

On Second-Order Necessary Conditions for Broken Extremals

Nikolai P. Osmolovskii

Received: 9 May 2014 / Accepted: 7 June 2014 / Published online: 8 July 2014
© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract We consider optimal control problems with initial–final state equality and inequality constraints and running mixed state-control equality constraints given by smooth functions. The mixed constraints satisfy the regularity assumption of linear independence of gradients with respect to the control. We present simple proofs of second-order necessary conditions for (extended) weak minimum for extremals with discontinuous controls in these problems.

Keywords Weak minimum · Euler equation · Pontryagin minimum principle · Critical cone · Quadratic form

Mathematics Subject Classification 49K15

1 Introduction

There is an extensive literature on second-order optimality conditions in optimal control. The classical no-gap second-order conditions have the following formulation: a necessary condition consists in that a certain quadratic form must be nonnegative on the so-called critical cone; a sufficient condition requires that the same quadratic form is positive definite on the critical cone. Note that, in simple cases, the quadratic

N. P. Osmolovskii
University of Technology and Humanities, ul. Malczewskiego 20A, 26-600 Radom, Poland

N. P. Osmolovskii
Systems Research Institute, ul. Newelska 6, 01-447 Warszawa, Poland

N. P. Osmolovskii (✉)
Moscow State University of Civil Engineering, Jaroslavscoe sh. 26, 129337 Moscow, Russia
e-mail: osmolovski@uph.edu.pl

form represents the second variation of the Lagrange function of the problem, and the critical cone consists of all directions tangent to the constraints. A natural question arises: how to check positive (semi-) definiteness of the quadratic form on the critical cone? The answer to this question can be typically given in terms of the Jacobi-type conditions or solutions to the corresponding Riccati-type (matrix) equation. But certainly, each new quadratic form, obtained for a new class of problems and/or a new type of optimal controls, requires a new answer to this question.

Historically, the study of second-order conditions in optimal control started from the degenerate case, where the control was singular on the whole time interval (singular controls appear, e.g. in problems, linear in control). In this case, the Pontryagin maximum principle is much less informative than in the general case. Necessary conditions for singular controls were obtained by Kelley, Kopp, Moyer, Gabasov, Kirillova, Goh, Krener, Gamkrelidze, Agrachev, Stefani and many others. Most of these conditions were not full in the sense that they did not possess natural strengthening to sufficient ones. It took some time to understand which quadratic form corresponds to singular controls. This understanding has been fully achieved in works of Milyutin and Dmitruk, who obtained (in the 70's–90's) no-gap necessary and sufficient second-order conditions for singular controls in a broad class of problems (see, e.g. [1] and [2]).

From around the middle of 70's, research efforts were also focused on studying the general, nondegenerate case. Fundamental results, for this case, were obtained by Milyutin and Osmolovskii (for problems with regular mixed state-control constraints), Maurer, Zeidan, Pickenhain, Malanowski, Bonnans (especially for problems with state constraints and lately, for problems with integral equations) and others. Most of the mentioned authors were interested in obtaining both necessary and sufficient no-gap conditions. Recently, a number of new results on second-order conditions were obtained by Frankowska and her collaborators. These results refer to problems with nonsmooth data, but, unfortunately, they have rather nontraditional and complicated formulation.

Note that the most part of the known second-order conditions in optimal control (with a few exceptions which will be discussed below) do not take into account discontinuities of optimal control. Moreover, second-order sufficient conditions, which guarantee the quadratic growth of the cost, usually presuppose the continuity of the control. Clearly, this property cannot be considered as natural for optimal control problems. Therefore, it was important to find a quadratic form which corresponds to discontinuous controls. Such a form was not known even in the classical calculus of variations, where an extremal with discontinuous control is simply a broken extremal. Probably, the first researchers, who obtained (independently) the quadratic form for a broken extremal, were Henrion [3] and Shnol' [4]. Shnol' did this for the simplest problem of the calculus of variations while Henrion—for a problem with an autonomous control system, without control constraints. The quadratic forms obtained by Henrion and Shnol' have different definitions but can be transformed to each other. Both quadratic forms are particular cases of the quadratic form obtained later by the author in [5] for the general problem of the calculus of variations (formulations of results [5] were earlier published in [6], Part 2). The latter is an optimal control problem (in Mayer form) of ordinary differential equations, considered on a fixed interval

of time, subject to boundary constraints of inequality and equality type and running mixed state-control constraints of equality type. The results [5] were extended in [7] and [8] to problems with regular mixed state-control constraints of both equality and inequality type (much earlier statements of results [7] and [8] were presented in [9]). Finally, in [10], the results [5] were extended to problems on a variable interval of time (statements of these extended results were first published in [6], Part 2). We will extensively use methods and results of [10] in the present paper.

It is worth to note that the results of papers [5] and [10] also provided new conditions of the second order, both necessary and sufficient, for bang-bang controls, in optimal control problems, linear in control, with a convex polyhedron as a control constraint. These conditions were first published in the book of Milyutin and Osmolovskii [6] and later developed in the book of Osmolovskii and Maurer [11]. In particular, book [11] contains criteria for positive (semi-) definiteness of the quadratic form, corresponding to a broken extremal, on the critical subspace.

Problems with constraints on the time and state variables, given in internal points of time interval, were considered by Hull in [12] (Chap. 16) and later by Dmitruk and Kaganovich in [13]. Such constraints are called internal or intermediate. Extremals with jumps of the control arise naturally in problems with such constraints. Hull considered the problem with free final time, equality constraint on the final time and state, and one internal point equality constraint. Dmitruk and Kaganovich considered the general problem of the calculus of variations on a nonfixed time interval with finitely many intermediate constraints of both equality and inequality type. First- and second-order necessary optimality conditions were obtained in both problems for extremals with jumps of the control. An important feature of these conditions is that the dual variable can have a jump at each intermediate point where the constraint is given, and the direction of the jump is determined by the gradients of intermediate constraints. It is very likely that the conditions obtained in these two papers can be arranged so that their relation with conditions obtained in [5], [6] and [10] becomes apparent.

The aim of the present publication is to show that rather difficult and lengthy proofs of necessary optimality conditions for extremals with discontinuous controls, given in [5] (and later in [11]), can be reduced to much shorter and simple proofs. To this end, we will study the relationship between necessary second-order conditions for a weak local minimum in the general problem of the calculus of variations on a fixed time interval and necessary second-order conditions for an extended weak local minimum in the general problem of the calculus of variations on a variable time interval. In contrast to the weak minimum, the extended weak minimum is determined by not only small (in absolute value) variations of the control variable but also with the help of small variations of discontinuity points of optimal control. The corresponding necessary conditions, which we obtain, take these variations into account. The relationship between two types of optimality conditions is based on a simple change of time variable. Just this change will allow us to obtain a new, relatively simple proof of necessary second-order conditions for an extended weak local minimum for extremals with discontinuous controls. Thus, we show a short road to the quadratic form corresponding to a broken extremal.

The paper is organized as follows. In Sect. 2, we formulate the second-order optimality conditions, derived in book [6] (Part 2) for broken extremals in the simplest problem of the calculus of variations. In Sect. 3, we consider the general problem of the calculus of variations on a fixed time interval and derive the (well-known) second-order necessary conditions for a weak local minimum in this problem, using Dubovitskii–Milyutin method of critical variations [14, 15]. In Sect. 4, for the general problem of the calculus of variations on a variable time interval, we derive the second-order necessary conditions for an extended weak local minimum. As it was said before, in the proof, we use a simple change of time variable and also the necessary conditions of a weak minimum from Sect. 3.

2 The Simplest Problem of the Calculus of Variations

Let us recall conditions [6] for a weak and an extended weak minimum in the simplest case. Consider the simplest problem of the calculus of variations:

$$\min \mathcal{J}(x) := \int_0^1 F(t, x(t), \dot{x}(t)) dt, \quad x(0) = a, \quad x(1) = b.$$

Here $x(t)$ is Lipschitz continuous, i.e. $x(\cdot)$ belongs to the Sobolev space $W^{1,\infty}$ endowed with the norm $\|x\|_{1,\infty} := |x(t_0)| + \text{ess sup}_{t \in [0,1]} |\dot{x}(t)|$. As is customary, for a measurable function $v(t) : [0, 1] \mapsto \mathbb{R}$, we put

$$\text{ess sup}_{t \in [0,1]} v(t) := \inf\{\rho \in \mathbb{R} : \text{meas}\{t \in [0, 1] : v(t) > \rho\} = 0\},$$

where “meas” means the Lebesgue measure. A local minimum in the space $W^{1,\infty}$ is exactly a *weak minimum*.

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

$$\frac{d}{dt} F_{\dot{x}}(t, x^0(t), \dot{x}^0(t)) = F_x(t, x^0(t), \dot{x}^0(t)).$$

Set $\dot{x} = u$, $w = (x, u)$. We call u the *control*. Set

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

Let

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

where $W^{1,2}$ is the Sobolev space of absolutely continuous functions with square integrable first derivative. Define a quadratic form in the space \mathcal{W}_2 :

$$\begin{aligned} \Omega^{sp}(w) &:= \int_0^1 \langle F_{ww}(t, w^0(t))w(t), w(t) \rangle dt \\ &= \int_0^1 (\langle F_{xx}x(t), x(t) \rangle + 2\langle F_{xu}u(t), x(t) \rangle + \langle F_{uu}u(t), u(t) \rangle) dt, \end{aligned}$$

where $F_{xx} = F_{xx}(t, w^0(t))$, etc., and the superscript “sp” means “the simplest problem”. Set

$$\mathcal{K}^{sp} := \{w \in \mathcal{W}_2 : \dot{x}(t) = u(t) \text{ a.e., } x(0) = x(1) = 0\}.$$

The following theorem is well-known.

Theorem 2.1 (a) *If x^0 is a weak minimum, then $\Omega^{sp}(w) \geq 0$ on \mathcal{K}^{sp} .*
 (b) *If $\Omega^{sp}(w)$ is positive definite on \mathcal{K}^{sp} , then x^0 is a (strict) weak minimum.*

Now, assume that the control $u^0(t)$ is *piecewise continuous* with one discontinuity point $t_* \in]0, 1[$. Moreover, assume that $u^0(t)$ is Lipschitz continuous on each of the two intervals $]0, t_*[$ and $]t_*, 1[$. Hence $x^0(t)$ is a *broken extremal* with a break at t_* . Which quadratic form corresponds to a broken extremal? The answer to this question is given, e.g. in the book [6], Part 2. Let us change the definition of a weak local minimum as follows. Set $\Theta := \{t_*\}$ and define the notion of a Θ -weak minimum, or extended weak minimum. Assuming that the control $u^0(t)$ is left continuous, denote by $\text{cl } u^0(\cdot)$ the closure of the graph of $u^0(t)$. Denote by V a neighbourhood of the compact set $\text{cl } u^0(\cdot)$ (see Fig. 1, where the neighbourhoods V_1^0 and V_2^0 serve for the definition of a weak minimum, while the neighbourhoods V_-^* and V_+^* are used in the definition of a Θ -weak minimum).

Definition 2.1 x^0 is a point of a Θ -weak minimum (or an *extended weak minimum*) iff there exists a neighbourhood V of the compact set $\text{cl } u^0(\cdot)$ such that $\mathcal{J}(x) \geq \mathcal{J}(x^0)$ for all $x \in W^{1,\infty}$ such that $x(0) = a, x(1) = b$, and $u(t) \in V$ a.e., where $u(t) = \dot{x}(t)$.

Let us note that a similar type of minimum (also called “extended weak”) was introduced by Dmitruk and Kaganovich in [13]. It is easy to compare these two types of minimum: the minimum, defined in Definition 2.1, always implies the minimum, defined in [13], but not vice versa. Moreover, we can claim that all necessary optimality conditions for an extended weak minimum, formulated in this paper, follows from the minimum, specified in [13]. We will omit the proof of this assertion.

Recall the Weierstrass–Erdmann necessary conditions for broken extremal:

(i) $\psi(t) := -F_u(t, w^0(t))$ is continuous at t_* , i.e. $[\psi] = 0$, where

$$[\psi] := \psi(t_*+) - \psi(t_*-) = \psi^+ - \psi^-$$

denotes the *jump* of ψ at t_* ;

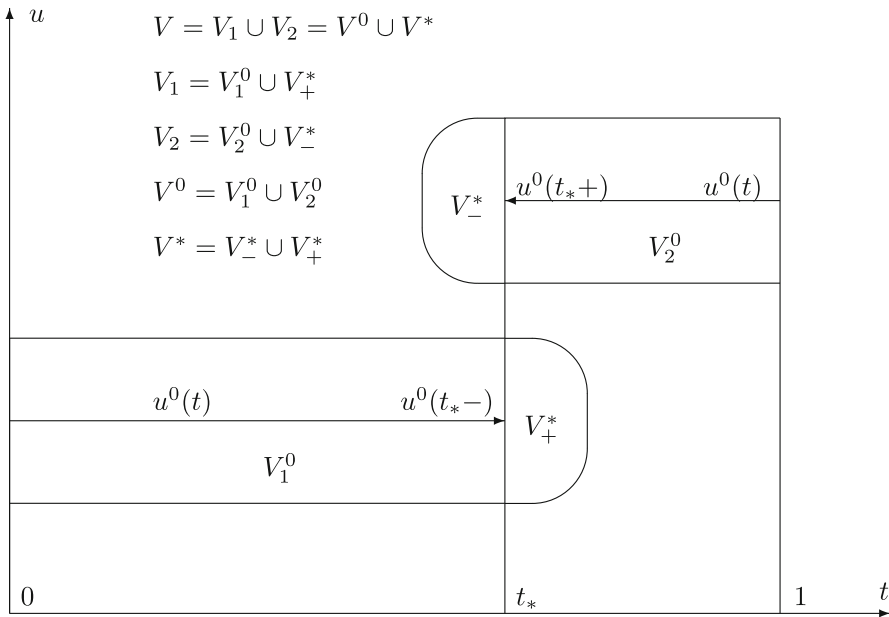


Fig. 1 Neighbourhoods of the graph of discontinuous control $u^0(t)$

(ii) $H(t) := \psi(t)u^0(t) + F(t, w^0(t))$ is continuous at t_* , i.e. $[H] = 0$.

We add one more necessary condition [6]:

(iii) $D(H) \geq 0$,

where $D(H)$ is equal to minus the derivative of the function

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

at t_* (the existence of this derivative is proved). One can show (see [6]) that

$$D(H) = \dot{\psi}^+ \dot{x}^{0-} - \dot{\psi}^- \dot{x}^{0+} + [\dot{\psi}_0], \tag{1}$$

where $\psi_0(t) := -H(t)$.

Denote by $P_\Theta W^{1,2}$ the Hilbert space of piecewise continuous functions $x(t)$, absolutely continuous on each of the two intervals $[0, t_*[$ and $]t_*, 1]$, and such that their first derivative is square integrable. Any $x \in P_\Theta W^{1,2}$ can have a nonzero jump $[x] := x(t_*+) - x(t_*-)$ at the point t_* . Let ξ be a numerical parameter. Denote by $Z_2(\Theta)$ the space of triples $z = (\xi, x, u) = (\xi, w)$ such that

$$\xi \in \mathbb{R}, \quad x(\cdot) \in P_\Theta W^{1,2}, \quad u(\cdot) \in L^2.$$

Thus,

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

In the space $Z_2(\Theta)$, define a quadratic form

$$\Omega_{\Theta}^{SP}(z) := D(H)\xi^2 + 2[F_x]x_{av}\xi + \int_0^1 \langle F_{ww}(t, w^0(t))w(t), w(t) \rangle dt,$$

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

$$x_{av} := \frac{1}{2} \left(x(t_*-) + x(t_*+) \right).$$

Set

$$\mathcal{K}_{\Theta}^{SP} := \{z \in Z_2(\Theta) : \dot{x}(t) = u(t) \text{ a.e., } [x] = [u^0]\xi, \quad x(0) = x(1) = 0\}.$$

Theorem 2.2 (a) *If x^0 is a Θ -weak minimum, then $\Omega_{\Theta}^{SP}(z) \geq 0$ on $\mathcal{K}_{\Theta}^{SP}$.*
 (b) *If $\Omega_{\Theta}^{SP}(z)$ is positive definite on $\mathcal{K}_{\Theta}^{SP}$, then x^0 is a (strict) Θ -weak minimum.*

A detailed proof of this theorem (based on the so-called “method of deciphering”) is given in [6]. It is rather long and technical. But it turned out that there was a relatively simple way to prove the necessary condition (a) of this theorem and thus to come to the quadratic form which corresponds to a broken extremal. This way will be shown in the present paper for the general problem of the calculus of variations. To this end we will first recall necessary second-order conditions for a weak minimum.

3 Necessary Second-Order Condition for a Weak Local Minimum in the General Problem of the Calculus of Variations on a Fixed Time Interval

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

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

$$\min \mathcal{J}(w) := J(x(t_0), x(t_f)) \tag{2}$$

subject to the constraints

$$F(x(t_0), x(t_f)) \leq 0, \quad K(x(t_0), x(t_f)) = 0, \tag{3}$$

$$\dot{x}(t) = f(t, x(t), u(t)), \tag{4}$$

$$g(t, x(t), u(t)) = 0, \tag{5}$$

$$(x(t_0), x(t_f)) \in \mathcal{P}, \quad (t, x(t), u(t)) \in \mathcal{Q}, \tag{6}$$

where \mathcal{P} and \mathcal{Q} are open sets, and x, u, F, K, f and g are vector-functions. We call (2)–(5) the Problem A.

We assume that the given functions J, F and K are twice continuously differentiable on \mathcal{P} , and the given functions f and g are twice continuously differentiable on \mathcal{Q} . It is also assumed that the gradients with respect to the control $g_{iu}(t, x, u), i = 1, \dots, d(g)$ are linearly independent at each point $(t, x, u) \in \mathcal{Q}$ such that $g(t, x, u) = 0$ (the *regularity assumption* for the equality constraint $g(t, x, u) = 0$). Here g_i are the components of the vector function g , and $d(g)$ is the dimension of this function.

The Problem A is considered in the space of pairs $w = (x, u)$ such that the *state variable* $x(t)$ is an absolutely continuous $d(x)$ -dimensional function, and the *control* $u(t)$ is a bounded measurable $d(u)$ -dimensional function on the interval $[t_0, t_f]$. Hence, the problem is considered in the space

$$\mathcal{W} := W^{1,1}([t_0, t_f], \mathbb{R}^{d(x)}) \times L^\infty([t_0, t_f], \mathbb{R}^{d(u)}).$$

Define a norm in this space as the sum of the norms:

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

We say that $w \in \mathcal{W}$ is an *admissible pair* iff it satisfies all constraints of the problem, and moreover, the graph of the function $w(t)$ lies in the set \mathcal{Q} with a certain ‘margin’. The latter means that for the pair $w = (x, u)$ there exists a compact set $\mathcal{C} \subset \mathcal{Q}$ such that $(t, x(t), u(t)) \in \mathcal{C}$ a.e. on $[t_0, t_1]$. Let $w^0 = (x^0, u^0) \in \mathcal{W}$ be a fixed admissible pair. We say that w^0 is a *weak local minimum* iff it is a local minimum in the space \mathcal{W} , i.e. there exists $\varepsilon > 0$ such that $\mathcal{J}(w) \geq \mathcal{J}(w^0)$ for all admissible pairs $w \in \mathcal{W}$ satisfying the condition $\|w - w^0\| \leq \varepsilon$.

3.2 Necessary Condition for a Weak Local Minimum

We introduce the *Pontryagin function*

$$H(t, x, u, \psi) := \psi f(t, x, u) \tag{7}$$

and the *augmented Pontryagin function*

$$\bar{H}(t, x, u, \psi, v) := H(t, x, u, \psi) + v g(t, x, u), \tag{8}$$

where ψ and v are row vectors of the dimensions $d(x)$ and $d(g)$, respectively. Note that the augmented Pontryagin function plays the role of the Lagrange function in the ‘local problem’ of minimization (with respect to the control u) of the function $H(t, x^0(t), u, \psi(t))$ subject to the ‘local constraint’ $g(t, x^0(t), u) = 0$. For brevity we set

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

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

$$L(p, \alpha_0, \alpha, \beta) := \alpha_0 J(p) + \alpha F(p) + \beta K(p), \tag{9}$$

where

$$\alpha_0 \in \mathbb{R}, \quad \alpha \in (\mathbb{R}^{d(F)})^*, \quad \beta \in (\mathbb{R}^{d(K)})^*.$$

Introduce a tuple of Lagrange multipliers

$$\lambda = (\alpha_0, \alpha, \beta, \psi(\cdot), \nu(\cdot)) \tag{10}$$

such that the function $\psi(\cdot) : [t_0, t_f] \mapsto (\mathbb{R}^{d(x)})^*$ is absolutely continuous, and the function $\nu(\cdot) : [t_0, t_f] \mapsto (\mathbb{R}^{d(g)})^*$ is measurable and essentially bounded.

Denote by Λ_0 the set of all tuples λ satisfying the following conditions at the point w^0 :

$$\begin{aligned} \alpha_0 \geq 0, \quad \alpha \geq 0, \quad \alpha F(p^0) = 0, \quad \alpha_0 + \sum_{i=1}^{d(F)} \alpha_i + \sum_{j=1}^{d(K)} |\beta_j| = 1, \\ \dot{\psi} = -\bar{H}_x, \quad \psi(t_0) = -L_{x_0}, \quad \psi(t_f) = L_{x_f}, \quad \bar{H}_u = 0, \end{aligned} \tag{11}$$

where $p^0 = (x^0(t_0), x^0(t_f))$, the derivatives L_{x_0} and L_{x_f} are at $(p^0, \alpha_0, \alpha, \beta)$ and the derivatives \bar{H}_x, \bar{H}_u are at $(t, x^0(t), u^0(t), \psi(t), \nu(t))$, $t \in [t_0, t_f]$. By α_i and β_j we denote the components of row vectors α and β , respectively.

The following well-known first-order necessary condition holds: if w^0 is a weak local minimum, then the set Λ_0 is nonempty. This condition is called the *local minimum principle* (or the *Euler–Lagrange equation*). From the regularity assumption for the constraint $g = 0$ and definition (11) it easily follows that Λ_0 is a finite-dimensional compact set, and the projector

$$(\alpha_0, \alpha, \beta, \psi(\cdot), \nu(\cdot)) \mapsto (\alpha_0, \alpha, \beta)$$

is injective on Λ_0 .

We will obtain the necessary condition of the first order: $\Lambda_0 \neq \emptyset$, together with a necessary condition of the second order, using the method of Dubovitskii and Milyutin[14, 15]. But of course, the condition of nonemptiness of Λ_0 can also be obtained by the direct application of abstract Lagrange multipliers rule to problem A (cf, e.g. [16]). Moreover, the Pontryagin function H , the augmented Pontryagin function \bar{H} , the endpoint Lagrange function L , the adjoint equation $-\dot{\psi} = \bar{H}_x$, the stationarity condition of the augmented Pontryagin function with respect to the control: $\bar{H}_u = 0$ —all these functions and conditions arise naturally from the stationarity condition (at w^0) of the Lagrange function \mathcal{L} of problem A .

Now let us formulate the second-order necessary condition at the point w^0 . Set

$$\mathcal{W}_2 := W^{1,2}([t_0, t_f], \mathbb{R}^{d(x)}) \times L^2([t_0, t_f], \mathbb{R}^{d(u)}),$$

where $W^{1,2}([t_0, t_f], \mathbb{R}^{d(x)})$ is the space of absolutely continuous functions such that their first derivative is square integrable, and $L^2([t_0, t_f], \mathbb{R}^{d(u)})$ is the space of square integrable functions. Hence \mathcal{W}_2 is a Hilbert space with a scalar product

$$(w, \tilde{w}) := \langle x(t_0), \tilde{x}(t_0) \rangle + \int_{t_0}^{t_f} \langle \dot{x}, \dot{\tilde{x}} \rangle dt + \int_{t_0}^{t_f} \langle u, \tilde{u} \rangle dt.$$

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

$$\begin{aligned} J'(p^0)\bar{p} &\leq 0, \quad F'_i(p^0)\bar{p} \leq 0 \quad \forall i \in I_F(p^0), \quad K'(p^0)\bar{p} = 0, \\ \dot{\tilde{x}}(t) &= f_w(t, w^0(t))\bar{w}(t), \quad \text{for a.a. } t \in [t_0, t_f], \\ g_w(t, w^0(t))\bar{w}(t) &= 0, \quad \text{for a.a. } t \in [t_0, t_f], \end{aligned} \tag{12}$$

where $I_F(p^0) := \{i : F_i(p^0) = 0\}$ is the set of active indices at the point p^0 , and $\bar{p} := (\bar{x}(t_0), \bar{x}(t_f))$. It is obvious that \mathcal{K} is a convex cone in the Hilbert space \mathcal{W}_2 . We call it the *critical cone*.

Let us introduce a quadratic form on the space \mathcal{W}_2 . For $\lambda \in \Lambda_0$ and $\bar{w} = (\bar{x}, \bar{u}) \in \mathcal{W}_2$, we set

$$\Omega(\lambda, \bar{w}) := \langle L_{pp}\bar{p}, \bar{p} \rangle + \int_{t_0}^{t_f} \langle \bar{H}_{ww}\bar{w}(t), \bar{w}(t) \rangle dt, \tag{13}$$

where $L_{pp} = L_{pp}(p^0, \alpha_0, \alpha, \beta)$ and $\bar{H}_{ww} = \bar{H}_{ww}(t, x^0(t), u^0(t), \psi(t), v(t))$.

Theorem 3.1 *If the trajectory \mathcal{T} yields a weak minimum, then the following Condition A holds: the set Λ_0 is nonempty and*

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

3.3 Proof of the Necessary Condition for a Weak Local Minimum

Here we present a short proof of Theorem 3.1 omitting some details. In this proof, we will use the Dubovitskii–Milyutin method of critical variations, cf. [14, 15]. Let w^0 be a weak local minimum. Without loss of generality we assume that $J(p^0) = 0$, and $F_i(p^0) = 0$ for all $i = 1, \dots, d(F)$.

Denote by $L^1([t_0, t_f], \mathbb{R}^{d(x)})$ the space of integrable functions. Consider the operator

$$G : w = (x, u) \in \mathcal{W} \mapsto \left(f(w) - \dot{x}, g(t, w), K(x(t_0), x(t_f)) \right) \in \mathcal{Y}, \tag{14}$$

where

$$\mathcal{Y} := L^1([t_0, t_f], \mathbb{R}^{d(x)}) \times L^\infty([t_0, t_f], \mathbb{R}^{d(g)}) \times \mathbb{R}^{d(K)}.$$

This operator is Fréchet continuously differentiable in a neighbourhood of the point w^0 , and its derivative $G'(w^0)$ at w^0 is a linear operator defined as follows:

$$w = (x, u) \in \mathcal{W} \mapsto (f_w(t, w^0)w - \dot{x}, g_w(t, w^0)w, K'(p^0)p) \in \mathcal{Y}. \tag{15}$$

The derivative $G'(w^0)$ has a closed image (see, e.g. [16]), since the linear operator, taking an element $w \in \mathcal{W}$ to an element

$$(f_w(t, w^0)w - \dot{x}, g_w(t, w^0)w) \in L^1([t_0, t_f], \mathbb{R}^{d(x)}) \times L^\infty([t_0, t_f], \mathbb{R}^{d(g)}),$$

is surjective (it easily follows from the regularity assumption for the constraint $g(t, w) = 0$), and the operator $w \in \mathcal{W} \mapsto K'(p^0)p \in \mathbb{R}^{d(K)}$ is finite dimensional. Consider two possible cases: $G'(w^0)\mathcal{W} \neq \mathcal{Y}$ and $G'(w^0)\mathcal{W} = \mathcal{Y}$.

(a) In the first case, the image $G'(w^0)\mathcal{W}$ is a closed subspace of \mathcal{Y} , not equal to \mathcal{Y} . Therefore, there exists a nonzero linear functional $l(w)$ vanishing on this image. The latter means that there exists a nonzero triple

$$\psi \in L^\infty([t_0, t_f], \mathbb{R}^{d(x)*}), \quad \nu \in (L^\infty([t_0, t_f], \mathbb{R}^{d(g)*}))^*, \quad \beta \in \mathbb{R}^{d(K)*}$$

such that

$$\int_{t_0}^{t_f} \psi(f_w(t, w^0)w - \dot{x}) dt + \langle \nu, g_w(t, w^0)w \rangle + \beta K'(p^0)p = 0 \quad \forall w \in \mathcal{W}. \tag{16}$$

On the subspace of $w \in \mathcal{W}$ such that $x = 0$ this condition takes the form:

$$\int_{t_0}^{t_f} \psi f_u(t, w^0)u dt + \langle \nu, g_u(t, w^0)u \rangle = 0 \quad \forall u \in L^\infty.$$

From this relation and the regularity assumption for g we easily obtain that the functional ν is absolutely continuous. Hence it is defined by an integrable function, which will be also denoted by ν . Then $\psi f_u(t, w^0) + \nu g_u(t, w^0) = 0$, i.e. $\bar{H}_u = 0$, where $\bar{H} = \psi f + \nu g$ (cf. (8)). It follows that ν is an essentially bounded function.

Now, setting $u = 0$ in (16) and taking into account that $\nu \in L^\infty$, we obtain

$$\begin{aligned} & \int_{t_0}^{t_f} (\psi(f_x(t, w^0)x - \dot{x}) + \nu g_x(t, w^0)x) dt \\ & + \beta(K_{x_0}(p^0)x(t_0) + K_{x_f}(p^0)x(t_f)) = 0 \quad \forall x \in W^{1,1}. \end{aligned} \tag{17}$$

It easily follows from (17) that the function ψ is absolutely continuous, and moreover, it satisfies the adjoint equation $-\dot{\psi} = \bar{H}_x$, and the transversality conditions $-\psi(t_0) = L_{x_0}$ and $\psi(t_f) = L_{x_f}$ with $L := \beta K$. If $\beta = 0$, then the conditions $\psi(t_f) = L_{x_f} = 0$, $-\dot{\psi} = \psi f_x$, $\psi f_u + \nu g_u = 0$ and the full rank condition for g_u imply that $\psi = 0$ and $\nu = 0$. Hence $\beta \neq 0$, and we can take a triple (β, ψ, ν) with $|\beta| = 1$. Set $\alpha_0 = 0$, $\alpha = 0$ and $\hat{\lambda} = (0, 0, \beta, \psi, \nu)$. We see that thus obtained tuple $\hat{\lambda}$ belongs to the set Λ_0 , and moreover, $-\hat{\lambda} \in \Lambda_0$. Then, for any element $w \in \mathcal{W}_2$ we have $\Omega(\hat{\lambda}, w) \geq 0$ or $\Omega(-\hat{\lambda}, w) \geq 0$. Thus, in the considered case, the set Λ_0 is nonempty, and $\max_{\Lambda_0} \Omega(\lambda, w) \geq 0$ on the whole space \mathcal{W}_2 . Hence condition \mathcal{A} trivially holds, although it is not informative in this case.

(b) Now, consider the main case: $G'(w^0)\mathcal{W} = \mathcal{Y}$. The following lemma holds in this case.

Lemma 3.1 *For any $\bar{w} \in \mathcal{K} \cap \mathcal{W}$ the following system of linear equalities and inequalities is inconsistent with respect to $\tilde{w} \in \mathcal{W}$:*

$$J'(p^0)\tilde{p} + \frac{1}{2}\langle J''(p^0)\bar{p}, \bar{p} \rangle < 0, \tag{18}$$

$$F'(p^0)\tilde{p} + \frac{1}{2}\langle F''(p^0)\bar{p}, \bar{p} \rangle < 0, \tag{19}$$

$$K'(p^0)\tilde{p} + \frac{1}{2}\langle K''(p^0)\bar{p}, \bar{p} \rangle = 0, \tag{20}$$

$$f_w(t, w^0)\tilde{w} - \tilde{x} + \frac{1}{2}\langle f_{ww}(t, w^0)\bar{w}, \bar{w} \rangle = 0, \tag{21}$$

$$g_w(t, w^0)\tilde{w} + \frac{1}{2}\langle g_{ww}(t, w^0)\bar{w}, \bar{w} \rangle = 0, \tag{22}$$

where $\tilde{p} = (\tilde{x}(t_0), \tilde{x}(t_f))$ and $\bar{p} = (\bar{x}(t_0), \bar{x}(t_f))$.

Proof Assume the contrary: let there exist $\bar{w} \in \mathcal{K} \cap \mathcal{W}$ and $\tilde{w} \in \mathcal{W}$ satisfying (18)–(22). Consider the curve $w^\varepsilon = w^0 + \varepsilon\bar{w} + \varepsilon^2\tilde{w}$ parameterized by $\varepsilon > 0$. From conditions (20)–(22) it easily follows that $\|G(w^\varepsilon)\| = o(\varepsilon)$. Then, by generalized Lyusternik’s theorem [17, 18], there exists a curve $\hat{w}^\varepsilon \in \mathcal{W}$ ($\varepsilon > 0$) such that $G(w^\varepsilon + \hat{w}^\varepsilon) = 0$ and $\|\hat{w}^\varepsilon\| = o(\varepsilon)$. Conditions (18)–(19) together with condition $\bar{w} \in \mathcal{K} \cap \mathcal{W}$ imply that $J(p^\varepsilon + \hat{p}^\varepsilon) < 0$ and $F(p^\varepsilon + \hat{p}^\varepsilon) < 0$ for small enough $\varepsilon > 0$. Since $\|w^\varepsilon + \hat{w}^\varepsilon - w^0\| \rightarrow 0$ ($\varepsilon \rightarrow 0$), the latter means that w^0 is not a local minimum in the problem. □

In order to analyse inconsistency of system (18)–(22), we will need the following well-known assertion (see, e.g. [16]).

Lemma 3.2 *Let X, Y be Banach spaces, $l_i : X \rightarrow \mathbb{R}$ linear functionals, a_i real numbers, $i = 1, \dots, k$, $A : X \rightarrow Y$ a liner surjective operator, $b \in Y$ a given element. The linear system (in $x \in X$)*

$$l_i(x) + a_i < 0, \quad i = 1, \dots, k, \quad Ax + b = 0$$

is inconsistent iff there exist numbers $\alpha_i \geq 0$ and a functional $y^ \in Y^*$ such that*

$$\sum_{i=1}^k \alpha_i l_i + y^* A = 0, \quad \sum_{i=1}^k \alpha_i > 0, \quad \sum_{i=1}^k \alpha_i a_i + y^* b \geq 0.$$

Applying this lemma to system (18)–(22), we obtain the following result: there exist $\alpha_i \geq 0, i = 0, 1, \dots, d(F), \beta \in (\mathbb{R}^{d(K)})^*, \psi \in L^\infty, \nu \in (L^\infty)^*$ such that $\sum_{i=0}^{d(F)} \alpha_i > 0$ and

$$L_p(p^0, \alpha_0, \alpha, \beta)p + \int_{t_0}^{t_f} \psi (f_w(t, w^0)w - \dot{x}) dt + \langle \nu, g_w(t, w^0)w \rangle = 0 \quad \forall w \in \mathcal{W}, \tag{23}$$

$$\langle L_{pp}(p^0, \alpha_0, \alpha, \beta)\bar{p}, \bar{p} \rangle + \int_{t_0}^{t_f} \langle \psi f_{ww}(t, w^0)\bar{w}, \bar{w} \rangle dt + \langle \nu, (g_{ww}(t, w^0)\bar{w}, \bar{w}) \rangle \geq 0, \tag{24}$$

where L is as in (9) and $\alpha = (\alpha_1, \dots, \alpha_{d(F)})$. Without any loss of generality we assume that

$$\sum_{i=0}^{d(F)} \alpha_i + \sum_{j=1}^{d(K)} |\beta_j| = 1.$$

The analysis of Eq. (23) is similar to that of (16). As a result we prove that ν is absolutely continuous functional given by a bounded measurable function (which we also denote by ν), the function ψ is absolutely continuous, and the tuple $\lambda = (\alpha_0, \alpha, \beta, \psi, \nu)$ satisfies all conditions in the definition of Λ_0 . Clearly, condition (24) means that $\Omega(\lambda, \bar{w}) \geq 0$. Thus, we have proved that for any $\bar{w} \in \mathcal{K} \cap \mathcal{W}$ there exists $\lambda \in \Lambda_0$ such that $\Omega(\lambda, \bar{w}) \geq 0$, and hence, $\max_{\Lambda_0} \Omega(\lambda, \cdot)$ is nonnegative on $\mathcal{K} \cap \mathcal{W}$. To get the same assertion on \mathcal{K} , it suffices to prove that the closure of the set $\mathcal{K} \cap \mathcal{W}$ in \mathcal{W}_2 is equal to \mathcal{K} . The latter easily can be proved using Hoffman’s lemma [19]. We omit this simple proof.

4 Necessary Second-Order Condition for an Extended Weak Local Minimum in the General Problem of the Calculus of Variations on a Variable Time Interval

4.1 General Problem of the Calculus of Variations on a Variable Time Interval. Main Results

Now, we consider a more general optimal control problem. 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 Δ is not fixed. For each process \mathcal{T} we denote here by

$$p := (t_0, x(t_0), t_f, x(t_f))$$

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

$$\min \mathcal{J}(\mathcal{T}) := J(p) \tag{25}$$

subject to the constraints

$$F(p) \leq 0, \quad K(p) = 0, \tag{26}$$

$$\dot{x}(t) = f(t, x(t), u(t)), \tag{27}$$

$$g(t, x(t), u(t)) = 0, \tag{28}$$

$$p \in \mathcal{P}, \quad (t, x(t), u(t)) \in \mathcal{Q}, \tag{29}$$

where \mathcal{P} and \mathcal{Q} are open sets, x, u, F, K, f and g are vector-functions. We call (25)–(28) the Problem B .

We assume that the given functions J, F and K are defined and twice continuously differentiable on \mathcal{P} , and the given functions f and g are defined and twice continuously differentiable on \mathcal{Q} . It is also assumed that the gradients with respect to the control $g_{iu}(t, x, u), i = 1, \dots, d(g)$ are linearly independent at each point $(t, x, u) \in \mathcal{Q}$ such that $g(t, x, u) = 0$ (the *regularity assumption* for g).

We say that a process \mathcal{T} is *admissible* iff it satisfies all constraints of the problem, and the graph of the function $w(t)$ lies in the set \mathcal{Q} with a certain margin (see Sec. 3.1). Let $\mathcal{T} = (x(t), u(t) \mid t \in [t_0, t_f])$ be a fixed admissible process. We say that \mathcal{T} is a *weak local minimum* iff there exists $\varepsilon > 0$ such that $\mathcal{J}(\tilde{\mathcal{T}}) \geq \mathcal{J}(\mathcal{T})$ for each admissible process $\tilde{\mathcal{T}} = (\tilde{x}(t), \tilde{u}(t) \mid t \in [\tilde{t}_0, \tilde{t}_f])$ satisfying the conditions

$$\begin{aligned} &|\tilde{t}_0 - t_0| < \varepsilon, \quad |\tilde{t}_f - t_f| < \varepsilon, \quad \max_{t \in \Delta \cap \tilde{\Delta}} |\tilde{x}(t) - x(t)| < \varepsilon, \\ &\text{ess sup}_{t \in \Delta \cap \tilde{\Delta}} |\tilde{u}(t) - u(t)| < \varepsilon, \end{aligned}$$

where $\tilde{\Delta} = [\tilde{t}_0, \tilde{t}_1]$.

In the sequel, we consider an admissible process

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

such that the control $u(\cdot)$ is a piecewise continuous function on the interval Δ with the set of discontinuity points

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

Moreover, we assume that the control $u(\cdot)$ is Lipschitz continuous on each interval $]t_{k-1}, t_k[, k = 1, \dots, s + 1$, where $t_{s+1} := t_f$ (in this case, we say that the function $u(\cdot)$ is *piecewise Lipschitz continuous* on Δ). Let us formulate the first-order necessary condition for optimality of the process \mathcal{T} . Again, we introduce the Pontryagin function $H(t, x, u, \psi)$, the augmented Pontryagin function $\bar{H}(t, x, u, \psi, \nu)$ and the endpoint

Lagrange function $L(p, \alpha_0, \alpha, \beta)$ defined as in (7), (8) and (9), respectively, but recall that now $p = (t_0, x(t_0), t_f, x(t_f))$. Introduce a tuple of Lagrange multipliers

$$\lambda = (\alpha_0, \alpha, \beta, \psi(\cdot), \psi_0(\cdot), \nu(\cdot)) \tag{30}$$

such that $\psi(\cdot) : \Delta \mapsto (\mathbb{R}^{d(x)})^*$ and $\psi_0(\cdot) : \Delta \mapsto \mathbb{R}^1$ are piecewise smooth functions, continuously differentiable on each interval of the set $\Delta \setminus \Theta$, and $\nu(\cdot) : \Delta \mapsto (\mathbb{R}^{d(g)})^*$ is a piecewise continuous function, Lipschitz continuous on each interval of the set $\Delta \setminus \Theta$.

Denote by Λ_0^Θ the set of all tuples λ satisfying the conditions

$$\begin{aligned} \alpha_0 \geq 0, \alpha \geq 0, \alpha F(p) = 0, \alpha_0 + \sum \alpha_i + \sum |\beta_j| = 1, \\ \dot{\psi} = -\bar{H}_x, \dot{\psi}_0 = -\bar{H}_t, \bar{H}_u = 0, t \in \Delta \setminus \Theta, \\ \psi(t_0) = -L_{x_0}, \psi(t_f) = L_{x_f}, \psi_0(t_0) = -L_{t_0}, \psi_0(t_f) = L_{t_f}, \\ H(t, x(t), u(t), \psi(t)) + \psi_0(t) = 0, t \in \Delta \setminus \Theta. \end{aligned} \tag{31}$$

The derivatives $L_{x_0}, L_{t_0}, L_{x_f}$ and L_{t_f} are at $(t_0, x(t_0), t_f, x(t_f), \alpha_0, \alpha, \beta)$, and the derivatives \bar{H}_x, \bar{H}_u and \bar{H}_t are at $(t, x(t), u(t), \psi(t), \nu(t))$, where $t \in \Delta \setminus \Theta$.

Let us give the definition of Θ -weak minimum in problem (25)–(28) on a variable interval $[t_0, t_f]$. For convenience, we assume that $u(\cdot)$ is left continuous at each point of discontinuity $t_k \in \Theta$. Denote by $\text{cl } u(\cdot)$ the closure of the graph of $u(t)$.

Definition 4.1 The process \mathcal{T} affords a Θ -weak minimum (or an extended weak minimum) iff there exist $\varepsilon > 0$ and a neighbourhood $V \subset \mathbb{R}^{d(u)+1}$ of the compact set $\text{cl } u(\cdot)$ 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) $\max_{\tilde{\Delta} \cap \Delta} |\tilde{x}(t) - x(t)| < \varepsilon, \quad \text{where } \tilde{\Delta} = [\tilde{t}_0, \tilde{t}_f],$
- (c) $(t, \tilde{u}(t)) \in V \quad \text{a.e. on } [\tilde{t}_0, \tilde{t}_f].$

The condition $\Lambda_0^\Theta \neq \emptyset$ is equivalent to the local minimum principle. It is the first-order necessary condition for a Θ -weak minimum for the process \mathcal{T} . Assume that Λ_0^Θ is nonempty. Using the definition of the set Λ_0^Θ and the full rank condition of the matrix g_u on the surface $g = 0$ one can easily prove that Λ_0^Θ is a finite-dimensional compact set, and the mapping $\lambda \mapsto (\alpha_0, \alpha, \beta)$ is injective on Λ_0^Θ .

Let us formulate the quadratic necessary condition for a Θ -weak minimum for the process \mathcal{T} . First, for this process, we introduce a Hilbert space $\mathcal{Z}_2(\Theta)$ and a ‘‘critical cone’’ $\mathcal{K}_\Theta \subset \mathcal{Z}_2(\Theta)$. As in Sect. 2, we denote here by $P_\Theta W^{1,2}(\Delta, \mathbb{R}^{d(x)})$ the Hilbert space of piecewise continuous functions $\tilde{x}(\cdot) : \Delta \mapsto \mathbb{R}^{d(x)}$, absolutely continuous on each interval of the set $\Delta \setminus \Theta$ and such that their first derivative is square integrable. For each $\tilde{x} \in P_\Theta W^{1,2}(\Delta, \mathbb{R}^{d(x)})$, $t_k \in \Theta$ we set

$$\tilde{x}^{k-} = \tilde{x}(t_k-), \quad \tilde{x}^{k+} = \tilde{x}(t_k+), \quad [\tilde{x}]^k = \tilde{x}^{k+} - \tilde{x}^{k-}.$$

Thus $[\tilde{x}]^k$ is the jump of the function $\tilde{x}(t)$ at the point $t_k \in \Theta$. Similar notation we will use to denote jumps of other functions at a point $t_k \in \Theta$.

Set

$$\tilde{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \tilde{x}, \tilde{u}),$$

where

$$\begin{aligned} \bar{t}_k &\in \mathbb{R}^1, \quad k = 0, 1, \dots, s, \quad \bar{t}_f \in \mathbb{R}^1, \\ \tilde{x} &\in P_\Theta W^{1,2}(\Delta, \mathbb{R}^{d(x)}), \quad \tilde{u} \in L^2(\Delta, \mathbb{R}^{d(u)}) \end{aligned}$$

(this notation will be convenient for the proofs). Thus,

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

Moreover, for given \tilde{z} we set

$$\tilde{w} := (\tilde{x}, \tilde{u}), \tag{32}$$

$$\bar{\bar{p}} := (\bar{t}_0, \tilde{x}(t_0) + \bar{t}_0 \dot{x}(t_0), \bar{t}_f, \tilde{x}(t_f) + \bar{t}_f \dot{x}(t_f)). \tag{33}$$

By $I_F(p) = \{i \in \{1, \dots, d(F)\} \mid F_i(p) = 0\}$ we denote the set of active indices of the constraints $F_i \leq 0$. Set $[\dot{x}]^k = \dot{x}(t_k+) - \dot{x}(t_k-)$.

Let \mathcal{K}_Θ be the set of all $\tilde{z} \in \mathcal{Z}_2(\Theta)$ satisfying the following conditions:

$$\begin{aligned} J'(p)\bar{\bar{p}} &\leq 0, \quad F'_i(p)\bar{\bar{p}} \leq 0 \quad \forall i \in I_F(p), \quad K'(p)\bar{\bar{p}} = 0, \\ \dot{\tilde{x}}(t) &= f_w(t, w(t))\tilde{w}(t), \quad \text{for a.a. } t \in [t_0, t_f], \\ [\tilde{x}]^k + [\dot{x}]^k \bar{t}_k &= 0, \quad k = 1, \dots, s \\ g_w(t, w(t))\tilde{w}(t) &= 0, \quad \text{for a.a. } t \in [t_0, t_f], \end{aligned} \tag{34}$$

where $p = (t_0, x(t_0), t_f, x(t_f))$, $w = (x, u)$. It is obvious that \mathcal{K}_Θ is a convex cone in the Hilbert space $\mathcal{Z}_2(\Theta)$. We call it again the *critical cone*.

Let us introduce a quadratic form on $\mathcal{Z}_2(\Theta)$. For $\lambda \in \Lambda_\Theta^\ominus$ and $\tilde{z} \in \mathcal{Z}_2(\Theta)$, we set

$$\begin{aligned} \Omega_\Theta(\lambda, \tilde{z}) &= \langle L_{pp}\bar{\bar{p}}, \bar{\bar{p}} \rangle + \int_{t_0}^{t_f} \langle \bar{H}_{ww}\tilde{w}(t), \tilde{w}(t) \rangle dt \\ &+ \sum_{k=1}^s [\dot{\psi}_0 + \dot{\psi}\dot{x}]^k \bar{t}_k^2 + 2[\dot{\psi}\tilde{x}]^k \bar{t}_k \\ &+ (\dot{\psi}_0(t_0) + \dot{\psi}(t_0)\dot{x}(t_0))\bar{t}_0^2 + 2\dot{\psi}(t_0)\tilde{x}(t_0)\bar{t}_0 \\ &- (\dot{\psi}_0(t_f) + \dot{\psi}(t_f)\dot{x}(t_f))\bar{t}_f^2 - 2\dot{\psi}(t_f)\tilde{x}(t_f)\bar{t}_f, \end{aligned} \tag{35}$$

where $L_{pp} = L_{pp}(p, \alpha_0, \alpha, \beta)$, $p = (t_0, x(t_0), t_f, x(t_f))$.

Theorem 4.1 *If the process \mathcal{T} yields a Θ -weak minimum, then the following Condition $\mathcal{A}(\Theta)$ holds: the set Λ_Θ^\ominus is nonempty, and*

$$\max_{\lambda \in \Lambda_\Theta^\ominus} \Omega_\Theta(\lambda, \tilde{z}) \geq 0 \text{ for all } \tilde{z} \in \mathcal{K}_\Theta.$$

Before starting the proof of this theorem, we make two remarks.

1. Condition $\mathcal{A}(\Theta)$ of Theorem 4.1, obtained for the case of optimal control with one break (Θ is a singleton), cannot be automatically extended to the case of finitely many breaks. Hence the case of multiple breaks requires a separate consideration.
2. There is a simple method of converting inequality constraints into equality constraints by adding new variables squared to the left-hand sides of inequalities. Often, this method leads to certain losses, even for necessary conditions of the first order: the conditions get weaker, while the assumptions get stronger. The more this is true for second-order conditions.

4.2 Proofs

The proofs are based on the quadratic necessary optimality conditions for a weak minimum, obtained in Sect. 3 for the problem on a fixed interval of time. We will give the proofs omitting some details. In order to extend the proofs to the case of a variable interval $[t_0, t_f]$ we use a simple change of the time variable. Namely, with the fixed admissible process

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

in problem (25)–(28) on a variable time interval we associate a process

$$\mathcal{T}^\tau = (t(\tau), x(\tau), u(\tau), v(\tau) \mid \tau \in [\tau_0, \tau_f]),$$

considered on a fixed interval $[\tau_0, \tau_f]$, where $\tau_0 = t_0, \tau_f = t_f, t(\tau) \equiv \tau, v(\tau) \equiv 1$. This is an admissible process in the following problem on a fixed interval $[\tau_0, \tau_f]$: to minimize the cost function

$$\min \mathcal{J}(\mathcal{T}^\tau) := J(t(\tau_0), x(\tau_0), t(\tau_f), x(\tau_f)) \tag{36}$$

subject to the constraints

$$F(t(\tau_0), x(\tau_0), t(\tau_f), x(\tau_f)) \leq 0, \tag{37}$$

$$K(t(\tau_0), x(\tau_0), t(\tau_f), x(\tau_f)) = 0, \tag{38}$$

$$\frac{dx(\tau)}{d\tau} = v(\tau)f(t(\tau), x(\tau), u(\tau)), \quad \frac{dt(\tau)}{d\tau} = v(\tau), \tag{39}$$

$$g(t(\tau), x(\tau), u(\tau)) = 0, \tag{40}$$

$$(t(\tau_0), x(\tau_0), t(\tau_f), x(\tau_f)) \in \mathcal{P}, \quad (t(\tau), x(\tau), u(\tau)) \in \mathcal{Q}, \quad v(\tau) > 0. \tag{41}$$

In this problem, $t(\tau)$ and $x(\tau)$ are state variables, and $u(\tau)$ and $v(\tau)$ are control variables. Remember that we refer to problem (25)–(28) as problem B (on a variable interval $\Delta = [t_0, t_f]$). We will refer to problem (36)–(40) as problem B^τ (on a fixed interval $[\tau_0, \tau_f]$). We denote by \mathcal{A}^τ the necessary quadratic condition \mathcal{A} (see Theorem 3.1) for process \mathcal{T}^τ in problem B^τ on a fixed interval $[\tau_0, \tau_f]$.

Recall that the control $u(\cdot)$ is a piecewise Lipschitz-continuous function on the interval $\Delta = [t_0, t_f]$ with the set of discontinuity points $\Theta = \{t_1, \dots, t_s\}$, where $t_0 < t_1 < \dots < t_s < t_f$. Hence, for each $\lambda \in \Lambda_0^\Theta$, the function $v(t)$ is also piecewise Lipschitz continuous on the interval Δ , and, moreover, all discontinuity points of v belong to Θ . Consequently, $\dot{u}(t)$ and $\dot{v}(t)$ are bounded measurable functions on Δ .

The proof of Theorem 4.1 is composed of the following chain of implications:

- (i) A Θ -weak minimum is attained on \mathcal{T} in Problem $B \implies$
- (ii) A weak minimum is attained on \mathcal{T}^τ in Problem $B^\tau \implies$
- (iii) Condition \mathcal{A}^τ for \mathcal{T}^τ in Problem $B^\tau \implies$
- (iv) Condition $\mathcal{A}(\Theta)$ for \mathcal{T} in Problem B .

The first implication is readily verified, the second follows from Theorem 3.1. The verification of the third implication (iii) \implies (iv) is not short and rather technical: we have to compare the sets of Lagrange multipliers, the critical cones and the quadratic forms in the both problems. This will be done below.

Comparison of the sets of Lagrange multipliers. Let us formulate the local minimum principle in problem B^τ for the process \mathcal{T}^τ . According to Eqs. (7)–(9), the endpoint Lagrange function L , the Pontryagin function H and the augmented Pontryagin function \bar{H} (all of them are equipped with the superscript τ) have the form:

$$\begin{aligned} L^\tau &= \alpha_0 J + \alpha F + \beta K = L, \\ H^\tau &= \psi f v + \psi_0 v = v(\psi f + \psi_0), \quad \bar{H}^\tau = H^\tau + v g. \end{aligned} \tag{42}$$

Let us explain that here the adjoint variable ψ corresponds to the first differential equation in (39), and the adjoint variable ψ_0 corresponds to the second one.

According to (11), the set Λ_0^τ in problem B^τ for the process \mathcal{T}^τ consists of all tuples of Lagrange multipliers $\lambda^\tau = (\alpha_0, \alpha, \beta, \psi, \psi_0, v)$ such that the following conditions hold:

$$\alpha_0 \geq 0, \alpha \geq 0, \alpha F(p) = 0, \alpha_0 + |\alpha| + |\beta| = 1, \tag{43}$$

$$-\frac{d\psi}{d\tau} = v\psi f_x + v g_x, \quad -\frac{d\psi_0}{d\tau} = v\psi f_t + v g_t, \tag{44}$$

$$\psi(\tau_0) = -L_{x_0}, \quad \psi(\tau_f) = L_{x_f}, \quad \psi_0(\tau_0) = -L_{t_0}, \quad \psi_0(\tau_f) = L_{t_f}, \tag{45}$$

$$\bar{H}_u^\tau = v\psi f_u + v g_u = 0, \quad \bar{H}_v^\tau = \psi f + \psi_0 = 0. \tag{46}$$

Recall that here $v(\tau) \equiv 1, t(\tau) \equiv \tau, \tau_0 = t_0, \tau_f = t_f$. In (44)–(46), the function f and the derivatives $f_x, f_u, f_t, g_x, g_u, g_t$ are taken at $(t(\tau), x(\tau), u(\tau))$, $\tau \in [\tau_0, \tau_f] \setminus \Theta$, while the derivatives $L_{t_0}, L_{x_0}, L_{t_f}, L_{x_f}$ are calculated at $(t(\tau_0), x(\tau_0), t(\tau_f), x(\tau_f)) = (t_0, x(t_0), t_f, x(t_f))$. This implies that $\Lambda_0^\tau = \Lambda_0^\Theta$ (cf. (31)).

Comparison of the critical cones. For brevity, we set

$$\varrho := (t, x, u, v) = (t, w, v).$$

Let us define the critical cone \mathcal{K}^τ in problem B^τ for the process T^τ . According to definition (12), it consists of all tuples $\bar{q} = (\bar{t}, \bar{x}, \bar{u}, \bar{v})$ satisfying the relations:

$$J_{t_0}\bar{t}(\tau_0) + J_{x_0}\bar{x}(\tau_0) + J_{t_f}\bar{t}(\tau_f) + J_{x_f}\bar{x}(\tau_f) \leq 0, \tag{47}$$

$$F_{it_0}\bar{t}(\tau_0) + F_{ix_0}\bar{x}(\tau_0) + F_{it_f}\bar{t}(\tau_f) + F_{ix_f}\bar{x}(\tau_f) \leq 0, \quad i \in I_F(p), \tag{48}$$

$$K_{t_0}\bar{t}(\tau_0) + K_{x_0}\bar{x}(\tau_0) + K_{t_f}\bar{t}(\tau_f) + K_{x_f}\bar{x}(\tau_f) = 0, \tag{49}$$

$$\frac{d\bar{x}}{d\tau} = \bar{v}f + v(f_t\bar{t} + f_x\bar{x} + f_u\bar{u}), \tag{50}$$

$$\frac{d\bar{t}}{d\tau} = \bar{v}, \tag{51}$$

$$g_t\bar{t} + g_x\bar{x} + g_u\bar{u} = 0, \tag{52}$$

where the derivatives $J_{t_0}, J_{x_0}, J_{t_f}, J_{x_f}$, etc. are calculated at the point

$$(t(\tau_0), x(\tau_0), t(\tau_f), x(\tau_f)) = (t_0, x(t_0), t_f, x(t_f)),$$

while $f, f_t, f_x, f_u, g_t, g_x$ and g_u are taken at the point

$$(t(\tau), x(\tau), u(\tau)), \quad \tau \in [\tau_0, \tau_f] \setminus \Theta.$$

Let $\bar{q} = (\bar{t}, \bar{x}, \bar{u}, \bar{v})$ be an element of the critical cone \mathcal{K}^τ . We will make use of the following change of variables:

$$\tilde{x} = \bar{x} - \bar{t}\dot{\bar{x}}, \quad \tilde{u} = \bar{u} - \bar{t}\dot{\bar{u}}, \tag{53}$$

or briefly

$$\tilde{w} = \bar{w} - \bar{t}\dot{\bar{w}}. \tag{54}$$

Since $v = 1, \dot{x} = f$ and $t = \tau$, Eq. (50) is equivalent to the equation

$$\frac{d\bar{x}}{dt} = \bar{v}\dot{\bar{x}} + f_t\bar{t} + f_w\bar{w}. \tag{55}$$

Using the relation $\bar{x} = \tilde{x} + \bar{t}\dot{\bar{x}}$ in this equation along with $\dot{\bar{t}} = \bar{v}$, we get

$$\dot{\tilde{x}} + \bar{t}\ddot{\bar{x}} = f_t\bar{t} + f_w\bar{w}. \tag{56}$$

By differentiating the equation $\dot{x}(t) = f(t, w(t))$, we obtain

$$\ddot{x} = f_t + f_w\dot{w}. \tag{57}$$

Using this relation in (56), we get

$$\dot{\tilde{x}} = f_w\dot{\tilde{w}}. \tag{58}$$

The relations

$$\bar{x} = \tilde{x} + \bar{t}\dot{x}, \quad [\bar{x}]^k = 0, \quad k = 1, \dots, s,$$

imply

$$[\tilde{x}]^k + [\dot{x}]^k \bar{t}_k = 0, \tag{59}$$

where

$$\bar{t}_k = \bar{t}(t_k), \quad k = 1, \dots, s. \tag{60}$$

Further, relation (52) can be written as $g_t \bar{t} + g_w \bar{w} = 0$. Differentiating the relation $g(t, w(t)) = 0$ we obtain

$$g_t + g_w \dot{w} = 0. \tag{61}$$

These relations along with (54) imply that

$$g_w \tilde{w} = 0. \tag{62}$$

Finally, note that since $\bar{x} = \tilde{x} + \bar{t}\dot{x}$, and $\tau_0 = t_0, \tau_f = t_f$, we have

$$\bar{p} = (\bar{t}_0, \bar{x}(t_0), \bar{t}_f, \bar{x}(t_f)) = (\bar{t}_0, \tilde{x}(t_0) + \bar{t}_0 \dot{x}(t_0), \bar{t}_f, \tilde{x}(t_f) + \bar{t}_f \dot{x}(t_f)), \tag{63}$$

where $\bar{t}_0 = \bar{t}(t_0)$ and $\bar{t}_f = \bar{t}(t_f)$. The vector in the right-hand side of the last equality has the same form as the vector \bar{p} in definition (33). Consequently, all relations in definition (34) of the critical cone \mathcal{K}_Θ in problem B are satisfied for the element $\tilde{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \tilde{w})$. We have proved that thus obtained element \tilde{z} belongs to the critical cone \mathcal{K}_Θ in problem B .

Vice versa, let $(\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \tilde{w})$ be an element of the critical cone \mathcal{K}_Θ in problem B . Let us take a Lipschitz-continuous function $\bar{t}(\tau)$ satisfying

$$\bar{t}(t_0) = \bar{t}_0, \quad \bar{t}(t_f) = \bar{t}_f, \quad \bar{t}(t_k) = \bar{t}_k, \quad k = 1, \dots, s;$$

e.g. one can take a continuous function \bar{t} , affine at each interval (t_{k-1}, t_k) , $k = 1, \dots, s + 1$, where $t_{s+1} = t_f$. Set

$$\bar{v} = \dot{\bar{t}}, \quad \bar{w} = \tilde{w} + \bar{t}\dot{w}.$$

Then, we obtain an element $(\bar{t}, \bar{w}, \bar{v})$ of the critical cone \mathcal{K}^τ (see (47)–(52)) in problem B^τ . Thus, we have proved the following lemma.

Lemma 4.1 *If $(\bar{t}, \bar{w}, \bar{v})$ is an element of the critical cone \mathcal{K}^τ , as in (47)–(52), in problem B^τ for the trajectory \mathcal{T}^τ and*

$$\bar{t}_0 = \bar{t}(t_0), \quad \bar{t}_f = \bar{t}(t_f), \quad \bar{w} = \tilde{w} - \bar{t}\dot{w}, \quad \bar{t}_k = \bar{t}(t_k), \quad k = 1, \dots, s, \tag{64}$$

then $(\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w})$ is an element of the critical cone \mathcal{K}_Θ , as in (34), in problem B for the trajectory \mathcal{T} . Moreover, relations (64) define a surjective mapping of the critical cone \mathcal{K}^τ on the critical cone \mathcal{K}_Θ .

We will say that an element $(\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w})$ of the critical cone \mathcal{K}_Θ in problem B corresponds to an element $(\bar{t}, \bar{w}, \bar{v})$ of the critical cone \mathcal{K}^τ in problem B^τ iff relations (64) hold.

Comparison of the quadratic forms. Let an element

$$\bar{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w})$$

of the critical cone \mathcal{K}_Θ in problem B corresponds to an element $\bar{\varrho} = (\bar{t}, \bar{w}, \bar{v})$ of the critical cone \mathcal{K}^τ in problem B^τ . Assume that $\lambda \in \Lambda_0^\tau$ (recall that $\Lambda_0^\tau = \Lambda_0^\Theta$). Let us show that the quadratic form $\Omega^\tau(\lambda, \cdot)$ (see (13)), calculated on the element $\bar{\varrho}$, can be transformed to the quadratic form $\Omega_\Theta(\lambda, \cdot)$ (as in (35)), calculated on the corresponding element \bar{z} .

(i) Set

$$\varrho = (t, w, v), \quad \bar{\varrho} = (\bar{t}, \bar{w}, \bar{v}).$$

It follows from (13) and (42) that

$$\Omega^\tau(\lambda, \bar{\varrho}) = \langle L_{pp} \bar{p}, \bar{p} \rangle + \int_{t_0}^{t_f} \langle \bar{H}_{\varrho\varrho}^\tau \bar{\varrho}, \bar{\varrho} \rangle dt, \tag{65}$$

where

$$\langle \bar{H}_{\varrho\varrho}^\tau \bar{\varrho}, \bar{\varrho} \rangle = \langle \bar{H}_{ww} \bar{w}, \bar{w} \rangle + 2\bar{H}_{tw} \bar{w} \bar{t} + \bar{H}_{tt} \bar{t}^2 + 2\bar{v}(\bar{H}_w \bar{w} + \bar{H}_t \bar{t}). \tag{66}$$

Since $\bar{w} = \tilde{w} + \bar{t}\dot{w}$, we have

$$\langle \bar{H}_{ww} \bar{w}, \bar{w} \rangle = \langle \bar{H}_{ww} \tilde{w}, \tilde{w} \rangle + 2\langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle \bar{t} + \langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle \bar{t}^2. \tag{67}$$

$$2\bar{H}_{tw} \bar{w} \bar{t} = 2\bar{H}_{tw} \tilde{w} \bar{t} + 2\bar{H}_{tw} \dot{w} \bar{t}^2, \tag{68}$$

Moreover, using the relations

$$\begin{aligned} H_w &= \bar{H}_w - v g_w, & H_t &= \bar{H}_t - v g_t, & g_t \bar{t} + g_w \bar{w} &= 0, \\ -\dot{\psi} &= \bar{H}_x, & -\dot{\psi}_0 &= \bar{H}_t, & \bar{H}_u &= 0, & g_t + g_w \dot{w} &= 0, \end{aligned}$$

we obtain

$$\begin{aligned} H_w \bar{w} + H_t \bar{t} &= \bar{H}_w \bar{w} + \bar{H}_t \bar{t} - v(g_w \bar{w} + g_t \bar{t}) \\ &= \bar{H}_w \bar{w} + \bar{H}_t \bar{t} = \bar{H}_x \bar{x} + \bar{H}_t \bar{t} \\ &= -\dot{\psi} \bar{x} - \dot{\psi}_0 \bar{t} = -\dot{\psi}(\bar{x} + \bar{t}\dot{x}) - \dot{\psi}_0 \bar{t} \\ &= -\dot{\psi} \tilde{x} - (\dot{\psi} \dot{x} + \dot{\psi}_0) \bar{t}. \end{aligned} \tag{69}$$

Relations (66)–(69) imply

$$\begin{aligned} \langle \bar{H}_{\bar{q}q}^{\tau} \bar{q}, \bar{q} \rangle &= \langle \bar{H}_{ww} \tilde{w}, \tilde{w} \rangle + 2 \langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle \bar{t} + \langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle \bar{t}^2 \\ &\quad + 2 \bar{H}_{tw} \tilde{w} \bar{t} + 2 \bar{H}_{tw} \dot{w} \bar{t}^2 + \bar{H}_{tt} \bar{t}^2 \\ &\quad - 2 \dot{\psi} \tilde{x} \tilde{v} - 2 (\dot{\psi}_0 + \dot{\psi} \dot{x}) \bar{t} \tilde{v}. \end{aligned} \tag{70}$$

Consequently,

$$\begin{aligned} \langle \bar{H}_{\bar{q}q}^{\tau} \bar{q}, \bar{q} \rangle &= \langle \bar{H}_{ww} \tilde{w}, \tilde{w} \rangle + 2 (\langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle + \bar{H}_{tw} \tilde{w}) \bar{t} \\ &\quad + (\langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle + \bar{H}_{tw} \dot{w}) \bar{t}^2 + (\bar{H}_{tw} \dot{w} + \bar{H}_{tt}) \bar{t}^2 \\ &\quad - 2 \dot{\psi} \tilde{x} \tilde{v} - 2 (\dot{\psi}_0 + \dot{\psi} \dot{x}) \bar{t} \tilde{v}. \end{aligned} \tag{71}$$

(ii) Let us transform the terms $2 (\langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle + \bar{H}_{tw} \tilde{w}) \bar{t}$ in (71). By differentiating the equation $-\dot{\psi} = \bar{H}_x$ with respect to t , we obtain

$$-\ddot{\psi} = \bar{H}_{tx} + (\dot{w})^* \bar{H}_{wx} + \dot{\psi} \bar{H}_{\psi x} + \dot{v} \bar{H}_{vx}.$$

Here $\bar{H}_{\psi x} = f_x$ and $\bar{H}_{vx} = g_x$. Therefore

$$-\ddot{\psi} = \bar{H}_{tx} + (\dot{w})^* \bar{H}_{wx} + \dot{\psi} f_x + \dot{v} g_x. \tag{72}$$

Similarly, by differentiating the equation $\bar{H}_u = 0$ with respect to t , we obtain

$$0 = \bar{H}_{tu} + (\dot{w})^* \bar{H}_{wu} + \dot{\psi} f_u + \dot{v} g_u. \tag{73}$$

Multiplying Eq. (72) by \tilde{x} and Eq. (73) by \tilde{u} and summing the results we get

$$-\ddot{\psi} \tilde{x} = \bar{H}_{tw} \tilde{w} + \langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle + \dot{\psi} f_w \tilde{w} + \dot{v} g_w \tilde{w}.$$

But $f_w \tilde{w} = \dot{\tilde{x}}$ and $g_w \tilde{w} = 0$. Therefore,

$$-\ddot{\psi} \tilde{x} = \bar{H}_{tw} \tilde{w} + \langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle + \dot{\psi} \dot{\tilde{x}},$$

whence

$$\bar{H}_{tw} \tilde{w} + \langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle = -\frac{d}{dt} (\dot{\psi} \tilde{x}). \tag{74}$$

This implies that

$$2 (\langle \bar{H}_{ww} \dot{w}, \tilde{w} \rangle + \bar{H}_{tw} \tilde{w}) \bar{t} = -2 \bar{t} \frac{d}{dt} (\dot{\psi} \tilde{x}). \tag{75}$$

(iii) Let us transform the terms $(\langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle + \bar{H}_{tw} \dot{w}) \bar{t}^2$ in (71). Multiplying Eq. (72) by \dot{x} and equation (73) by \dot{u} and summing the results we obtain

$$-\ddot{\psi} \dot{x} = \bar{H}_{tw} \dot{w} + \langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle + \dot{\psi} f_w \dot{w} + \dot{v} g_w \dot{w}. \tag{76}$$

From (57) and (61) we get $f_w \dot{w} = \ddot{x} - f_t$ and $g_w \dot{w} = -g_t$, respectively. Then (76) implies

$$\bar{H}_{tw} \dot{w} + \langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle = -\frac{d}{dt}(\dot{\psi} \dot{x}) + (\dot{\psi} f_t + \dot{v} g_t). \tag{77}$$

Multiplying this relation by \bar{t}^2 we get

$$\langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle + \bar{H}_{tw} \dot{w} \bar{t}^2 = -\bar{t}^2 \frac{d}{dt}(\dot{\psi} \dot{x}) + (\dot{\psi} f_t + \dot{v} g_t) \bar{t}^2. \tag{78}$$

(iv) Finally, let us transform the terms $(\bar{H}_{tw} \dot{w} + \bar{H}_{tt}) \bar{t}^2$ in (71). Differentiating the equation $-\dot{\psi}_0 = \bar{H}_t$ with respect to t and using the relations $\bar{H}_{\psi t} = f_t$ and $\bar{H}_{vt} = g_t$, we get

$$-\ddot{\psi}_0 = \bar{H}_{tt} + \bar{H}_{tw} \dot{w} + (\dot{\psi} f_t + \dot{v} g_t). \tag{79}$$

Consequently,

$$(\bar{H}_{tw} \dot{w} + \bar{H}_{tt}) \bar{t}^2 = -\ddot{\psi}_0 \bar{t}^2 - (\dot{\psi} f_t + \dot{v} g_t) \bar{t}^2. \tag{80}$$

(v) Summing Eqs. (78) and (80) we obtain

$$\langle \bar{H}_{ww} \dot{w}, \dot{w} \rangle \bar{t}^2 + \bar{H}_{tt} \bar{t}^2 + 2\bar{H}_{tw} \dot{w} \bar{t}^2 = -\ddot{\psi}_0 \bar{t}^2 - \bar{t}^2 \frac{d}{dt}(\dot{\psi} \dot{x}). \tag{81}$$

(vi) Using relations (75) and (81) in (70) we get

$$\begin{aligned} \langle \bar{H}_{\bar{q}\bar{q}} \bar{q}, \bar{q} \rangle &= \langle \bar{H}_{ww} \tilde{w}, \tilde{w} \rangle - 2\bar{t} \frac{d}{dt}(\dot{\psi} \tilde{x}) \\ &\quad - \ddot{\psi}_0 \bar{t}^2 - \bar{t}^2 \frac{d}{dt}(\dot{\psi} \dot{x}) - 2\dot{\psi} \tilde{x} \bar{v} - 2(\dot{\psi}_0 + \dot{\psi} \dot{x}) \bar{t} \bar{v}. \end{aligned} \tag{82}$$

But

$$\ddot{\psi}_0 \bar{t}^2 + 2\bar{v} \bar{t} \dot{\psi}_0 = \frac{d}{dt}(\dot{\psi}_0 \bar{t}^2), \quad \bar{t} \frac{d}{dt}(\dot{\psi} \tilde{x}) + \bar{v}(\dot{\psi} \tilde{x}) = \frac{d}{dt}(\bar{t} \dot{\psi} \tilde{x}),$$

$$2\bar{t} \bar{v}(\dot{\psi} \dot{x}) + \bar{t}^2 \frac{d}{dt}(\dot{\psi} \dot{x}) = \frac{d}{dt}(\dot{\psi} \dot{x} \bar{t}^2).$$

Therefore,

$$\langle \bar{H}_{\bar{q}\bar{q}} \bar{q}, \bar{q} \rangle = \langle \bar{H}_{ww} \tilde{w}, \tilde{w} \rangle - \frac{d}{dt} \left((\dot{\psi} \dot{x}) \bar{t}^2 + \dot{\psi}_0 \bar{t}^2 + 2\dot{\psi} \tilde{x} \bar{t} \right). \tag{83}$$

We have proved the following lemma.

Lemma 4.2 *Let $\bar{\varrho} = (\bar{t}, \bar{w}, \bar{v}) \in \mathcal{K}^\tau$ and $\bar{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w}) \in \mathcal{K}_\Theta$ be such that the relations (64) hold, and let $\lambda \in \Lambda_0^\Theta$. Then formula (83) holds.*

(vii) Recall that $\tau_0 = t_0, \tau_f = t_f, t(\tau) = \tau, dt = d\tau$. Since the functions $\dot{\psi}_0, \dot{\psi}, \dot{x}$ and \tilde{x} can have discontinuities only at the points of the set Θ , the following formula holds:

$$\int_{t_0}^{t_f} \frac{d}{dt} \left((\dot{\psi}_0 + \dot{\psi}\dot{x})\bar{t}^2 + 2\dot{\psi}\tilde{x}\bar{t} \right) dt = \left((\dot{\psi}_0 + \dot{\psi}\dot{x})\bar{t}^2 + 2\dot{\psi}\tilde{x}\bar{t} \right) \Big|_{t_0}^{t_f} - \sum_{k=1}^s \left([\dot{\psi}_0 + \dot{\psi}\dot{x}]^k \bar{t}(t_k)^2 + 2[\dot{\psi}\tilde{x}]^k \bar{t}(t_k) \right). \tag{84}$$

Formula (83) along with formula (84) gives the following transformation of quadratic form Ω^τ , as in (65), on the element $\bar{\varrho}$ of the critical cone \mathcal{K}^τ :

$$\begin{aligned} \Omega^\tau(\lambda, \bar{\varrho}) &= \langle L_{pp}\bar{p}, \bar{p} \rangle + \int_{t_0}^{t_f} \langle \bar{H}_{ww}\bar{w}, \bar{w} \rangle dt \\ &\quad - \left((\dot{\psi}_0 + \dot{\psi}\dot{x})\bar{t}^2 + 2\dot{\psi}\tilde{x}\bar{t} \right) \Big|_{t_0}^{t_f} \\ &\quad + \sum_{k=1}^s \left([\dot{\psi}_0 + \dot{\psi}\dot{x}]^k \bar{t}(t_k)^2 + 2[\dot{\psi}\tilde{x}]^k \bar{t}(t_k) \right). \end{aligned} \tag{85}$$

Taking into account (64) and the relation $\bar{p} = \bar{\bar{p}}$, we see that the right-hand side of (85) is the quadratic form $\Omega_{\Theta}(\lambda, \bar{z})$ (35) in problem B for the process \mathcal{T} , where $\bar{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w})$ is the corresponding element of the critical cone \mathcal{K}_Θ . Thus, we have proved the following theorem.

Theorem 4.2 *Let $\bar{\varrho} = (\bar{t}, \bar{w}, \bar{v})$ be an element of the critical cone \mathcal{K}^τ in problem B^τ for the process \mathcal{T}^τ . Let $\bar{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w})$ be the corresponding element of the critical cone \mathcal{K}_Θ in problem B for the process \mathcal{T} , i.e. relations (64) hold. Then for any $\lambda \in \Lambda_0^\tau$ the following equality holds: $\Omega^\tau(\lambda, \bar{\varrho}) = \Omega_\Theta(\lambda, \bar{z})$.*

This theorem proves the implication (iii) \Rightarrow (iv) (see the beginning of this subsection). Indeed, suppose that Condition \mathcal{A}^τ holds for the process \mathcal{T}^τ in problem B^τ , and let $\bar{z} = (\bar{t}_0, \bar{t}_1, \dots, \bar{t}_s, \bar{t}_f, \bar{w})$ be an arbitrary element of the critical cone \mathcal{K}_Θ in problem B . Then by Lemma 4.1 there exists an element $\bar{\varrho} = (\bar{t}, \bar{w}, \bar{v})$ of the critical cone \mathcal{K}^τ in problem B^τ for the process \mathcal{T}^τ such that relations (64) hold. Since Λ_0^Θ is a compact set, and Condition \mathcal{A}^τ holds in problem B^τ , there exists an element $\lambda \in \Lambda_0^\Theta$ such that $\Omega^\tau(\lambda, \bar{\varrho}) \geq 0$. By Theorem 4.2 we have $\Omega^\tau(\lambda, \bar{\varrho}) = \Omega_\Theta(\lambda, \bar{z})$. Consequently, $\Omega_\Theta(\lambda, \bar{z}) \geq 0$, i.e. Condition $\mathcal{A}(\Theta)$ holds for the process \mathcal{T} in problem B . Thus, we have proved the implication (iii) \Rightarrow (iv). This completes the proof of Theorem 4.1.

4.3 Equivalent Formulation of Main Results

In [5], [10] and [11], similar results were presented in another form. We will show that the critical cone and the quadratic form defined in the present work can be transformed to those in [5], [10] and [11].

Let us transform the terms related to the discontinuity points t_k of the control $u(\cdot)$, $k = 1, \dots, s$. Set, as in [5], [10] and [11] (cf. (1)),

$$D^k(\bar{H}) = \dot{\psi}^{k+} \dot{x}^{k-} - \dot{\psi}^{k-} \dot{x}^{k+} + [\dot{\psi}_0]^k, \tag{86}$$

where $\dot{\psi}^{k-} := \dot{\psi}(t_k-)$, $\dot{\psi}^{k+} := \dot{\psi}(t_k+)$, etc.

Lemma 4.3 *Let $\lambda \in M_0$ be an arbitrary element. Then for any $k = 1, \dots, s$ the following formula holds:*

$$[\dot{\psi}_0 + \dot{\psi} \dot{x}]^k \bar{t}(t_k)^2 + 2[\dot{\psi} \tilde{x}]^k \bar{t}(t_k) = D^k(\bar{H}) \bar{\xi}_k^2 - 2[\dot{\psi}]^k \tilde{x}_{av}^k \bar{\xi}_k, \tag{87}$$

where $\bar{\xi}_k = -\bar{t}(t_k)$.

Proof Everywhere in this proof we will omit the subscript and superscript k . We will also write \bar{t} instead of $\bar{t}(t_k)$. Taking into account that

$$D(\bar{H}) = \dot{\psi}^+ \dot{x}^- - \dot{\psi}^- \dot{x}^+ + [\dot{\psi}_0],$$

we obtain

$$\begin{aligned} & [\dot{\psi}_0 + \dot{\psi} \dot{x}] \bar{t}^2 + 2[\dot{\psi} \tilde{x}] \bar{t} \\ &= \bar{\xi}^2 [\dot{\psi}_0] + \bar{\xi}^2 (\dot{\psi}^+ \dot{x}^+ - \dot{\psi}^- \dot{x}^-) - 2\bar{\xi} (\dot{\psi}^+ \tilde{x}^+ - \dot{\psi}^- \tilde{x}^-) \\ &= \bar{\xi}^2 [\dot{\psi}_0] + \bar{\xi}^2 (\dot{\psi}^+ \dot{x}^- - \dot{\psi}^- \dot{x}^+) - 2\bar{\xi} \left(\dot{\psi}^+ (\tilde{x}_{av} + \frac{1}{2}[\dot{x}]\bar{\xi}) - \dot{\psi}^- (\tilde{x}_{av} - \frac{1}{2}[\dot{x}]\bar{\xi}) \right) \\ &= \bar{\xi}^2 [\dot{\psi}_0] + \bar{\xi}^2 (\dot{\psi}^+ \dot{x}^+ - \dot{\psi}^- \dot{x}^- - \dot{\psi}^+ [\dot{x}] + \dot{\psi}^- [\dot{x}]) - 2\bar{\xi} [\dot{\psi}] \tilde{x}_{av} \\ &= \bar{\xi}^2 [\dot{\psi}_0] + \bar{\xi}^2 (\dot{\psi}^+ \dot{x}^+ - \dot{\psi}^- \dot{x}^- - \dot{\psi}^+ (\dot{x}^+ - \dot{x}^-) + \dot{\psi}^- (\dot{x}^+ - \dot{x}^-)) - 2\bar{\xi} [\dot{\psi}] \tilde{x}_{av} \\ &= \bar{\xi}^2 [\dot{\psi}_0] + \bar{\xi}^2 (\dot{\psi}^+ \dot{x}^- + \dot{\psi}^- \dot{x}^+) - 2\bar{\xi} [\dot{\psi}] \tilde{x}_{av} \\ &= D(\bar{H}) \bar{\xi}^2 - 2[\dot{\psi}] \tilde{x}_{av} \bar{\xi}. \end{aligned}$$

□

This lemma shows the correspondence between quadratic form (35) and that in [10]. Similar relations $\bar{\xi}_k = -\bar{t}_k$, $k = 1, \dots, s$ define a 1:1 correspondence between the elements of critical cone (34) and that in [10].

5 Conclusions and Perspectives

We study an optimal control problem, in Bolza form, of ordinary differential equations, considered on a nonfixed interval of time, subject to the initial–final time-state constraints of equality and inequality type and running mixed state-control constraints of equality type. We call this problem—the general problem of the calculus of variations, since it does not contain control (or mixed state-control) constraints of the *inequality* type which are characteristic to optimal control problems. The absence of such constraints considerably simplifies the study of the problem. For an extremal with piecewise continuous control, we define the notion of an extended weak minimum which differs from the notion of a weak minimum in that the first one takes

(additionally) into account the variations of discontinuity points of the control. Under the assumption that the gradients with respect to the control of running mixed constraints are linearly independent, we derive a second-order necessary condition for an extended weak minimum, for extremals with piecewise Lipschitz-continuous controls. The method of the proof is based on the study of relationship between necessary conditions for a weak and for an extended weak minimum, respectively. Let us briefly recall the stages and the logic of this proof.

First, using the Dubovitskii–Milyutin method of critical variations, we derive a second-order necessary condition for a weak local minimum in a problem on a fixed time interval. Then, we make a change of time variable which converts an extremal with a piecewise continuous control in the problem on a variable interval to an extremal in a problem on a fixed interval. We claim that if an extended weak minimum is attained on the first extremal (in the problem on a variable interval), then a weak local minimum is attained on the second extremal (in the problem on a fixed interval). Hence, the necessary second-order condition for a weak minimum in a problem on a fixed interval, applied to the transformed extremal in the transformed problem, is also necessary for an extended weak minimum in the initial problem. Further, we write down this necessary condition and transform it to the form, where the change of time variable does not participate any more, and everything is presented in terms of the initial problem. This transformation is the essence of the proof. As a result, we get the proof of the necessary second-order condition for an extremal with discontinuous control, which is much simpler and shorter than that in [5] (see also [11], Chapters 1–3).

It should be pointed out that, in [5] and [11], it was obtained a bit finer necessary second-order condition for an extended weak minimum. It had a smaller set of Lagrange multipliers than the set Λ_0^Θ . Let us define this smaller set.

An element $\lambda \in \Lambda_0^\Theta$ is said to be *Legendre* iff the following conditions hold:

- (a) $D^k(\bar{H}) \geq 0$ for all $t_k \in \Theta$, where $D^k(\bar{H})$ is as in (86);
- (b) for any $t \in [t_0, t_f] \setminus \Theta$ the inequality

$$\langle \bar{H}_{uu}(t, x(t), u(t), \psi(t), v(t))u, u \rangle \geq 0$$

holds for all $u \in \mathbb{R}^{d(u)}$ such that $g(t, x(t), u(t))u = 0$.

Here (b) is the well-known Legendre–Clebsch condition. Both conditions (a) and (b) follow from the Pontryagin minimum principle (see, e.g. [5] and [11]). Denote by $\text{Leg}(\Lambda_0^\Theta)$ the subset of all Legendre elements $\lambda \in \Lambda_0^\Theta$. Then Theorem 4.1 can be strengthened as follows (see [5] and [11]):

If the process \mathcal{T} yields a Θ -weak minimum, then the following Condition $\mathcal{A}'(\Theta)$ holds: the set $\text{Leg}(\Lambda_0^\Theta)$ is nonempty, and

$$\max_{\lambda \in \text{Leg}(\Lambda_0^\Theta)} \Omega_\Theta(\lambda, \tilde{z}) \geq 0 \text{ for all } \tilde{z} \in \mathcal{K}_\Theta.$$

Is there a simple proof of this *strengthened* version of Theorem 4.1? (The proof of this version, available in [5] and [11], is neither simple nor short.) It is worth noting that

just Condition $\mathcal{A}'(\Theta)$ (but not Condition $\mathcal{A}(\Theta)$) possesses a natural strengthening to a (no-gap) sufficient optimality condition (see, e.g. [11]).

We also mention some other open problems, related to this article.

1. We hope that the method of changing of time variable, which was used in the present paper, can be also used in optimal control problems with Volterra-type integral equations and can yield new second-order conditions for extremals with discontinuous controls in such problems.
2. Certainly, a similar approach can be used in problems with regular mixed state-control *inequality* constraints. The regularity of the constraints means that the gradients with respect to the control of the active inequalities are (positively) linearly independent. But the proof of the necessary second-order conditions for a weak minimum in such problems is much more difficult than that was demonstrated here for the case of mixed equality-type constraints.
3. The switchings of the optimal control are possible not only at the interior points of the time interval (as it was assumed in the present paper) but also at the ends of the interval. For example, in a time optimal control problem, the switching may happen at the final time t_f . More precisely, this means that, at the point t_f , there is a value of control u_f for which the (nonmaximized) Hamiltonian has the same value as for the optimal control $u(t_f)$ and $u_f \neq u(t_f)$. The concept of an extended weak minimum can be modified in a way that the values t_f and u_f have been taken into account. Which quadratic form corresponds to such modification of the minimum?
4. Is it possible to transform the second-order necessary optimality conditions obtained in [13] for extremals with discontinuous controls in the problem with finitely many internal constraints $\theta_i(t_k, x(t_k)) = 0, \vartheta_j(t_k, x(t_k)) \leq 0, k = 1, \dots, s$ ($t_0 < t_1 < \dots < t_s < t_f$) to such a form that their relation with conditions obtained in [6], [5] and [10] become obvious? Similar question arises with respect to the conditions obtained in Chapter 16 of book [12].

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Milyutin, A.A.: Quadratic conditions for an extremum in smooth problems with a finite-dimensional image. In: Levin, V.L. (ed.): *Metody teorii ekstremal'nykh zadach v ekonomike* (Methods of the theory of extremal problems in economics), pp. 138–177. Nauka, CEMI, Moscow (1981) [in Russian]
2. Dmitruk, A.V.: Quadratic conditions for a local minimum for singular extremals in a general optimal control problem. In: G. Ferreyra, et al. (eds.): *Proceedings of Symposia in Pure Mathematics*, vol 64 "Differential Geometry and Control", pp. 163–198. American Mathematical Society (1999)
3. Henrion, R.: *La Théorie de la Variation Seconde et ses Applications en Commande Optimale*. Palais des Academies, Academie Royal de Belgique, Bruxelles (1979)
4. Shnol', E.E.: On the degeneration in the simplest problem of the calculus of variations. *Mat. Zametki* **24**, 707–716 (1978)
5. Osmolovskii, N.P.: Quadratic optimality conditions for broken extremals in the general problem of calculus of variations. *J. Math. Sci.* **123**, 3987–4122 (2004)

6. Milyutin, A.A., Osmolovskii, N.P.: *Calculus of Variations and Optimal Control*. Translations of Mathematical Monographs, vol. 180. American Mathematical Society, Providence (1998)
7. 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)
8. 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)
9. Osmolovskii, N.P.: Higher-order necessary and sufficient conditions for Pontryagin and boundedly strong minima in the optimal control problem. *Dokl. Akad. Nauk SSSR* **303**, 1052–1056 (1988)
10. Osmolovskii, N.P.: Second order conditions in optimal control problems with mixed equality-type constraints on a variable time interval. *Control Cybern.* **38**, 1535–1556 (2009)
11. 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, Philadelphia (2012)
12. Hull, D.G.: *Optimal Control Theory for Applications*. Springer, New York (2003)
13. Dmitruk, A.V., Kaganovich, A.M.: Quadratic order conditions for an extended weak minimum in optimal control problems with intermediate and mixed constraints. *Discr. Contin. Dyn. Syst.* **29**, 523–545 (2011)
14. Dubovitski, A.Ya., Milyutin, A.A.: Problems for extremum under constraints. *Zh. Vychislit. Mat. i Mat. Fiz.* **5**, 395–453 (1965)
15. Dubovitski, A.Ya., Milyutin, A.A.: English transl. in: *U.S.S.R. Comput. Math. Math. Phys.* **5**, 1–80 (1965)
16. Milyutin, A.A., Dmitruk, A.V., Osmolovskii, N.P.: *Maximum Principle in Optimal Control*. Moscow State University, Moscow (2004). [in Russian]
17. Dmitruk, A.V., Milyutin, A.A., Osmolovskii, N.P.: Lyusternik’s theorem and the theory of extrema. *Uspekhi matem. nauk* **35**, 11–46 (1980)
18. Dmitruk, A.V., Milyutin, A.A., Osmolovskii, N.P.: English transl. *Russian Math. Surveys* **35**, 11–51 (1980)
19. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. *J. Res. Nat’l Bur. Standarts* **49**, 263–265 (1952)