 & , \text{ if } t_1 < t \leq t_f \end{cases}. \tag{122}$$

The corresponding IOP has the two optimization variables  $t_1, t_f$  and the scalar equality constraint  $x_1(t_f) = 0$ . The code NUDOCCCS yields the following results:

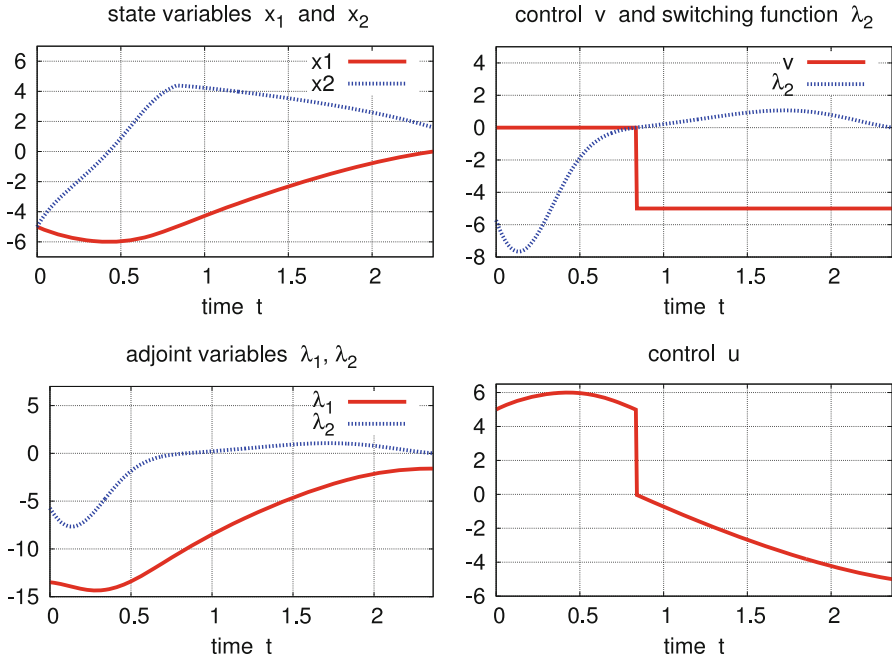
$$\begin{aligned} \mathcal{J}(x, u) &= 60.72697, & t_1 &= 0.8343100, & t_f &= 2.364688, \\ x_1(t_f) &= 0.0, & x_2(t_f) &= 1.600202, \\ \lambda_1(0) &= -13.4868, & \lambda_1(t_f) &= -1.60020, \\ \lambda_2(0) &= -5.72868, & \lambda_2(t_f) &= 0.0. \end{aligned}$$

The extremal solution with state, control, adjoint variables, and switching function is shown in Fig. 11.

The reduced Hessian (64) of the Lagrangian for the IOP is a scalar which we compute as the positive number 4.3525. Moreover, Fig. 11, top row, right, shows that the strict bang-bang property is fulfilled; cf. also Remark 5.1:

$$\begin{aligned} \lambda_2(t) < 0 & \quad \forall 0 \leq t < t_1, & \lambda_2(t_1) = 0, & \dot{\lambda}_2(t_1) > 0, \\ \lambda_2(t) > 0 & \quad \forall t_1 < t < t_f, & \lambda_2(t_f) = 0, & \dot{\lambda}_2(t_f) < 0. \end{aligned}$$

Hence, the extremal shown in Fig. 11 satisfies the SSC in Theorem 17 and thus provides a strict strong minimum.



**Fig. 11** Objective (113) with free terminal time  $t_f$  and constraints  $-5 \leq v = u + x_1 \leq 0, x_1(t_f) = 0$ . Top row: (left) state variables  $x_1, x_2$ , (right) transformed control  $v$  and switching function  $\sigma = \lambda_2$ . Bottom row: (left) adjoint variables  $\lambda_1, \lambda_2$ , (right) control  $u = v - x_1$

## 10 Conclusion

We presented no-gap necessary and sufficient second-order optimality conditions for extremals with discontinuous controls in the simplest problem of the Calculus of Variations and the general optimal control problem with regular mixed constraint  $g(t, x, u) = 0$  on a variable time interval  $[t_0, t_f]$ . We formulated similar conditions for bang-bang controls in an optimal control problem with a Mayer functional, where the dynamical system is affine in control variable and the control constraint is given by a convex polyhedron. Bang-bang controls induce an optimization problem with respect to the switching times of the control, the so-called Induced Optimization Problem IOP. We showed that the classical second-order sufficient condition for the IOP together with the strict bang-bang property of the switching function ensure second-order sufficient conditions (IOP) for the bang-bang control problem. The verification of SSC for bang-bang controls was illustrated on two numerical examples. First, we studied extremals in the optimal control of the chemotherapy of HIV. Then, following [36], we investigated extremals in time-optimal control problems of two-link robots. We also discussed optimal control problems with running mixed control-state constraints and control appearing linearly. Taking the

mixed constraint as a new control variable we converted such problems to bang-bang control problems. As an example, we studied extremals in the optimal control problem for the Rayleigh equation.

The results on SSC naturally lend themselves to sensitivity results for the IOP and the underlying bang-bang control problem using the well-known sensitivity results for finite-dimensional optimization problems developed by Fiacco [11]. Sensitivity results for bang-bang controls may be found in Felgenhauer [10], Kim and Maurer [17] and Maurer and Vossen [25]. Sensitivity results also allow to develop real-time control techniques as indicated already in Büskens et al. [7]. These issues will be the topic of a future paper. Related sensitivity results may be obtained for bang-singular controls as suggested in Vossen [40, 41]. This approach still needs a practical method of verifying the more abstract SSC for bang-singular controls given in Aronna et al. [2].

## References

1. Agrachev, A.A., Stefani, G., Zezza, P.L.: Strong optimality for a bang-bang trajectory. *SIAM J. Control Optim.* **41**, 991–1014 (2002)
2. Aronna, M.S., Bonnans, F., Dmitruk, A., Lotito, P.A.: Quadratic order conditions for bang-singular extremals. *Numer. Algebra Control Optim.* **2**, 511–546 (2012)
3. Betts, J.T.: *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*. Advances in Design and Control, vol. DC 19, 2nd edn. SIAM Publications, Philadelphia (2010)
4. Büskens, C.: *Optimierungsmethoden und Sensitivitätsanalyse für optimale Steuerprozesse mit Steuer- und Zustands-Beschränkungen*. Dissertation, Institut für Numerische Mathematik, Universität Münster (1998)
5. Büskens, C., Maurer, H.: SQP-methods for solving optimal control problems with control and state constraints: adjoint variables, sensitivity analysis and real-time control. *J. Comput. Appl. Math.* **120**, 85–108 (2000)
6. Büskens, C., Maurer, H.: Sensitivity analysis and real-time optimization of parametric nonlinear programming problems. In: Grötschel, M., Krumke, S.O., Rambau, J. (eds.) *Online Optimization of Large Scale Systems*, pp. 3–16. Springer, Berlin (2001)
7. Büskens, C., Pesch, H.J., Winderl, S.: Real-time solutions of bang-bang and singular optimal control problems. In: Grötschel, M., Krumke, S.O., Rambau, J. (eds.) *Online Optimization of Large Scale Systems*, pp. 129–142. Springer, Berlin (2001)
8. Chen, J., Gerds, M.: Smoothing techniques of nonsmooth Newton methods for control-state constrained optimal control problems. *SIAM J. Numer. Anal.* **50**, 1982–2011 (2012)
9. Chernousko, F.L., Akulenko, L.D., Bolotnik, N.N.: Time-optimal control for robotic manipulators. *Opt. Control Appl. Methods* **10**, 293–311 (1989)
10. Felgenhauer, U.: On stability of bang-bang type controls. *SIAM J. Control Optim.* **41**, 1843–1867 (2003)
11. Fiacco, A.V.: *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. Mathematics in Science and Engineering, vol. 165. Academic, New York (1983)
12. Fourer, R., Gay, D.M., Kernighan, B.W.: *AMPL: A Modeling Language for Mathematical Programming*. Duxbury Press/Brooks-Cole Publishing Company, Monterey (1993)
13. Geering, H.P., Guzella, L., Hepner, S.A.R., Onder, C.: Time-optimal motions of robots in assembly tasks. *IEEE Trans. Autom. Control* **AC-31**, 512–518 (1986)

14. Göllmann, L.: Numerische Berechnung zeitoptimaler Trajektorien für zweigliedrige Roboterarme. Diploma thesis, Institut für Numerische Mathematik, Universität Münster (1991)
15. Hannemann, R.: Diploma thesis, Institut für Numerische und Angewandte Mathematik, Universität Münster (2004)
16. Kaya, C.Y., Noakes, J.L.: Computational method for time-optimal switching control. *J. Optim. Theory Appl.* **117**, 69–92 (2003)
17. Kim, J.-H.R., Maurer, H.: Sensitivity analysis of optimal control problems with bang-bang controls. In: Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, pp. 3281–3286 (2003)
18. Kirschner, D., Lenhart, S., Serbin, S.: Optimal control of the chemotherapy of HIV. *J. Math. Biol.* **35**, 775–792 (1996)
19. Malanowski, K., Maurer, H.: Sensitivity analysis for parametric control problems with control-state constraints. *Comput. Optim. Appl.* **5**, 253–283 (1996)
20. Maurer, H., Augustin, D.: Sensitivity analysis and real-time control of parametric optimal control problems using boundary value methods. In: Grötschel, M., Krumke, S.O., Rambau, J. (eds.) *Online Optimization of Large Scale Systems*, pp. 17–55. Springer, Berlin (2001)
21. Maurer, H., Osmolovskii, N.P.: Quadratic sufficient optimality conditions for bang-bang control problems. *Control Cybern.* **33**, 555–584 (2003)
22. Maurer, H., Osmolovskii, N.P.: Second order sufficient conditions for time-optimal bang-bang control problems. *SIAM J. Control Optim.* **42**, 2239–2263 (2004)
23. Maurer, H., Osmolovskii, N.: Second-order conditions for optimal control problems with mixed control-state constraints and control appearing linearly. In: Proceedings of the 52nd IEEE Conference on Control and Design, Firenze, 9–12 Dec, pp. 514–519 (2013)
24. Maurer, H., Pickenhain, S.: Second order sufficient conditions for optimal control problems with mixed control-state constraints. *J. Optim. Theory Appl.* **86**, 649–667 (1995)
25. Maurer, H., Vossen, G.: Sufficient conditions and sensitivity analysis for bang-bang control problems with state constraints. In: Korytowski, A., Szymkat, M. (eds.) *Proceedings of the 23rd IFIP Conference on System Modeling and Optimization*, Cracow, pp. 82–99. Springer, Berlin (2009)
26. Maurer, H., Büskens, C., Kim, J.-H.R., Kaya, Y.: Optimization methods for the verification of second-order sufficient conditions for bang-bang controls. *Opt. Control Methods Appl.* **26**, 129–156 (2005)
27. Maurer, H., Tarnopolskaya, T., Fulton, N.: Computation of optimal bang-bang and singular controls in planar collision avoidance. *J. Ind. Manag. Optim.* **10**(2), 443–460 (2014) [Special Issue on Computational Methods for Optimization and Control]
28. Milyutin, A.A., Osmolovskii, N.P.: *Calculus of Variations and Optimal Control*. Translations of Mathematical Monographs, vol. 180, American Mathematical Society, Providence, RI (1998)
29. Noble, J., Schättler, H.: Sufficient conditions for relative minima of broken extremals. *J. Math. Anal. Appl.* **269**, 98–128 (2002)
30. Oberle, H.J.: Numerical computation of singular control functions for a two-link robot arm. *Opt. Control Lect. Notes Control Inf. Sci.* **95**, 244–253 (1987)
31. Osmolovskii, N.P.: Quadratic optimality conditions for broken extremals in the general problem of calculus of variations. *J. Math. Sci.* **123**(3), 3987–4122 (2004)
32. Osmolovskii, N.P.: Second order conditions in optimal control problem with mixed equality-type constraints on a variable time interval. *Control Cybern.* **38**(4) 1535–1556 (2009)
33. Osmolovskii, N.P.: Sufficient quadratic conditions of extremum for discontinuous controls in optimal control problem with mixed constraints. *J. Math. Sci.* **173**, 1–106 (2011)
34. Osmolovskii, N.P.: Necessary quadratic conditions of extremum for discontinuous controls in optimal control problem with mixed constraints. *J. Math. Sci.* **183**, 435–576 (2012)
35. Osmolovskii, N.P., Maurer, H.: Equivalence of second order optimality conditions for bang-bang control problems. Part 2: Proof, variational derivatives and representations. *Control and Cybern.* **36**, 5–45 (2007); Part 1: Main result. *Control Cybern.* **34**, 927–950 (2005)

36. Osmolovskii, N.P., Maurer, H.: Applications to Regular and Bang-Bang Control: Second-Order Necessary and Sufficient Optimality Conditions in Calculus of Variations and Optimal Control. SIAM Series Design and Control, vol. DC 24. SIAM Publications, Philadelphia (2012)
37. Perelson, A., Kirschner, D., DeBoer, R.: The dynamics of HIV infection of CD4<sup>+</sup> T cells. *Math. Biosci.* **114**, 81–125 (1993)
38. Schättler, H., Ledzewicz, U.: Geometric Optimal Control: Theory, Methods and Examples. Interdisciplinary Applied Mathematics, vol. 38. Springer, New York (2012)
39. Schättler, H., Ledzewicz, U., Maurer, H.: Sufficient conditions for strong local optimality in optimal control problems with  $L_2$ -type objectives and control constraints. *Discrete Contin. Dyn. Syst. B* **19**(8), 2657–2679 (2014)
40. Vossen, G.: Numerische Lösungsmethoden, hinreichende Optimalitätsbedingungen und Sensitivitätsanalyse für optimale bang-bang und singuläre Steuerungen. Dissertation, Institut für Numerische und Angewandte Mathematik, Universität Münster (2005)
41. Vossen, G.: Switching time optimization for bang-bang and singular controls. *J. Optim. Theory Appl.* **144**(2), 409–429 (2010)
42. Wächter, A., Biegler, L.T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **106**, 25–57 (2006); cf. IPOPT home page (C. Laird and A. Wächter): <https://projects.coin-or.org/Ipopt>