Dedicated to
Professor Czesław Olech

Control and Cybernetics

vol. **34** (2005) No. 3

Equivalence of second order optimality conditions for bang–bang control problems. Part 1: Main results*

by

N.P. Osmolovskii[1] and H. Maurer[2]

[1]Systems Research Institute, Polish Academy of Sciences
ul. Newelska 6, 01-447 Warszawa, Poland
and
Institute of Mathematics and Physics, The Academy of Podlasie
3 Maja 54, 08-110 Siedlce, Poland
e-mail: nikolai@osmolovskii.msk.ru

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

Abstract: Second order optimality conditions have been derived in the literature in two different forms. Osmolovskii (1988a, 1995, 2000, 2004) obtained second order necessary and sufficient conditions requiring that a certain quadratic form be positive (semi)-definite on a critical cone. Agrachev, Stefani, Zezza (2002) first reduced the bang-bang control problem to finite-dimensional optimization and then show that 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. In this paper, we establish the equivalence of both forms of sufficient conditions and give explicit relations between corresponding Lagrange multipliers and elements of critical cones. Part 1 summarizes the main results while detailed proofs will be given in Part 2.

Keywords: bang–bang control, second order necessary and suffi-cient conditions, critical cone, quadratic forms and equivalence.

*This work was partly supported by NSh-304.2003.1 and RFBR 04-01-00482

1. Introduction

We study bang–bang control problems in a very general form admitting free final time and mixed initial and terminal conditions of equality and inequality type. Second order necessary and sufficient optimality conditions for bang-bang controls were obtained by Osmolovskii; see Milyutin and Osmolovskii (1998), Part 2, Chapter 3, Section 12. These conditions require the postive (semi)-definiteness of a certain quadratic form on the finite-dimensional critical cone. Using a suitable transformation via a linear matrix ODE, we have developed numerical methods of testing the positive definiteness of the quadratic form; see Maurer, Osmolovskii (2003,2004).

A different approach to second order sufficient conditions (SSC) was presented in Agrachev, Stefani, Zezza (2002). These authors consider a more specialized form of a bang-bang control problem with fixed final time and separated boundary conditions of equality type. Their approach consists in transforming the bang-bang control problem into a finite-dimensional optimization problem where the assumed finitely many switching times and the (possibly free) initial point are taken as optimization variables. (We will refer to this problem as to the *induced optimization problem*). Their main result can be summarized by stating that a combination of finite dimensional SSC for the induced problem and the so-called strict bang-bang property imply SSC for the bang-bang control problem. An extension of this result to control problems with free final time was discussed in Poggioloni, Stefani (2003). However, no numerical applications have been given by these authors so far, though they present a conceptual algorithm. Recently, a numerical implementation of the optimization approach has been discussed in Maurer, Büskens, Kim and Kaya (2004) on a variety of bang-bang control problems. The basic ideas for applying SSC to the sensitivity analysis of bang-bang controls may be found in Kim, Maurer (2003).

When comparing the approaches of Agrachev, Stefani, Zezza (2002) and Osmolovskii (Milyutin and Osmolovskii, 1998) we strongly suspected that the results in both publications are mathematically equivalent. Indeed, it is the purpose of this paper to establish the equivalence of quadratic forms in both works and to give explicit relations between the corresponding Lagrange multipliers and elements of the critical cones. An interesting side-effect of this analysis is that elements of the Hessian of the Lagrangian associated to the optimization problem can be computed solely on the basis of first order variations of the trajectory. The results obtained in this paper extend those in Agrachev, Stefani, Zezza (2002) to the general class of bang-bang control problems, more precisely, to the problem on a nonfixed time interval, with mixed initial and terminal conditions of equality and inequality type, see the main problem (1)–(4).

Due to space restrictions we present only the basic methodology and a summary of the results. The proofs are given in the second part of the paper. They involve a detailed and rather lengthy study of the first and second order derivatives of the trajectories with respect to variations of the switching times, the free final time and the free initial point.

In Section 2 we give a statement of the general bang-bang control problem (main problem), formulations of the minimum principle (the first order necessary optimality condition) and the notion of a regular (or strict) bang-bang control. In Section 3 we formulate second order optimality conditions, both necessary and sufficient, for a regular bang-bang control in the main problem which are given in Milyutin and Osmolovskii (1998) and briefly discuss the proofs of these conditions. Section 4 contains the main results of this paper: a statement of the finite-dimensional induced optimization problem introduced by Agrachev, Stefani and Zezza, second order necessary and sufficient optimality conditions for the induced problem, relationships between Lagrange multipliers, critical cones and quadratic forms in the main and induced problems.

2. Bang–bang control problems on nonfixed time intervals

2.1. The main problem

We consider optimal control problems with control appearing linearly. Let $x(t) \in \mathbb{R}^{d(x)}$ denote the state variable and $u(t) \in \mathbb{R}^{d(u)}$ the control variable in the time interval $t \in [t_0, t_1]$ with non–fixed initial time t_0 and final time t_1. We shall refer to the following optimal control problem (1)–(4) as the *main problem*:

$$\text{Minimize} \qquad \mathcal{J}(t_0, t_1, x, u) = J(t_0, x(t_0), t_1, x(t_1)) \tag{1}$$

subject to the constraints

$$\dot{x}(t) = f(t, x(t), u(t)), \quad u(t) \in U, \quad (t, x(t)) \in \mathcal{Q}, \quad t_0 \le t \le t_1, \tag{2}$$

$$\begin{gathered} F(t_0, x(t_0), t_1, x(t_1)) \le 0, \quad K(t_0, x(t_0), t_1, x(t_1)) = 0, \\ (t_0, x(t_0), t_1, x(t_1)) \in \mathcal{P}, \end{gathered} \tag{3}$$

where the control variable appears linearly in the system dynamics,

$$f(t, x, u) = a(t, x) + B(t, x)u\,. \tag{4}$$

Here, F, K, a are column-vector functions, B is a $d(x) \times d(u)$ matrix function, $\mathcal{P} \subset \mathbb{R}^{2+2d(x)}$, $\mathcal{Q} \subset \mathbb{R}^{1+d(x)}$ are open sets and $U \subset \mathbb{R}^{d(u)}$ is a convex polyhedron. The functions J, F, K are assumed to be twice continuously differentiable on $\mathcal{P}$ and the functions a, B are twice continuously differentiable on $\mathcal{Q}$. The dimensions of F, K are denoted by $d(F)$, $d(K)$. By $\Delta = [t_0, t_1]$ we shall denote the interval of control.

We shall use the abbreviations

$$x_0 = x(t_0),\ x_1 = x(t_1),\ p = (t_0, x_0, t_1, x_1).$$

A trajectory

$$\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_1])$$

is said to be *admissible*, if $x(\cdot)$ is absolutely continuous, $u(\cdot)$ is measurable bounded on Δ and the pair of functions $(x(t), u(t))$ together with the end-points $p = (t_0, x(t_0), t_1, x(t_1))$ satisfies the constraints (2), (3).

Definition 2.1 *The trajectory $\mathcal{T}$ affords a Pontryagin local minimum, if there is no sequence of admissible trajectories*

$$\mathcal{T}^n = (x^n(t), u^n(t) \mid t \in [t_0^n, t_1^n]), \quad n = 1, 2, \ldots,$$

such that the following properties hold with $\Delta^n = [t_0^n, t_1^n]$:

(a) $\mathcal{J}(\mathcal{T}^n) < \mathcal{J}(\mathcal{T}) \quad \forall\, n$ *and* $t_0^n \to t_0,\ t_1^n \to t_1$ *for* $n \to \infty$;

(b) $\max\limits_{\Delta^n \cap \Delta} |x^n(t) - x(t)| \to 0$ *for* $n \to \infty$;

(c) $\int\limits_{\Delta^n \cap \Delta} |u^n(t) - u(t)|\, dt \to 0$ *for* $n \to \infty$.

Note that for a fixed time interval Δ, a Pontryagin minimum corresponds to an L_1–local minimum with respect to the control variable.

2.2. First order necessary optimality conditions

Let

$$\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$$

be a fixed admissible pair of functions such that the control $\hat{u}(\cdot)$ is a piecewise constant function on the interval $\hat{\Delta} = [\hat{t}_0, \hat{t}_1]$. Denote by

$$\hat{\theta} = \{\hat{\tau}_1, \ldots, \hat{\tau}_s\}, \quad \hat{t}_0 < \hat{\tau}_1 < \ldots < \hat{\tau}_s < \hat{t}_1$$

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

$$[\hat{u}]^k = \hat{u}^{k+} - \hat{u}^{k-}$$

to denote the jump of function $\hat{u}(t)$ at the point $\hat{\tau}_k \in \hat{\theta}$, where

$$\hat{u}^{k-} = \hat{u}(\hat{\tau}_k - 0), \quad \hat{u}^{k+} = \hat{u}(\hat{\tau}_k + 0)$$

are the left hand and the right hand side values of the control $\hat{u}(t)$ at $\hat{\tau}_k$, respectively. Similarly, we denote by $[\dot{\hat{x}}]^k$ the jump of the function $\dot{\hat{x}}(t)$ at the same point.

Let us formulate first-order necessary conditions of optimality for $\hat{\mathcal{T}}$ in the form of the Pontryagin minimum principle. To this end we introduce the Pontryagin or Hamiltonian function

$$H(t, x, \psi, u) = \psi f(t, x, u) = \psi a(t, x) + \psi B(t, x) u, \tag{5}$$

where ψ is a row-vector of dimension $d(\psi) = d(x)$ while x, u, f and a are column-vectors. The factor of the control u in the Pontryagin function is called the *switching function*

$$\sigma(t, x, \psi) = \psi B(t, x) \tag{6}$$

which is a row vector of dimension $d(u)$. Denote by l the end-point Lagrange function

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

where α and β are row-vectors with $d(\alpha) = d(F)$, $d(\beta) = d(K)$, and α_0 is a number. We introduce a tuple of Lagrange multipliers

$$\lambda = (\alpha_0, \alpha, \beta, \psi, \psi_0)$$

such that

$$\alpha_0 \in \mathbb{R}^1,\ \alpha \in \mathbb{R}^{d(F)},\ \beta \in \mathbb{R}^{d(K)},\ \psi \in W^{1,1}(\hat{\Delta}, \mathbb{R}^{d(x)}),\ \psi_0 \in W^{1,1}(\hat{\Delta}, \mathbb{R}^1),$$

where $W^{1,1}(\hat{\Delta}, \mathbb{R}^{d(x)})$ is the space of absolutely continuous functions

$$\psi : \hat{\Delta} \to \mathbb{R}^{d(x)}.$$

Thus, λ is an element of the space

$$\mathcal{Y} := \mathbb{R}^{\infty} \times \mathbb{R}^{\lceil(\mathcal{F})} \times \mathbb{R}^{\lceil(\mathcal{K})} \times \mathcal{W}^{\infty,\infty}(\mathbb{A}, \mathbb{R}^{\lceil(\S)}) \times \mathcal{W}^{\infty,\infty}(\mathbb{A}, \mathbb{R}^{\infty}).$$

In the sequel, we shall denote first or second order partial derivatives by subscripts referring to the variables.

Denote by M_0 the set of the normalized multipliers $\lambda \in \mathcal{Y}$ satisfying the minimum principle conditions for the trajectory $\hat{\mathcal{T}}$:

$$\alpha_0 \geq 0, \quad \alpha \geq 0, \quad \alpha F(\hat{p}) = 0, \quad \alpha_0 + \sum \alpha_i + \sum |\beta_j| = 1, \tag{7}$$

$$\dot{\psi} = -H_x, \quad \dot{\psi}_0 = -H_t \quad \forall t \in \hat{\Delta} \setminus \hat{\theta}, \tag{8}$$

$$\psi(\hat{t}_0) = -l_{x_0}, \quad \psi(\hat{t}_1) = l_{x_1}, \quad \psi_0(\hat{t}_0) = -l_{t_0}, \quad \psi_0(\hat{t}_1) = l_{t_1}, \tag{9}$$

$$\min_{u \in U} H(t, \hat{x}(t), \psi(t), u) = H(t, \hat{x}(t), \psi(t), \hat{u}(t)) \quad \forall t \in \hat{\Delta} \setminus \hat{\theta}, \tag{10}$$

$$H(t, \hat{x}(t), \psi(t), \hat{u}(t)) + \psi_0(t) = 0 \quad \forall t \in \hat{\Delta} \setminus \hat{\theta}. \tag{11}$$

The derivatives l_{x_0} and l_{x_1} are taken at the point $(\alpha_0, \alpha, \beta, \hat{p})$, where $\hat{p} = (\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1))$, and the derivatives H_x, H_t are evaluated at the point $(t, \hat{x}(t), \hat{u}(t), \psi(t))$, $t \in \hat{\Delta} \setminus \hat{\theta}$. The condition $M_0 \neq \emptyset$ constitutes the first order necessary condition for a Pontryagin minimum of the trajectory $\hat{\mathcal{T}}$, which is the so called Pontryagin minimum principle, see, e.g. Pontryagin et al. (1961), Hestenes (1966), Milyutin, Osmolovskii (1998).

THEOREM 2.1 *If the trajectory $\hat{\mathcal{T}}$ affords a Pontryagin minimum, then the set M_0 is nonempty. The set M_0 is a finite-dimensional compact set and the projector $\lambda \mapsto (\alpha_0, \alpha, \beta)$ is injective on M_0.*

In the sequel, it will be convenient to use the simple abbreviation (t) for indicating all arguments $(t, \hat{x}(t), \hat{u}(t), \psi(t))$, e.g., $H(t) = H(t, \hat{x}(t), \hat{u}(t), \psi(t))$, $\sigma(t) = \sigma(t, \hat{x}(t), \psi(t))$. The continuity of the pair of functions $(\psi_0(t), \psi(t))$ at the points $\hat{\tau}_k \in \hat{\theta}$ constitutes the Weierstrass–Erdmann necessary conditions for nonsmooth extremals. We formulate one more condition of this type which is important for the statement of second-order conditions for extremal with jumps in the control. Namely, for $\lambda \in M_0$, $\hat{\tau}_k \in \hat{\theta}$ consider the function

$$(\Delta_k H)(t) = H(t, \hat{x}(t), \psi(t), \hat{u}^{k+}) - H(t, \hat{x}(t), \psi(t), \hat{u}^{k-}) = \sigma(t)\, [\hat{u}]^k. \quad (12)$$

PROPOSITION 2.1 *For each $\lambda \in M_0$ the following equalities hold*

$$\frac{d}{dt}(\Delta_k H)\big|_{t=\hat{\tau}_k - 0} = \frac{d}{dt}(\Delta_k H)\big|_{t=\hat{\tau}_k + 0}, \quad k = 1, \ldots, s.$$

Consequently, for each $\lambda \in M_0$ the function $(\Delta_k H)(t)$ is differentiable at the point $\hat{\tau}_k \in \hat{\theta}$. Define the quantity

$$D^k(H) = -\frac{d}{dt}(\Delta_k H)(\hat{\tau}_k).$$

Then the minimum condition (10) implies the following property:

PROPOSITION 2.2 *For each $\lambda \in M_0$ the following conditions hold:*

$$D^k(H) \geq 0, \quad k = 1, \ldots, s. \quad (13)$$

The value $D^k(H)$ can be written in the form

$$\begin{aligned} D^k(H) &= -H_x^{k+} H_\psi^{k-} + H_x^{k-} H_\psi^{k+} - [H_t]^k \\ &= \dot{\psi}^{k+} \dot{\hat{x}}^{k-} - \dot{\psi}^{k-} \dot{\hat{x}}^{k+} + [\psi_0]^k, \end{aligned}$$

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

$$D^k(H) = -\dot{\sigma}(\hat{\tau}_k - 0)[\hat{u}]^k = -\dot{\sigma}(\hat{\tau}_k + 0)[\hat{u}]^k. \quad (14)$$

2.3. Integral cost function, unessential variables, strong minimum

It is well known that any control problem with a cost functional in the integral form

$$\mathcal{J} = \int_{t_0}^{t_1} f_0(t, x(t), u(t))\, dt \quad (15)$$

can be represented in the canonical form (1) by introducing a new state variable y defined by the state equation

$$\dot{y} = f_0(t, x, u), \quad y(t_0) = 0\,. \tag{16}$$

This yields the cost function $\mathcal{J} = y(t_1)$. The component y is called an *unessential* component in the augmented problem. The general definition of an unessential component is as follows.

DEFINITION 2.2 *The state variable x_i, i.e., the i–th component of the state vector x is called* unessential *if the function f does not depend on x_i and if the functions F, J, K are affine in $x_{i0} = x_i(t_0)$ and $x_{i1} = x_i(t_1)$.*

Unessential components should not be taken into consideration in the definition of a minimum. This leads to the definition of a *strong minimum* which is stronger than the Pontryagin minimum in Definition 1. The strong minimum refers to the proximity of the state components in the trajectory only. In the following, let $\underline{x}$ denote the vector of all essential components of the state vector x.

DEFINITION 2.3 *We say that the trajectory $\mathcal{T}$ affords a strong minimum if there is no sequence of admissible trajectories*

$$\mathcal{T}^n = (x^n(t), u^n(t) \mid t \in [t_0^n, t_1^n]), \quad n = 1, 2, \ldots$$

such that
(a) $\mathcal{J}(\mathcal{T}^n) < \mathcal{J}(\mathcal{T})$,
(b) $t_0^n \to t_0, \quad t_1^n \to t_1, \quad x^n(t_0) \to x(t_0)$ *($n \to \infty$),*
(c) $\max\limits_{\Delta^n \cap \Delta} |\underline{x}^n(t) - \underline{x}(t)| \to 0$ *($n \to \infty$), where* $\Delta^n = [t_0^n, t_1^n]$.

The *strict* strong minimum is defined in a similar way by replacing the inequality $\mathcal{J}(\mathcal{T}^n) < \mathcal{J}(\mathcal{T})$ in (a) by $\mathcal{J}(\mathcal{T}^n) \leq \mathcal{J}(\mathcal{T})$ where the trajectory $\mathcal{T}^n$ is required to be different from $\mathcal{T}$ for each n.

2.4. Bang-bang control

By definition, a bang–bang control assumes values only in the set of vertices of the admissible polyhedron U in (2). We shall need a slightly more restrictive property to obtain the sufficient conditions in Theorem 3.2 (see definition (12.53) and Theorem 12.9 in Milyutin and Osmolovskii, 1998, Part 2, Chapter 3, Section 12.3). For a given $\lambda \in M_0$, let

$$\text{Arg}\min_{v \in U} \sigma(t)v$$

be the set of points $v \in U$ where the minimum of the linear function $\sigma(t)v$ is attained. For a given extremal trajectory $\hat{\mathcal{T}} = \{\, (\hat{x}(t), \hat{u}(t)) \mid t \in \hat{\Delta}\,\}$ with a piecewise constant control $\hat{u}(t)$ we say that $\hat{u}(t)$ is a *regular* (or *strict*, or *nonsingular*) *bang-bang control* if there exists $\lambda = (\alpha_0, \alpha, \beta, \psi, \psi_0) \in M_0$ such that

$$\text{Arg min}_{v \in U} \sigma(t) v = [\hat{u}(t-0), \hat{u}(t+0)] \quad \forall\, t \in \hat{\Delta}, \tag{17}$$

where $[\hat{u}(t-0), \hat{u}(t+0)]$ denotes the line segment spanned by the vectors $\hat{u}(t-0)$ and $\hat{u}(t+0)$ in $\mathbb{R}^{d(u)}$. Note that $[\hat{u}(t-0), \hat{u}(t+0)]$ is a singleton $\{\hat{u}(t)\}$ at each continuity point of the control $\hat{u}(t)$ with $\hat{u}(t)$ being a vertex of the polyhedron U. Only at the points $\hat{\tau}_k \in \hat{\theta}$ the line segment $[\hat{u}^{k-}, \hat{u}^{k+}]$ does coincide with an edge of the polyhedron.

If the control is *scalar*, $d(u) = 1$ and $U = [u_{\min}, u_{\max}]$, then the regular bang–bang property is equivalent to

$$\sigma(t) \neq 0 \quad \forall\, t \in \hat{\Delta} \setminus \hat{\theta},$$

which yields the control law

$$\hat{u}(t) = \left\{ \begin{array}{lll} u_{\min}, & \text{if} & \sigma(t) > 0 \\ u_{\max}, & \text{if} & \sigma(t) < 0 \end{array} \right\} \quad \forall\, t \in \hat{\Delta} \setminus \hat{\theta}. \tag{18}$$

For vector–valued control inputs, condition (17) imposes further restrictions. For example, if U is the unit cube in $\mathbb{R}^{d(u)}$, condition (17) precludes simultaneous switching of any two different components of the control. This property holds in many examples. However, a modification of this condition, which would include the situation of simultaneous switching of control components, is an interesting question (see Felgenhauer, 2005).

3. Quadratic necessary and sufficient optimality conditions for bang-bang control problems

In this section, we shall formulate a quadratic necessary optimality condition of a Pontryagin minimum (Definition 2.1) for a given bang–bang control. A strengthening of this quadratic condition yields a quadratic sufficient condition for a strong minimum (Definition 2.3). These quadratic conditions are based on some properties of a quadratic form on the so called *critical cone* whose elements are first order variations along a given trajectory $\hat{\mathcal{T}}$. The main results of this section (Theorems 3.1 and 3.2) are due to Osmolovskii; see Milyutin and Osmolovskii (1998), Part 2, Chapter 3. The proofs missing in that book are given in Osmolovskii (2004).

3.1. Critical cone

Denote by $P_{\hat{\theta}} C^1(\hat{\Delta}, \mathbb{R}^{d(x)})$ the space of piecewise continuous functions

$$\bar{x}(\cdot) : \hat{\Delta} \to \mathbb{R}^{d(x)},$$

that are continuously differentiable on each interval of the set $\hat{\Delta} \setminus \hat{\theta}$. For each $\bar{x} \in P_{\hat{\theta}}C^1(\hat{\Delta}, \mathbb{R}^{d(x)})$ and for $\hat{\tau}_k \in \hat{\theta}$ we set

$$\bar{x}^{k-} = \bar{x}(\hat{\tau}_k - 0), \quad \bar{x}^{k+} = \bar{x}(\hat{\tau}_k + 0), \quad [\bar{x}]^k = \bar{x}^{k+} - \bar{x}^{k-}.$$

Set $\bar{z} = (\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x})$, where $\bar{t}_0, \bar{t}_1 \in \mathbb{R}^1$, $\bar{\xi} \in \mathbb{R}^s$, $\bar{x} \in P_{\hat{\theta}}C^1(\hat{\Delta}, \mathbb{R}^{d(x)})$. Thus,

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

For each $\bar{z}$ we set

$$\tilde{x}_0 = \bar{x}(\hat{t}_0) + \bar{t}_0\dot{\hat{x}}(\hat{t}_0), \quad \tilde{x}_1 = \bar{x}(\hat{t}_1) + \bar{t}_1\dot{\hat{x}}(\hat{t}_1), \quad \tilde{p} = (\bar{t}_0, \tilde{x}_0, \bar{t}_1, \tilde{x}_1). \tag{19}$$

The vector $\tilde{p}$ is understood as a column vector. Note that $\bar{t}_0 = 0$ and $\bar{t}_1 = 0$ hold for a fixed initial time t_0 and a fixed final time t_1, respectively. Let $I_F(\hat{p}) = \{i \in \{1, \ldots, d(F)\} \mid F_i(\hat{p}) = 0\}$ be the set of indices of all active endpoint inequalities $F_i(\hat{p}) \leq 0$ at the point $\hat{p} = (\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1))$. Denote by $\mathcal{K}$ the set of all $\bar{z} \in \mathcal{Z}(\hat{\theta})$ satisfying the following conditions, where the prime denotes the total derivative:

$$J'(\hat{p})\tilde{p} \leq 0, \quad F_i'(\hat{p})\tilde{p} \leq 0 \; \forall\, i \in I_F(\hat{p}), \quad K'(\hat{p})\tilde{p} = 0, \tag{20}$$

$$\dot{\bar{x}}(t) = f_x(t, \hat{x}(t), \hat{u}(t))\bar{x}(t), \quad [\bar{x}]^k = [\dot{\hat{x}}]^k\bar{\xi}_k, \quad k = 1, \ldots, s. \tag{21}$$

It is obvious that $\mathcal{K}$ is a convex finite-dimensional cone with finitely many faces in the space $\mathcal{Z}(\hat{\theta})$. We call $\mathcal{K}$ the *critical cone*. Each element $\bar{z} \in \mathcal{K}$ is uniquely determined by numbers $\bar{t}_0$, $\bar{t}_1$, a vector $\bar{\xi}$ and an initial value $\bar{x}(\hat{t}_0)$ of the function $\bar{x}(t)$. The following two important properties of the critical cone were proved in Maurer and Osmolovskii (2003).

Proposition 3.1 *For any $\lambda \in M_0$ and $\bar{z} \in \mathcal{K}$, we have*

$$\alpha_0 J'(\hat{p})\tilde{p} = 0, \quad \alpha_i F_i'(\hat{p})\tilde{p} = 0 \; \forall\, i \in I_F(\hat{p}).$$

Proposition 3.2 *Suppose that there exist $\lambda \in M_0$ with $\alpha_0 > 0$. Then adding the equalities $\alpha_i F_i'(\hat{p})\tilde{p} = 0 \; \forall i \in I_F(\hat{p})$, to the system (20), (21) defining $\mathcal{K}$, one can omit the inequality $J'(\hat{p})\tilde{p} \leq 0$, in that system without affecting $\mathcal{K}$.*

Thus, $\mathcal{K}$ is defined by condition (21) and by the condition $\tilde{p} \in \mathcal{K}_0$, where $\mathcal{K}_0$ is the cone in $\mathbb{R}^{2d(x)+2}$ given by (20). But if there exists $\lambda \in M_0$ with $\alpha_0 > 0$, then we can put

$$\mathcal{K}_0 = \{\tilde{p} \in \mathbb{R}^{d(x)+2} \mid F_i'(p)\tilde{p} \leq 0, \; \alpha_i F_i'(p)\tilde{p} = 0 \; \forall\, i \in I_F(p), \; K'(p)\tilde{p} = 0\}.$$

If, in addition, $\alpha_i > 0$ holds for all $i \in I_F(p)$, then $\mathcal{K}_0$ is a subspace in $\mathbb{R}^{d(x)+2}$.

An explicit representation of the variations $\bar{x}(t)$ in (21) is obtained as follows. For each $k = 1, ..., s$, define the vector functions $y^k(t)$ as the solutions to the system

$$\dot{y} = f_x(t)y, \quad y(\hat{\tau}_k) = [\dot{\hat{x}}]^k, \quad t \in [\hat{\tau}_k, \hat{t}_1].$$

For $t < \hat{\tau}_k$ we put $y^k(t) = 0$ which yields the jump $[y^k]^k = [\dot{\hat{x}}]^k$. Moreover, define $y^0(t)$ as the solution to the system

$$\dot{y} = f_x(t)y, \quad y(\hat{t}_0) = \bar{x}(\hat{t}_0) =: \bar{x}_0.$$

By the superposition principle for linear ODEs it is obvious that we have

$$\bar{x}(t) = \sum_{k=1}^{s} y^k(t)\bar{\xi}_k + y^0(t)$$

from which we obtain the representation

$$\tilde{x}_1 = \sum_{k=1}^{s} y^k(\hat{t}_1)\bar{\xi}_k + y^0(\hat{t}_1) + \dot{x}(\hat{t}_1)\bar{t}_1.$$

Furthermore, denote by $x(t; \tau_1, ..., \tau_s)$ the solution of the state equation (2) using the values of the optimal bang–bang control with switching points $\tau_1, ...\tau_s$. It easily follows from elementary properties of ODEs that the partial derivatives of state trajectories w.r.t. the switching points are given by

$$\frac{\partial x}{\partial \tau_k}(t; \hat{\tau}_1, ..., \hat{\tau}_s) = -y^k(t) \quad \text{for } \ t \geq \hat{\tau}_k, \ \ k = 1, ..., s.$$

This gives the following expression for $\bar{x}(t)$:

$$\bar{x}(t) = -\sum_{k=1}^{s} \frac{\partial x}{\partial \tau_k}(t)\bar{\xi}_k + y^0(t).$$

3.2. Quadratic necessary optimality conditions

Let us introduce a quadratic form on the critical cone $\mathcal{K}$ defined by conditions (20), (21). For each $\lambda \in M_0$ and $\bar{z} \in \mathcal{K}$ we set

$$\Omega(\lambda, \bar{z}) = \langle A\tilde{p}, \tilde{p}\rangle + \sum_{k=1}^{s} \big(D^k(H)\bar{\xi}_k^2 + 2[H_x]^k \bar{x}_{\mathrm{av}}^k \bar{\xi}_k\big) + \int_{\hat{\Delta}} \langle H_{xx}\bar{x}(t), \bar{x}(t)\rangle \, dt, \quad (22)$$

where

$$\begin{aligned} \langle A\tilde{p}, \tilde{p}\rangle \;\; &= \;\; \langle l_{pp}\tilde{p}, \tilde{p}\rangle + 2\dot{\psi}(\hat{t}_0)\tilde{x}_0\bar{t}_0 + (\dot{\psi}_0(\hat{t}_0) - \dot{\psi}(\hat{t}_0)\dot{\hat{x}}(\hat{t}_0))\bar{t}_0^2 \\ &\quad -2\dot{\psi}(\hat{t}_1)\tilde{x}_1\bar{t}_1 - (\dot{\psi}_0(\hat{t}_1) - \dot{\psi}(\hat{t}_1)\dot{\hat{x}}(\hat{t}_1))\bar{t}_1^2, \qquad (23) \end{aligned}$$

$$l_{pp} = l_{pp}(\alpha_0, \alpha, \beta, \hat{p}), \ \ \hat{p} = (\hat{t}_0, \hat{x}(\hat{t}_0), \hat{t}_1, \hat{x}(\hat{t}_1)), \ \ \tilde{p} = (\bar{t}_0, \tilde{x}_0, \bar{t}_1, \tilde{x}_1),$$

$$\bar{x}_{\mathrm{av}}^k = \frac{1}{2}(\bar{x}^{k-} + \bar{x}^{k+}), \ \ H_{xx} = H_{xx}(t, \hat{x}(t), \psi(t), \hat{u}(t)).$$

Note that the functional $\Omega(\lambda, \bar{z})$ is linear in λ and quadratic in $\bar{z}$. Also note that for a problem on a fixed time interval $[t_0, t_1]$ we have $\bar{t}_0 = \bar{t}_1 = 0$

and, hence, the quadratic form (23) reduces to $\langle A\tilde{p}, \tilde{p}\rangle = \langle l_{pp}\bar{p}, \bar{p}\rangle$, where $\bar{p} = (\bar{t}_0, \bar{x}(t_0), \bar{t}_1, \bar{x}(t_1)) = (0, \bar{x}(t_0), 0, \bar{x}(t_1))$.

The following theorem gives the main second order necessary condition of optimality in the main problem (1)–(4) (see Theorem 12.7 in Milyutin and Osmolovskii, 1998, Part 2, Chapter 3, Section 12.3, p. 306).

THEOREM 3.1 *If the trajectory $\hat{\mathcal{T}}$ affords a Pontryagin minimum, then the following Condition $\mathcal{A}$ holds: the set M_0 is nonempty and*

$$\max_{\lambda \in M_0} \Omega(\lambda, \bar{z}) \geq 0 \quad \textit{for all} \;\; \bar{z} \in \mathcal{K}.$$

We call Condition $\mathcal{A}$ the necessary quadratic condition, although it is truly quadratic only if M_0 is a singleton.

3.3. Quadratic sufficient optimality conditions

A natural strengthening of the necessary Condition $\mathcal{A}$ turns out to be a sufficient optimality condition not only for a Pontryagin minimum, but also for a strong minimum in the main problem (1)–(4); see Definition 2.3. The following theorem was obtained by Osmolovskii; see Milyutin and Osmolovskii (1998), Part 2, Chapter 3, Section 12.3, Theorem 12.9, p. 307. The proofs missing in that book are given in Osmolovskii (2004).

THEOREM 3.2 *Let the following Condition $\mathcal{B}$ be fulfilled for $\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$:*

(a) $\hat{u}(t)$ is a regular bang-bang control (hence M_0 is nonempty and condition (17) holds for some $\lambda \in M_0$),

(b) there exists $\lambda \in M_0$ such that $D^k(H) > 0,\; k = 1, \dots, s$,

(c) $\max_{\lambda \in M_0} \Omega(\lambda, \bar{z}) > 0$ for all $\bar{z} \in \mathcal{K} \setminus \{0\}$.

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

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

$$\Omega(\lambda, \bar{z}) > 0 \text{ for all } \bar{z} \in \mathcal{K} \setminus \{0\}. \tag{24}$$

3.4. Discussion of the proofs of quadratic conditions

The complete proofs of Theorems 3.1 and 3.2 are given in the book of Milyutin and Osmolovskii (1998), Part 2, Chapter 3, Section 12, and in the paper by Osmolovskii (2004). Below we shall briefly recall the general results on second order conditions for broken extremals and show how these results were used in Milyutin and Osmolovskii (1998) to get Theorems 3.1 and 3.2, first for a fixed and then for a variable interval of control. The subsection can be omitted by those readers, who are not interested in the details of the proofs.

Quadratic optimality conditions for broken extremal in the general problem of the calculus of variations. Consider the following problem on a fixed time interval $[t_0, t_1]$ with a pointwise equality-type constraint:

$$\mathcal{J}(w) = J(x_0, x_1) \to \min, \tag{25}$$

$$F(x_0, x_1) \le 0, \qquad K(x_0, x_1) = 0, \quad (x_0, x_1) \in \mathcal{P}, \tag{26}$$

$$\dot{x} = f(t, x, u), \qquad g(t, x, u) = 0, \quad (t, x, u) \in \mathcal{Q}, \tag{27}$$

where, by definition, $x_0 = x(t_0)$, $x_1 = x(t_1)$, $w = (x, u)$. It is assumed that the functions J, F, and K are twice continuously differentiable on an open set $\mathcal{P} \subset \mathbb{R}^{2d(x)}$, and f and g are twice continuously differentiable on an open set $\mathcal{Q} \subset \mathbb{R}^{1+d(x)+d(u)}$. Moreover, the following local full rank condition is assumed to be satisfied:

$$\text{rank } g_u(t, x, u) = d(g) \tag{28}$$

for all $(t, x, u) \in \mathcal{Q}$ such that $g(t, x, u) = 0$.

We are looking for the minimum in the set of pairs of functions $w(t) = (x(t), u(t))$ such that $x(t)$ is an absolutely continuous function on $[t_0, t_1]$ and $u(t)$ is a bounded measurable function on $\Delta = [t_0, t_1]$. Hence the minimum is sought over all pairs $w = (x, u)$ in the space

$$W := W^{1,1}(\Delta, \mathbb{R}^{d(x)}) \times L^\infty(\Delta, \mathbb{R}^{d(u)}).$$

Consider an admissible trajectory $w^0(t) = (x^0(t), u^0(t))$ in the space W such that $u^0(t)$ is a piecewise Lipschitz continuous function on the interval $\Delta = [t_0, t_1]$, i.e. $u^0(t)$ is a piecewise continuous and Lipschitz continuous on each continuity interval. The set of discontinuity points of $u^0(t)$ will be denoted here by

$$\theta = \{t^1, \ldots, t^s\}, \quad t_0 < t^1 < \ldots < t^s < t_1.$$

Denote by H and $\bar{H}$ the Pontryagin function and the extended Pontryagin function, respectively:

$$H(t, x, u, \psi) = \psi f(t, x, u), \quad \bar{H}(t, x, u, \psi, \nu) = H(t, x, u, \psi) + \nu g(t, x, u),$$

where $\nu \in \mathbb{R}^{d(g)}$. Denote by l the endpoint Lagrange function:

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

where $p = (x_0, x_1)$. Denote by M_0 the set of normalized Lagrange multipliers $\lambda = (\alpha_0, \alpha, \beta, \psi(\cdot), \nu(\cdot))$ satisfying the minimum principle conditions for the trajectory $w^0(\cdot) = (x^0(\cdot), u^0(\cdot))$:

$$\alpha_0 \ge 0, \quad \alpha \ge 0, \quad \alpha F(p^0) = 0, \quad \alpha_0 + \sum \alpha_i + \sum |\beta_j| = 1,$$

$$\dot{\psi} = -\bar{H}_x, \quad \psi(t_0) = -l_{x_0}, \quad \psi(t_1) = l_{x_1}, \quad \bar{H}_u = 0,$$

$$\min_{u \in U(t, x^0(t))} H(t, x^0(t), u, \psi(t)) = H(t, x^0(t), u^0(t), \psi(t)),$$

where

$$U(t,x) = \{u \in \mathbb{R}^{d(u)} \mid (t,x,u) \in \mathcal{Q}, \quad g(t,x,u) = 0\}.$$

Denote by $\mathcal{K}$ the set of triples $\bar{z} = (\bar{\xi}, \bar{x}(\cdot), \bar{u}(\cdot))$ such that $\bar{\xi} \in \mathbb{R}^s$, $\bar{u}(\cdot) \in L^2$, $\bar{x}(\cdot)$ is a piecewise continuous function, absolutely continuous on each interval of the set $\Delta \setminus \theta$, and the following conditions are fulfilled:

$$J'(p^0)\bar{p} \le 0, \quad F_i'(p^0)\bar{p} \le 0,\ i \in I_F(p^0), \quad K_p'(p^0)\bar{p} = 0,$$

$$\dot{\bar{x}}(t) = f_x(t, w^0(t))\bar{x}(t) + f_u(t, w^0(t))\bar{u}(t), \quad [\bar{x}]^k = [\dot{x}^0]^k \bar{\xi}_k, \quad k = 1, \dots, s,$$

$$g_x(t, w^0(t))\bar{x}(t) + g_u(t, w^0(t))\bar{u}(t) = 0,$$

where $I_F(p^0) = \{i \in \{1, \dots, d(F)\} \mid F_i(p^0) = 0\}$ is the set of active indices, $\bar{p} = (\bar{x}(t_0), \bar{x}(t_1))$, $p^0 = (x^0(t_0), x^0(t_1))$. For $\lambda \in M_0$ define the following quadratic form $\Omega(\lambda, \cdot)$ on the critical cone $\mathcal{K}$:

$$\Omega(\lambda, \bar{z}) = \langle l_{pp}\bar{p}, \bar{p}\rangle + \int_{t_0}^{t_1} \langle \bar{H}_{ww}\bar{w}, \bar{w}\rangle\, dt + \sum_{k=1}^{s} \left(D^k(\bar{H})\bar{\xi}_k^2 + 2[\bar{H}_x]^k \bar{x}_{\mathrm{av}}^k \bar{\xi}_k \right),$$

where $w = (x,u)$, $\bar{w} = (\bar{x}, \bar{u})$, $D^k(\bar{H}) = -\frac{d}{dt}(\Delta^k \bar{H})|_{t=t^k}$, $\Delta^k \bar{H} = \bar{H}(t, x^0(t), u^0(t^k+0), \psi(t), \nu(t^k+0)) - \bar{H}(t, x^0(t), u^0(t^k-0), \psi(t), \nu(t^k-0))$,
$l_{pp} = l_{pp}(\alpha_0, \alpha, \beta, p^0)$, $\bar{H}_{ww} = \bar{H}_{ww}(t, w^0(t))$.

We say that w^0 is a *point of a Pontryagin minimum* for the problem (25)-(27) if there is no sequence of admissible points $w^n = (x^n, u^n) \in W$ such that $\mathcal{J}(w^n) < \mathcal{J}(w^0)$ $\forall n$ and the sequence $\{w^n\}$ converges to w^0 in the Pontryagin sense. The latter means that $\max_{t \in [t_0,t_1]} |x^n(t) - x^0(t)| \to 0$, $\int_{t_0}^{t_1} |u^n(t) - u^0(t)|\, dt \to 0$ $(n \to \infty)$ and there exists a compact set $\mathcal{C} \subset \mathcal{Q}$ such that for all sufficiently large n we have $(t, x^n(t), u^n(t)) \in \mathcal{C}$ a.e. on Δ.

THEOREM 3.3 *If w^0 is a point of a Pontryagin minimum in problem (25)-(27), then the set M_0 is nonempty and*

$$\max_{\lambda \in M_0} \Omega(\lambda, \bar{z}) \ge 0 \quad \forall \bar{z} \in \mathcal{K}.$$

The proof of this theorem is given in Osmolovskii (2004), see Chapter III, Section 11.5, Theorem 11.1.

Now we proceed to formulations of sufficient optimality conditions in problem (25)-(27) at the point w^0. Assume that for this point the set M_0 is nonempty. Note that for each $\lambda = (\alpha_0, \alpha, \beta, \psi(\cdot), \nu(\cdot)) \in M_0$ the function $H(t, x^0(t), u^0(t), \psi(t)$ is continuous in $t \in \Delta$; in particular $H^{k-} = H^{k+}$ for all $t^k \in \theta$, where $H^{k-} := H(t^k, x^0(t^k), u^0(t^k - 0), \psi(t^k))$, $H^{k+} := H(t^k, x^0(t^k), u^0(t^k + 0), \psi(t^k))$. For given $\lambda \in M_0$, $t^k \in \theta$ we set $H^k := H^{k-} = H^{k+}$.

Denote by M_0^+ the subset of all elements $\lambda \in M_0$ satisfying the *strict minimum principle* determined by conditions:

a) $H(t, x^0(t), u, \psi(t)) > H(t, x^0(t), u^0(t), \psi(t))$ for all $t \in \Delta \backslash \theta$, $u \in U(t, x^0(t))$, $u \neq u^0(t)$, and

b) $H(t^k, x^0(t^k), u, \psi(t^k)) > H^k$ for all $t^k \in \theta$, $u \in U(t^k, x^0(t^k))$, $u \neq u^0(t^k - 0)$, $u \neq u^0(t^k + 0)$.

Let us define a notion of a strictly Legendrian element. An element $\lambda \in M_0$ will be called *strictly Legendrian* if the following conditions are fulfilled:

$$D^k(\bar{H}) > 0, \quad k = 1, \dots, s,$$

and the strengthened Legendre-Clebsch condition holds, namely,

- for any $t \in \Delta \setminus \theta$ the quadratic form
  $\langle \bar{H}_{uu}(t, x^0(t), u^0(t), \psi(t), \nu(t))\bar{u}, \bar{u}\rangle$
  is positive definite on the subspace of vectors $\bar{u} \in \mathbb{R}^{d(u)}$ such that
  $g_u(t, x^0(t), u^0(t))\bar{u} = 0$;
- for any $t^k \in \theta$ the quadratic form
  $\langle \bar{H}_{uu}(t^k, x^0(t^k), u^0(t^k - 0), \psi(t^k), \nu(t^k - 0))\bar{u}, \bar{u}\rangle$
  is positive definite on the subspace of vectors $\bar{u} \in \mathbb{R}^{d(u)}$ such that
  $g_u(t^k, x^0(t^k), u^0(t^k - 0))\bar{u} = 0$;
- for any $t^k \in \theta$ the quadratic form
  $\langle \bar{H}_{uu}(t^k, x^0(t^k), u^0(t^k + 0), \psi(t^k), \nu(t^k + 0))\bar{u}, \bar{u}\rangle$
  is positive definite on the subspace of vectors $\bar{u} \in \mathbb{R}^{d(u)}$ such that
  $g_u(t^k, x^0(t^k), u^0(t^k + 0))\bar{u} = 0$.

Denote by $\mathrm{Leg}_+(M_0^+)$ the set of all strictly Legendrian elements $\lambda \in M_0^+$.

Set

$$\sigma(w) = (J(p) - J(p^0))_+ + \sum_{i=1}^{d(F)} (F_i(p))_+ + |K(p)| + \int_{t_0}^{t_1} |\dot{x}(t) - f(t, x(t), u(t))|\, dt,$$

where $w = (x, u)$, $p = (x(t_0), x(t_1))$, $a_+ = \max\{a, 0\}$. Denote

$$\gamma_1(w - w^0) = \max_{t \in \Delta} |x(t) - x^0(t)|^2 + \left(\int_{t_0}^{t_1} |u(t) - u^0(t)|\, dt \right)^2.$$

We say that the *bounded-strong γ_1-sufficiency* holds in the problem (25)-(27) at the point w^0 if there is no sequence $\{w^n\} = \{(x^n, u^n)\} \subset W$ such that $\sigma(w^n) = o(\gamma(w^n - w^0))$, and the following conditions are fulfilled:

(a) we have $\max_{t \in \Delta} |\underline{x}^n(t) - \underline{x}^0(t)| \to 0$, where the vector function $\underline{x}^n$ consists of essential components of the vector function x^n;

(b) there exists a compact set $\mathcal{C} \subset \mathcal{Q}$ such that for any n we have $(t, x^n(t), u^n(t)) \in \mathcal{C}$ a.e. on Δ;

(c) for any n we have $g(t, x^n(t), u^n(t)) = 0$ a.e. on Δ

If $\{(t, x, u) \in \mathcal{Q} \mid g(t, x, u) = 0\}$ is a compact set, then the property (b) is automatically fulfilled. In this case instead of the term "bounded-strong γ_1-sufficiency" we shall use the term "strong γ_1-sufficiency".

THEOREM 3.4 *Assume that, at a point w^0 the set $\mathrm{Leg}_+(M_0^+)$ is nonempty and there exist $\varepsilon > 0$ and a nonempty compact set $M \subset \mathrm{Leg}_+(M_0^+)$ such that*

$$\max_{\lambda \in M} \Omega(\lambda, \bar{z}) \geq \varepsilon \bar{\gamma}(\bar{z}) \quad \forall \bar{z} \in \mathcal{K},$$

where

$$\bar{\gamma}(\bar{z}) = \langle \bar{\xi}, \bar{\xi} \rangle + \langle \bar{x}(t_0), \bar{x}(t_0) \rangle + \int_{t_0}^{t_1} \langle \bar{u}(t), \bar{u}(t) \rangle \, dt.$$

Then the bounded-strong γ_1-sufficiency holds at the point w^0.

The proof of Theorem 3.4 follows from Theorem 12.1 in Osmolovskii (2004), Chapter III, Section 12.3, as well as from an estimate derived in Milyutin and Osmolovskii (1998), Part 2, Proposition 12.2, p. 300.

Linear in control problem on a fixed time interval. Consider the simplified version of the main problem, where the interval $[t_0, t_1]$ is fixed. Namely:

$$J(x_0, x_1) \to \min, \quad F(x_0, x_1) \leq 0, \quad K(x_0, x_1) = 0, \quad (x_0, x_1) \in \mathcal{P}, \tag{29}$$

$$\dot{x} = a(t, x) + B(t, x)u, \quad (t, x) \in \mathcal{Q}_{tx}, \quad u \in U. \tag{30}$$

We use the same notations and assumptions as in (1)–(4).

Let u^i, $i = 1, \ldots, m$ be the vertices of the polyhedron U, and denote $V = \{u^1, \ldots, u^m\}$. Consider the admissible pair $w^0 = (x^0, u^0)$ where the control $u^0(t)$ is a piecewise constant function taking values in the vertices of U, i.e. $u^0(t) \in V$ for all $t \in \Delta$. Again, we denote by $\theta = \{t^1, \ldots, t^s\}$ the set of discontinuity points (switchings) of the control $u^0(t)$.

Let $Q_u^i \subset \mathbb{R}^{d(u)}$, $i = 1, \ldots, m$ be disjoint open neighborhoods of the vertices $u^i \in V$. Set $Q_u = \bigcup_{i=1}^{m} Q_u^i$. Define the function $g(u) : Q_u \to \mathbb{R}^{d(u)}$ by setting $g(u) = u - u^i$ on each $Q_u^i \subset Q_u$, $i = 1, \ldots, m$. Then $g(u)$ is a function of class C^∞ on Q_u defining the set of vertices of U, i.e. $V = \{u \in Q_u \mid g(u) = 0\}$. Moreover, $g'_u(u) = I$ for all $u \in Q_u$. Hence the full rank condition (28) is fulfilled.

Now, consider the following problem P:

$$J(x_0, x_1) \to \min, \quad F(x_0, x_1) \leq 0, \quad K(x_0, x_1) = 0, \quad (x_0, x_1) \in \mathcal{P}, \tag{31}$$

$$\dot{x} = a(t, x) + B(t, x)u, \quad (t, x) \in \mathcal{Q}_{tx}, \quad g(u) = 0, \quad u \in Q_u. \tag{32}$$

By virtue of relations $V = \{u \in Q_u \mid g(u) = 0\}$ and $U = \mathrm{co}\, V$ (where $\mathrm{co}\, V$ denotes the convex hull of V) problem (29), (30) can be considered as the convexification of the problem P. Therefore we refer to the problem (29), (30) as to the problem $\operatorname{co} P$.

It is easy to see that we can use here Theorem 3.3 to derive the necessary conditions for w^0 to be a Pontryagin minimum in problem P. Obviously, the necessary optimality conditions for Pontryagin minimum in the problem P are also necessary in the problem $\operatorname{co} P$. It leads to the proof of Theorem 3.1 in the case of a fixed time interval $[t_0, t_1]$ (see Milyutin and Osmolovskii, 1998, Part 2, Chapter 3, Section 12.1, Theorem 12.1).

Now we turn to quadratic sufficient conditions. Applying Theorem 3.4, we derive condition under which a bounded-strong γ_1-sufficiency holds at w^0 for P. Since V is a compact set, a bounded-strong γ_1-sufficiency for P is equivalent to a strong γ_1-sufficiency.

There are examples where the convexification leads to the loss of the minimum. However, a remarkable fact is that the convexification of the constraint $u \in V$ turns a strong γ_1-sufficiency into a strong minimum (see Theorem 12.3 in Milyutin and Osmolovskii, 1998, Part 2, Section 12.1). It leads to the proof of Theorem 3.2 in the case of a fixed time interval $[t_0, t_1]$. The above very simple but somewhat unexpected way of using equality constraints in the problem with constraints on the control specified by a polyhedron is due to A.A. Milyutin.

Linear in control problem on a variable time interval. In order to extend the proofs to the case of a variable time interval $[t_0, t_1]$ we used (see Milyutin and Osmolovskii, 1998, Part 2, Section 12.2) a simple change of the time variable. Namely, with the admissible control process $\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ in problem (1)-(4), we associate the process $(x^0(\tau), t^0(\tau), v^0(\tau), u^0(\tau))$, $\tau \in [\tau_0, \tau_1]$, where $x^0(\tau), t^0(\tau), v^0(\tau)$ are state variables, $u^0(\tau)$ is a control, $\tau_0 = \hat{t}_0$, $\tau_1 = \hat{t}_1$, $t^0(\tau) = \tau$, $v^0(\tau) = 1$, $x^0(\tau) = \hat{x}(\tau)$, $u^0(\tau) = \hat{u}(\tau)$. Thus, we get an admissible process for the following problem on a fixed interval $[\tau_0, \tau_1]$:

$$J(t(\tau_0), x(\tau_0), t(\tau_1), x(\tau_1)) \to \min, \quad F(t(\tau_0), x(\tau_0), t(\tau_1), x(\tau_1)) \le 0,$$

$$K(t(\tau_0), x(\tau_0), t(\tau_1), x(\tau_1)) = 0, \quad (t(\tau_0), x(\tau_0), t(\tau_1), x(\tau_1)) \in \mathcal{P},$$

$$\frac{dx(\tau)}{d\tau} = v(\tau)\left(a(t(\tau), x(\tau)) + B(t(\tau), x(\tau))u(\tau)\right),$$

$$\frac{dt(\tau)}{d\tau} = v(\tau), \quad \frac{dv(\tau)}{d\tau} = 0, \quad (t(\tau), x(\tau)) \in Q_{tx}, \quad u(\tau) \in U.$$

Necessary and sufficient quadratic optimality conditions written for the process $(x^0(\tau), t^0(\tau), v^0(\tau), u^0(\tau)$ in this new problem on a fixed time interval give us the corresponding quadratic optimality conditions for the process $\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ in problem (1)-(4) on a variable time interval.

4. Main results

4.1. Induced optimization problem

Again, let $\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ be an admissible trajectory for the main problem (1)-(3). Assume that $\hat{u}(t)$ is a bang-bang control in $\hat{\Delta} = [\hat{t}_0, \hat{t}_1]$ taking values in the set of vertices V of the polyhedron U,

$$\hat{u}(t) = u^i \in V \text{ for } t \in (\hat{\tau}_{i-1}, \hat{\tau}_i), \quad i = 1, \ldots, s+1,$$

where $\hat{\tau}_0 = \hat{t}_0$, $\hat{\tau}_{s+1} = \hat{t}_1$. Thus, $\hat{\theta} = \{\hat{\tau}_1, \ldots, \hat{\tau}_s\}$ is the set of switching points of the control $\hat{u}(\cdot)$ with $\hat{\tau}_i < \hat{\tau}_{i+1}$ for $i = 0, 1, ..., s$. Assume now that the set M_0 of multipliers is nonempty for the trajectory $\hat{\mathcal{T}}$. Put

$$\hat{x}(\hat{t}_0) = \hat{x}_0, \qquad \hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_s), \qquad \hat{\zeta} = (\hat{t}_0, \hat{t}_1, \hat{x}_0, \hat{\tau}). \tag{33}$$

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

Take a small neighbourhood $\mathcal{V}$ of the point $\hat{\zeta}$ in $\mathbb{R}^2 \times \mathbb{R}^n \times \mathbb{R}^s$, and let $\zeta = (t_0, t_1, x_0, \tau) \in \mathcal{V}$, where $\tau = (\tau_1, \ldots, \tau_s)$ satisfies $t_0 < \tau_1 < \tau_2 < \ldots < \tau_s < t_1$. Define the function $u(t; \tau)$ by the condition

$$u(t; \tau) = u^i \text{ for } t \in (\tau_{i-1}, \tau_i), \quad i = 1, \ldots, s+1, \tag{34}$$

where $\tau_0 = t_0$, $\tau_{s+1} = t_1$. The values $u(\tau_i; \tau)$, $i = 1, \ldots, s$, may be chosen in U arbitrarily. For definiteness, define them by the condition of continuity of the control from the left: $u(\tau_i; \tau) = u(\tau_i - 0; \tau)$, $i = 1, \ldots, s$.

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

$$\dot{x} = f(t, x, u(t; \tau)), \quad t \in [t_0, t_1], \qquad x(t_0) = x_0. \tag{35}$$

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

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

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 $\zeta = (t_0, t_1, x_0, \tau)$:

$$\begin{aligned}
&\mathcal{F}_0(\zeta) := J(t_0, x_0, t_1, x(t_1; t_0, x_0, \tau)) \to \min, \\
&\mathcal{F}(\zeta) := F(t_0, x_0, t_1, x(t_1; t_0, x_0, \tau)) \le 0, \\
&\mathcal{G}(\zeta) := K(t_0, x_0, t_1, x(t_1; t_0, x_0, \tau)) = 0.
\end{aligned} \tag{36}$$

We shall call (36) *the induced optimization problem of Agrachev, Stefani, Zezza* or simply the *induced problem* (see Agrachev, Stefani, Zezza, 2002). The following assertion is almost obvious.

THEOREM 4.1 *Let the trajectory $\hat{\mathcal{T}}$ be a Pontryagin local minimum for the main control problem (1)-(3). Then the point $\hat{\zeta}$ is a local minimum for the induced optimization problem (36), and hence it satisfies first and second order necessary conditions for this problem.*

Proof. Assume that $\hat{\zeta}$ is not a local minimum in problem (36). Then there exists a sequence of admissible points $\zeta^\nu = (t_0^\nu, t_1^\nu, x_0^\nu, \tau^\nu)$ in problem (36) such that $\zeta^\nu \to \hat{\zeta}$ for $\nu \to \infty$ and $\mathcal{F}_0(\zeta^\nu) < \mathcal{F}_0(\hat{\zeta})$. Take the corresponding sequence of admissible trajectories

$$\mathcal{T}^\nu = \{x(t; t_0^\nu, x_0^\nu, \tau^\nu), u(t; \tau^\nu) \mid t \in [t_0^\nu, t_1^\nu]\}$$

in problem (1)-(3). Then the conditions $t_0^\nu \to \hat{t}_0$, $t_1^\nu \to \hat{t}_1$, $x_0^\nu \to \hat{x}_0$, $\tau^\nu \to \hat{\tau}$ imply that

$$\int\limits_{\Delta^\nu \cap \hat{\Delta}} |u(t; \tau^\nu) - \hat{u}(t)|\, dt \to 0, \quad \max_{\Delta^\nu \cap \hat{\Delta}} |x(t; t_0^\nu, x_0^\nu, \tau^\nu) - \hat{x}(t)| \to 0,$$

where $\Delta^\nu = [t_0^\nu, t_1^\nu]$. Moreover,

$$\mathcal{J}(\mathcal{T}^\nu) = \mathcal{F}_0(\zeta^\nu) < \mathcal{F}_0(\hat{\zeta}) = \mathcal{J}(\hat{\mathcal{T}}).$$

It means that the trajectory $\hat{\mathcal{T}}$ is not a Pontryagin local minimum for the main problem (1)-(3). ■

We shall clarify a relationship between the second order conditions for the induced optimization problem (36) at the point $\hat{\zeta}$ and those in the main bang-bang control problem (1)-(3) for the trajectory $\hat{\mathcal{T}}$. We shall state that there is an one-to-one correspondence between Lagrange multipliers in these problems and an 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 will allow us to express the necessary and sufficient quadratic optimality conditions for bang-bang control, formulated in Theorems 3.1 and 3.2, in terms of the induced problem (36), and thus to establish the equivalence between our quadratic sufficient conditions and those due to Agrachev, Stefani, Zezza.

First, for the sake of convenience, we shall recall second order necessary and sufficient conditions for a smooth finite dimensional optimization problem with inequality and equality type constraints.

4.2. Second order necessary and sufficient conditions for a local minimum in a smooth optimization problem with inequality and equality constraints

Consider the problem in $\mathbb{R}^n$

$$f_0(x) \to \min; \qquad f_i(x) \le 0, \quad i = 1, \ldots, k; \quad g_j(x) = 0, \quad j = 1, \ldots, m, \qquad (37)$$

where $f_0, \ldots, f_k, g_1, \ldots, g_m$ are C^2–functions in $\mathbb{R}^n$. Let $\hat{x}$ be an admissible point in this problem. Define, at this point, the set of normalized vectors

$$\mu = (\alpha_0, \ldots, \alpha_k, \beta_1, \ldots, \beta_m)$$

of Lagrange multipliers

$$\begin{aligned}\Lambda_0 \quad = \quad & \{\mu \in \mathbb{R}^{k+1+m} \mid \alpha_i \geq 0,\ i = 0, \ldots, k; \quad \alpha_i f_i(\hat{x}) = 0,\ i = 1, \ldots, k; \\ & \sum_{i=0}^{k} \alpha_i + \sum_{j=1}^{m} |\beta_j| = 1; \quad L_x(\mu, \hat{x}) = 0\},\end{aligned}$$

where

$$L(\mu, x) = \sum_{i=0}^{k} \alpha_i f_i(x) + \sum_{j=1}^{m} \beta_j g_j(x)$$

is the Lagrange function. Define the set of indices of active inequality constraints at the point $\hat{x}$

$$I = \{i \in \{1, \ldots, k\} \mid f_i(\hat{x}) = 0\}$$

and the critical cone

$$\mathcal{K}_0 = \{\bar{x} \mid f_0'(\hat{x})\bar{x} \leq 0, \quad f_i'(\hat{x})\bar{x} \leq 0,\ i \in I, \quad g_j'(\hat{x})\bar{x} = 0,\ j = 1, \ldots, m\}.$$

THEOREM 4.2 *Let $\hat{x}$ be a local minimum in problem (37). Then, at this point, the set Λ_0 is nonempty and the following inequality holds*

$$\max_{\mu \in \Lambda_0} \langle L_{xx}(\mu, \hat{x})\bar{x}, \bar{x} \rangle \geq 0 \quad \forall\, \bar{x} \in \mathcal{K}_0.$$

THEOREM 4.3 *Let the set Λ_0 be nonempty at the point $\hat{x}$ and*

$$\max_{\mu \in \Lambda_0} \langle L_{xx}(\mu, \hat{x})\bar{x}, \bar{x} \rangle > 0 \quad \forall\, \bar{x} \in \mathcal{K}_0 \setminus \{0\}.$$

Then $\hat{x}$ is a local minimum in problem (37).

These conditions were obtained in Levitin, Milyutin, and Osmolovskii (1974), (1978); see also Ben-Tal, Zowe (1982).

4.3. The relationship between second-order conditions for the main and induced problem

Let $\hat{\mathcal{T}} = (\hat{x}(t), \hat{u}(t) \mid t \in [\hat{t}_0, \hat{t}_1])$ be an admissible trajectory in the main problem with the properties assumed in Section 4.1 and let $\hat{\zeta} = (\hat{t}_0, \hat{t}_1, \hat{x}_0, \hat{\tau})$ be the corresponding admissible point in the induced problem.

Lagrange multipliers. Let us define the set $\Lambda_0 \subset \mathbb{R}^{1+d(F)+d(K)}$ of the triples $\mu = (\alpha_0, \alpha, \beta)$ of normalized Lagrange multipliers at the point $\hat{\zeta}$ for the induced problem. The Lagrange function for the induced problem is

$$\begin{aligned} L(\mu,\zeta) &= L(\mu,t_0,t_1,x_0,\tau) = \alpha_0 J(t_0,x_0,t_1,x(t_1;t_0,x_0,\tau)) \\ &\quad +\alpha F(t_0,x_0,t_1,x(t_1;t_0,x_0,\tau)) + \beta K(t_0,x_0,t_1,x(t_1;t_0,x_0,\tau)) \\ &= l(\mu,t_0,x_0,t_1,x(t_1;t_0,x_0,\tau)), \end{aligned} \tag{38}$$

where $l = \alpha_0 J + \alpha F + \beta K$. By definition, Λ_0 is the set of multipliers $\mu = (\alpha_0, \alpha, \beta)$ such that

$$\alpha_0 \geq 0,\ \alpha \geq 0,\ \alpha_0 + |\alpha| + |\beta| = 1,\ \alpha F(\hat{p}) = 0,\ L_\zeta(\mu,\hat{\zeta}) = 0, \tag{39}$$

where $\hat{p} = (\hat{t}_0, \hat{x}_0, \hat{t}_1, \hat{x}_1)$, $\hat{x}_0 = \hat{x}(\hat{t}_0)$, $\hat{x}_1 = \hat{x}(\hat{t}_1) = x(\hat{t}_1; \hat{t}_0, \hat{x}_0, \hat{\tau})$. Now, let us define the corresponding set of normalized Lagrange multipliers for the trajectory $\hat{\mathcal{T}}$ in the main problem. Denote by Λ the set of multipliers $\lambda = (\alpha_0, \alpha, \beta, \psi, \psi_0) \in \mathcal{Y}$ such that

$$\begin{aligned} &\alpha_0 \geq 0, \quad \alpha \geq 0, \quad \alpha_0 + |\alpha| + |\beta| = 1, \quad \alpha F(\hat{p}) = 0, \\ &-\dot{\psi}(t) = \psi(t) f_x(t,\hat{x}(t),\hat{u}(t)), \quad -\dot{\psi}_0(t) = \psi(t) f_t(t,\hat{x}(t),\hat{u}(t)), \\ &\psi(\hat{t}_0) = -l_{x_0}(\mu,\hat{p}), \quad \psi(\hat{t}_1) = l_{x_1}(\mu,\hat{p}), \\ &\psi_0(\hat{t}_0) = -l_{t_0}(\mu,\hat{p}), \quad \psi_0(\hat{t}_1) = l_{t_1}(\mu,\hat{p}), \\ &\psi(t) f(t,\hat{x}(t),\hat{u}(t)) + \psi_0(t) = 0 \quad \forall t \in \hat{\Delta} \setminus \hat{\theta}, \end{aligned} \tag{40}$$

where $\hat{\Delta} = [\hat{t}_0, \hat{t}_1]$, $\hat{\theta} = \{\hat{\tau}_0, \ldots, \hat{\tau}_s\}$.

Proposition 4.1 *The projector*

$$\pi_0 : (\alpha_0, \alpha, \beta, \psi, \psi_0) \to (\alpha_0, \alpha, \beta) \tag{41}$$

maps one-to-one the set Λ *onto the set* Λ_0*.*

Let us define the inverse mapping. Take an arbitrary multiplier $\mu = (\alpha_0, \alpha, \beta) \in \Lambda_0$. This tuple defines the gradient $l_{x_1}(\mu, \hat{p})$, and hence the system

$$-\dot{\psi} = \psi f_x(t,\hat{x}(t),\hat{u}(t)), \quad \psi(t_1) = l_{x_1}(\mu,\hat{p}) \tag{42}$$

defines $\psi(t)$. Define $\psi_0(t)$ by the equality

$$\psi(t) f(t,\hat{x}(t),\hat{u}(t)) + \psi_0(t) = 0. \tag{43}$$

Proposition 4.2 *The inverse mapping*

$$\pi_0^{-1} : (\alpha_0, \alpha, \beta) \in \Lambda_0 \to (\alpha_0, \alpha, \beta, \psi, \psi_0) \in \Lambda \tag{44}$$

is defined by formulas (42) and (43).

We note that $M_0 \subset \Lambda$ holds, because the system of conditions(7)-(9) and (11) is equivalent to system (40). But it may happen that $M_0 \neq \Lambda$, since in the definition of Λ there is no requirement that its elements satisfy minimum condition (10). Let us denote $\Lambda_0^{MP} := \pi_0(M_0)$ (MP=minimum principle).

We shall say that multipliers $\mu = (\alpha_0, \alpha, \beta)$ and $\lambda = (\alpha_0, \alpha, \beta, \psi, \psi_0)$ *correspond to each other* if they have the same components α_0, α and β, i.e. $\pi_0(\alpha_0, \alpha, \beta, \psi, \psi_0) = (\alpha_0, \alpha, \beta)$.

Critical cones. We denote by $\mathcal{K}_0$ the critical cone at the point $\hat{\zeta}$ in the induced problem. Thus, $\mathcal{K}_0$ is the set of tuples $\bar{\zeta} = (\bar{t}_0, \bar{t}_1, \bar{x}_0, \bar{\tau})$ such that

$$\mathcal{F}_0'(\hat{\zeta})\bar{\zeta} \leq 0, \quad \mathcal{F}_i'(\hat{\zeta})\bar{\zeta} \leq 0,\ i \in I, \quad \mathcal{G}'(\hat{\zeta})\bar{\zeta} = 0, \tag{45}$$

where I is the set of indices of the inequality constraints active at the point $\hat{\zeta}$. Let $\mathcal{K}$ be the critical cone for the trajectory $\hat{\mathcal{T}}$ in the main problem, i.e. the set of all tuples $\bar{z} = (\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x}) \in \mathcal{Z}(\hat{\theta})$, satisfying conditions (19)-(21).

Proposition 4.3 *The operator*

$$\pi_1 : (\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x}) \to (\bar{t}_0, \bar{t}_1, \bar{x}_0, \bar{\tau})$$

defined by

$$\bar{\tau} = -\bar{\xi}, \quad \bar{x}_0 = \bar{x}(t_0) \tag{46}$$

is an one-to-one mapping of the critical cone $\mathcal{K}$ (for the trajectory $\hat{\mathcal{T}}$ in the main problem) onto the critical cone $\mathcal{K}_0$ (at the point $\hat{\zeta}$ in the induced problem).

We say that elements $\bar{\zeta} = (\bar{t}_0, \bar{t}_1, \bar{x}_0, \bar{\tau}) \in \mathcal{K}_0$ and $\bar{z} = (\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x}) \in \mathcal{K}$ *correspond to each other* if $\bar{\tau} = -\bar{\xi}$ and $\bar{x}_0 = \bar{x}(t_0)$, i.e. $\pi_1(\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x}) = (\bar{t}_0, \bar{t}_1, \bar{x}_0, \bar{\tau})$.

Now we shall give explicit formulas for the inverse mapping for π_1. Let $V(t)$ be $n \times n$ matrix-valued function ($n = d(x)$) which is absolutely continuous in $\hat{\Delta} = [\hat{t}_0, \hat{t}_1]$ and satisfies the system

$$\dot{V}(t) = f_x(t, \hat{x}(t), \hat{u}(t))V(t), \quad V(\hat{t}_0) = E, \tag{47}$$

where E is the identity matrix.

For each $i = 1, \ldots, s$ denote by $y^i(t)$ the n-dimensional vector function which is equal to zero in $[\hat{t}_0, \hat{\tau}_i)$ and in $[\hat{\tau}_i, \hat{t}_1]$ it is the solution to the initial value problem

$$\dot{y}^i = f_x(t, \hat{x}(t), \hat{u}(t))y^i, \quad y^i(\hat{\tau}_i) = -[\dot{\hat{x}}]^i. \tag{48}$$

Hence y^i is a piecewise continuous function with one jump $[y^i]^i = -[\dot{\hat{x}}]^i$ at the point $\hat{\tau}_i$.

PROPOSITION 4.4 *The inverse mapping*

$$\pi_1^{-1} : (\bar{t}_0, \bar{t}_1, \bar{x}_0, \bar{\tau}) \in \mathcal{K}_0 \to (\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x}) \in \mathcal{K}$$

is given by the formulas

$$\bar{\xi} = -\bar{\tau}, \qquad \bar{x}(t) = V(t)\left(\bar{x}_0 - \dot{\hat{x}}(\hat{t}_0)\bar{t}_0\right) + \sum_{i=1}^{s} y^i(t)\bar{\tau}_i. \tag{49}$$

Quadratic forms. For $\mu \in \Lambda_0$ the quadratic form, of the induced optimization problem, is equal to $\langle L_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta}\rangle$.

The main result of this paper is the following:

THEOREM 4.4 *Let the Lagrange multipliers $\mu = (\alpha_0, \alpha, \beta) \in \Lambda_0^{MP}$ and $\lambda = (\alpha_0, \alpha, \beta, \psi, \psi_0) \in M_0$ correspond to each other, i.e. $\pi_0\lambda = \mu$, and let the elements of the critical cones $\bar{\zeta} = (\bar{t}_0, \bar{t}_1, \bar{x}_0, \bar{\tau}) \in \mathcal{K}_0$ and $\bar{z} = (\bar{t}_0, \bar{t}_1, \bar{\xi}, \bar{x}) \in \mathcal{K}$ correspond to each other, i.e. $\pi_1\bar{z} = \bar{\zeta}$. Then the quadratic forms in the main and induced problems take equal values: $\langle L_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta}\rangle = \Omega(\lambda, \bar{z})$. Consequently,*

$$\max_{\mu \in \Lambda_0^{MP}} \langle L_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta}\rangle = \max_{\lambda \in M_0} \Omega(\lambda, \bar{z})$$

for each pair of elements of the critical cones $\bar{\zeta} \in \mathcal{K}_0$ and $\bar{z} \in \mathcal{K}$ such that $\pi_1\bar{z} = \bar{\zeta}$.

Theorems 3.1, 4.4, and Proposition 4.3 imply the following second order necessary optimality condition for the main problem.

THEOREM 4.5 *If the trajectory $\mathcal{T}$ affords a Pontryagin minimum in the main problem, then the following Condition $\mathcal{A}_0$ holds: the set M_0 is nonempty and*

$$\max_{\mu \in \Lambda_0^{MP}} \langle L_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta}\rangle \geq 0 \quad \textit{for all} \quad \bar{\zeta} \in \mathcal{K}_0.$$

Theorems 3.2, 4.4, and Proposition 4.3 imply the following second order sufficient optimality condition for the main control problem.

THEOREM 4.6 *Let the following Condition $\mathcal{B}_0$ be fulfilled for the trajectory $\mathcal{T}$ in the main problem:*

- *(a) $\hat{u}(t)$ is a regular bang-bang control (hence, the set M_0 is nonempty and condition (17) holds for some $\lambda \in M_0$),*
- *(b) there exists $\lambda \in M_0$ such that $D^k(H) > 0, k = 1, \ldots, s$,*
- *(c) $\max_{\mu \in \Lambda_0^{MP}} \langle L_{\zeta\zeta}(\mu, \hat{\zeta})\bar{\zeta}, \bar{\zeta}\rangle > 0$ for all $\bar{\zeta} \in \mathcal{K}_0 \setminus \{0\}$.*

Then $\mathcal{T}$ is a strict strong minimum in the main problem.

Theorem 4.6 is a generalization of sufficient optimality conditions for bang-bang controls obtained in Agrachev et al. (2002).

Due to space limitations a detailed proof of the preceding theorems will be given in the second part of this paper. Let us point out that the proof reveals the useful fact that all elements of the Hessian $L_{\zeta\zeta}(\mu, \hat{\zeta})$ can be computed explicitly on the basis of the transition matrix $V(\hat{t}_1)$ in (47) and of the *first order* variations y^i defined by (48).

References

Agrachev, A.A., Stefani, G. and Zezza, P.L. (2002) Strong optimality for a bang–bang trajectory. *SIAM J. Control and Optimization* **41**, 991-1014.

Ben-Tal, A. and Zowe, J. (1982) A unified theory of first and second order conditions for extremum problems in topological vector spaces. *Mathematical Programming Study* **19**, 39–76.

Ben-Tal, A. and Zowe, J. (1982) Necessary and sufficient conditions optimality conditions for a class of nonsmooth minimization problems. *Mathematical Programming* **24**, 70-91.

Felgenhauer, U. (2003) On stability of bang–bang type controls. *SIAM J. Control and Optimization* **41**, 1843–1867.

Felgenhauer, U. (2005) Optimality properties of controls with bang-bang components in problems with semilinear state equation. *Control and Cybernetics* **34**, in this issue.

Hestens (1966) *Calculus of Variations and Optimal Control*. Kohn Wiley, New York.

Kim, J.-H.R., Lippi, G.L. and Maurer, H. (2004) Minimizing the transition time in lasers by optimal control methods. Single mode semiconductor lasers with homogeneous transverse profile. *Physica-D – Nonlinear phenomena* **191**, 238–260.

Kim, J.-H.R. and Maurer, H. (2004) Sensitivity analysis of optimal control problems with bang-bang controls. *Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Dec. 9-12, 2003*, IEEE Control Society, 3281–3286.

Levitin, E.S., Milyutin, A.A. and Osmolovskii, N.P. (1974) Conditions for a local minimum in a problem with constraints. In: *Mathematical economics and functional analysis*. Nauka, Moscow, 130-202.

Levitin, E.S., Milyutin, A.A. and Osmolovskii, N.P. (1978) Conditions of high order for a local minimum in problems with constraints. *Russian Math. Surveys* **33** (6), 97–168.

Maurer, H., Büskens, C., Kim, J.-H.R. and Kaya, Y. (2004) Optimization methods for the verification of second-order sufficient conditions for bang-bang controls. Submitted to *Optimal Control Applications and Methods*.

MAURER, H. and OSMOLOVSKII, N.P. (2004) Second order sufficient conditions for time optimal bang–bang control problems. *SIAM J. Control and Optimization* **42**, 2239–2263.

MAURER, H. and OSMOLOVSKII, N.P. (2003) Second order optimality conditions for bang–bang control problems. *Control and Cybernetics* **32** (3), 555-584.

MILYUTIN, A.A. and OSMOLOVSKII, N.P. (1998) *Calculus of Variations and Optimal Control*. Translations of Mathematical Monographs, **180**, American Mathematical Society, Providence.

OSMOLOVSKII, N.P. (1988a) High-order necessary and sufficient conditions for Pontryagin and bounded-strong minima in the optimal control problems. *Dokl. Akad. Nauk SSSR, Ser. Cybernetics and Control Theory* **303**, 1052–1056, English transl., *Sov. Phys. Dokl.* **33** (12), 1988, 883–885.

OSMOLOVSKII, N.P. (1988b) *Theory of higher order conditions in optimal control*. Doctor of Science Thesis (in Russian), Moscow.

OSMOLOVSKII, N.P. (1995) Quadratic conditions for nonsingular extremals in optimal control (A theoretical treatment). *Russian J. of Mathematical Physics* **2**, 487–516.

OSMOLOVSKII, N.P. (2000) Second order conditions for broken extremal. In: A. Ioffe, S. Reich and I. Shafir, eds., *Calculus of variations and optimal control* (Technion 1998), Chapman and Hall/CRC, Boca Raton, Florida, 198-216.

OSMOLOVSKII, N.P. (2002) Second-order sufficient conditions for an extremum in optimal control. *Control and Cybernetics* **31** (3), 803-831.

OSMOLOVSKII, N.P. (2004) Quadratic optimality conditions for broken extremals in the general problem of calculus of variations. *Journal of Math. Science* **123** (3), 3987-4122.

OSMOLOVSKII, N.P. and LEMPIO, F. (2000) Jacobi–type conditions and Riccati equation for broken extremal. *Journal of Math. Science* **100** (5), 2572–2592.

OSMOLOVSKII, N.P. and LEMPIO, F. (2002) Transformation of quadratic forms to perfect squares for broken extremals. *Journal of Set Valued Analysis* **10**, 209-232.

POGGIOLONI, L. and STEFANI, G. (2003) State-local optimality of a bang-bang trajectory: a Hamiltonian approach. Report.

PONTRYAGIN, L.S., BOLTYANSKII, V.G., GAMKRELIDZE, R.V. and MISHCHENKO, E.F. (1961) *The Mathematical Theory of Optimal Processes*. Fizmatgiz, Moscow; English translation: Pergamon Press, New York, 1964.