QUADRATIC ORDER CONDITIONS OF A LOCAL MINIMUM FOR SINGULAR EXTREMALS IN A GENERAL OPTIMAL CONTROL PROBLEM

A. V. DMITRUK

Abstract

We consider the class of optimal control problems, linear in the control, with control bounded by linear inequalities, and with terminal equality and inequality constraints. Both the control and state variables are multidimensional, and the examined control is totally singular. For such problems we suggest quadratic-order necessary and sufficient conditions for a weak and a so-called Pontryagin minimum, the last being a minimum of an intermediate type between classic weak and strong minima. Necessary conditions transform into sufficient ones only by strengthening an inequality, what is similar to conditions in the classical analysis and calculus of variations (adjoint pairs of conditions).

Key words: singular extremal, weak minimum, Pontryagin minimum, strong minimum, quadratic order of estimation, necessary and sufficient conditions, critical cone, second variations, Legendre conditions.

1 Statement of the problem and preliminaries

Let us consider the control system:

$$\dot{x} = f_0(x,t) + F(x,t)u = f_0(x,t) + \sum_{i=1}^{k} u_i f_i(x,t), \tag{1}$$

where all the functions $f_0, \ldots, f_k$ are defined in an open set $\mathcal{D}_{x,t}$ in $R^n \times R$ and take values in R^n, they are assumed to be C^2-smooth in (x,t). Here $f_i(x,t)$, $i = 1, \ldots, k$, are the columns of the matrix $F(x,t)$.

We consider this system on a fixed time interval $[t_0, t_1]$. Denote $x_0 = x(t_0)$, $x_1 = x(t_1)$, and $p = (x_0, x_1)$, and consider the following optimal control problem A:

$$J = \varphi_0(p) \longrightarrow \min, \tag{2}$$

1991 *Mathematics Subject Classification.* Primary 49K15.

This research was supported by the Russian Foundation for Basic Research under grants 97-01-00135 and 96-15-96072.

$$\varphi_i(p) \le 0, \quad i = 1,\dots,\nu, \tag{3}$$

$$K(p) = 0, \tag{4}$$

$$u(t) \in U, \tag{5}$$

$$(x(t),t) \in \mathcal{D}_x, \qquad p \in \mathcal{D}_p, \tag{6}$$

where $\varphi_0,\dots,\varphi_\nu$, K are C^2-smooth functions (K is multidimensional), defined in an open set $\mathcal{D}_p$ in R^{2n}, and U is a closed convex solid set in R^k. The sets $\mathcal{D}_{x,t}$ and $\mathcal{D}_p$ are usually just implied, but not indicated explicitly.

Without loss of generality we may put $t_0 = 0$ and $t_1 = T$, denoting then $x(t_1) = x_T$. Note a well known important property of system (1) (see e.g. [41, 33]).

Theorem 1.1. *Let $u_m \in L^k_\infty[0,T]$, and x_m be the solution of (1) for u_m with an initial condition $x_m(0) = a_m$. Let $u_m \overset{\text{weak-*}}{\longrightarrow} u_0 \in L^k_\infty[0,T]$, $a_m \to a_0$, and x be the solution of (1) for u_0 with the initial condition $x_0(0) = a_0$.*

Then $x_m \rightrightarrows x_0$ (uniformly).

In Problem A we will seek a minimum among all absolutely continuous n-vector functions $x(t)$ and bounded measurable k-vector functions $u(t)$. We do not study here the problem of existence of minimum. At least we can say, that if U is a convex compactum, x_0 is fixed, every solution $x(t)$ of system (1) entirely lies in $\mathcal{D}_{x,t}$, and its endpoints (x_0, x_T), satisfying (3), (4), belong to $\mathcal{D}_p$, then the existence follows from the known Filippov's lemma [42], which, in its turn, follows from the theorem of Alaoglu and Theorem 1.1.

Introduce the space $W = AC^n \times L^k_\infty[0,T]$ with elements $w = (x,u)$, and equip it with the norm $||w|| = |x(0)| + ||\dot x||_1 + ||u||_\infty$. (Here and throughout the paper $||\cdot||_p$ stands for the norm in the Banach space $L^k_p[0,T]$, and $||\cdot||_C$ stands for the norm in the space of continuous n-vector functions $C^n[0,T]$.) One can easily see, that for solutions of (1) the convergence of a sequence w_m to a point w_0 in the norm of W is equivalent to its convergence to w_0 in the norm $||w||_\infty = ||x||_C + ||u||_\infty$, hence a local minimum w.r.t. the norm of W is a common weak minimum. As usual, by a strong minimum we call a minimum w.r.t. the seminorm $||w||' = ||x||_C$, the u being free.

Now we introduce also another type of minimum, intermediate between the weak and strong minima. Let $\hat w = (\hat x, \hat u)$ be an admissible pair.

Definition 1.1. We say that $\hat w$ is a Pontryagin minimum point in Problem A, if for all N there exists an $\varepsilon > 0$ such that $\hat w$ is a minimum point in Problem A on the set

$$||x - \hat x||_C < \varepsilon, \quad ||u - \hat u||_1 < \varepsilon, \quad ||u - \hat u||_\infty \le N. \tag{7}$$

In other words, there can exist no sequence $w_m = (x_m, u_m)$ such that

$$||x_m - \hat{x}||_C \to 0, \quad ||u_m - \hat{u}||_1 \to 0, \quad ||u_m - \hat{u}||_\infty \le O(1), \tag{8}$$

all constraints are satisfied, and for all $m \quad J(p_m) < J(\hat{p})$.

Sequences that satisfy (8), we call *Pontryagin sequences,* converging to $\hat{w}$. The set of all such sequences we denote by $\Pi(\hat{w})$. The differences $\delta w_m = w_m - \hat{w}$ are Pontryagin sequences, converging to zero; we call them Pontryagin variations. Thus, the Pontryagin minimum (briefly, the Π-minimum) is the minimum in the class of all Pontryagin sequences (or, respectively, in the class of all Pontryagin variations). This type of minimum was introduced by A.Ja.Dubovitskii and A.A.Milyutin as a natural extension of minimum in the class of uniformly small and needle-type variations of the control. The importance of this notion is that the Pontryagin Maximum Principle (MP) is a necessary condition of the first order for the Pontryagin minimum, i.e. for the minimum in the class of Pontryagin variations. (It is said often that the MP is a necessary condition for the *strong* minimum, but this assertion is weaker than the above one.) Moreover, Dubovitskii and Milyutin showed [3], [10], that the fulfillment of MP for a trajectory is *equivalent* to its stationarity in the class of all possible Pontryagin variations. As we told already, the following relations for these three types of minimum are true:

$$\text{weak min} \;<\; \text{Pontryagin min} \;<\; \text{strong min},$$

and both the "inequalities" are strict: it is not difficult to provide corresponding counterexamples, see e.g. [33]. In particular case, when the admissible control set U is bounded, the Π-minimum is in fact the minimum w.r.t. the norm $||w||_1 = ||x||_C + ||u||_1$, and we call it the L_1-minimum w.r.t. the control.

The Π-minimum for Problem A is, generally, rather far from the strong minimum (unlike the case of problems with the general control system, nonlinear in the control, where these two types of minimum are rather close to each other; see [6], [7]). However, note here a special case, when these two types of minimum are equivalent: it is when U is bounded, and $(\forall\, t)\ \hat{u}(t)$ is an extreme point of U. (The set of all extreme points of U is denoted by $ex\, U$.)

Proposition 1.1. *Let U be an arbitrary convex compactum, and $\hat{u}(t) \in ex\, U$ a.e. Suppose also that $\operatorname{rank} F(\hat{x}(t), t) = k \quad \forall t$, i.e. the vectors $f_i(\hat{x}(t), t), \quad i = 1, ..., k$ are linearly independent. Then the Π -minimum is equivalent to the strong minimum.*

The proof readily follows from the next

Lemma 1.1. *Let $\hat{x}, \hat{u}$ satisfy equation (1), $x_m \Rightarrow \hat{x}$, $u_m \in U$, U is a convex compactum, and $\hat{u}(t) \in exU$ a.e. Then $||u_m - \hat{u}||_1 \longrightarrow 0$.*

The proof follows from another two lemmas.

Lemma 1.2. *Under the first four conditions of Lemma 1.1, $u_m \overset{\text{weak-*}}{\longrightarrow} \hat{u}$ (as elements of L_∞ w.r.t. L_1).*

Proof. Since U is bounded, then due to the theorem of Alaoglu, passing if necessary to a subsequence, one can consider $u_m(t) \overset{\text{weak-*}}{\longrightarrow} u_*(t)$. Since U is closed and convex, $u_*(t) \in U$ a.e. Let x_* be the solution of system (1) for this u_* with the initial condition $x_*(0) = \hat{x}(0)$. By Theorem 1.1 $x_m \Rightarrow x_*$, whence due to the uniqueness of the limit we have $x_* = \hat{x}$, thus $\hat{x}$ is the solution of (1) for u_* with the initial value $\hat{x}(0)$. But then, since the vectors $f_i(\hat{x}(t), t)$, $i = 1, ..., k$ are linearly independent, the value of control is uniquely determined by $\hat{x}(t)$, and so $u_*(t) = \hat{u}(t)$ a.e. Thus, $u_m \overset{\text{weak-*}}{\longrightarrow} \hat{u}$ (for the chosen subsequence; hence, for any subsequence of the initial sequence there is a sub-subsequence with this property, hence it is true for the whole initial sequence).

Lemma 1.3. *Let U be a convex compactum, $u_m(t) \in U$, and $u_m(t) \overset{\text{weak-*}}{\longrightarrow} \hat{u}(t)$, where $(\forall\, t)$ $\hat{u}(t) \in ex\, U$. Then $||u_m - \hat{u}||_1 \longrightarrow 0$.*

Proof. Let $p(t)$, $|p(t)| = 1$ $\forall t$, be a supporting vector for U at the point $\hat{u}(t)$, i.e.

$$(p(t), U) \leq (p(t), \hat{u}(t)). \tag{9}$$

(Such a $p(t)$ does exist by the measurable selection theorem [43].) Since $\int p(t)(u_m - \hat{u})dt \to 0$, then from (9), from the extremality of $\hat{u}(t)$ and the boundedness of U one can show that $\int |u_m - \hat{u}|dt \to 0$. To understand this effect, one may consider, for example, $U = \{u \in R^2 |\ u_1^2 \leq u_2\}$, $\hat{u} = (0,0)$, $p(t) = (0,-1)$. If $\int p u_m dt = \int u_{2,m} dt \to 0$, then $\int u_{1,m}^2 dt \to 0$, whence obviously $\int |u_{1,m}|dt \to 0$ too, i.e. $\int (|u_{1,m}| + |u_{2,m}|)dt \to 0$, q.e.d. The general case involves a bit more technical details, see [41].

2 Maximum Principle and singular controls

Let $\hat{w} = (\hat{x}, \hat{u})$ be a Π-minimum point in Problem A. Then it satisfies the Pontryagin Maximum Principle (MP), which says, that there exist Lagrange multipliers $\alpha = (\alpha_0, \dots, \alpha_\nu) \geq 0$, $\beta \in R^{d(K)}$, and a Lipschitz $n-$vector function $\psi(t)$ (the adjoint, or costate variable), the collection of which we denote by $\lambda = (\alpha, \beta, \psi)$, and which generate the terminal Lagrange function $l[\lambda](p) = \alpha \cdot \varphi(p) + \beta \cdot K(p)$, where $\varphi = (\varphi_0, \ldots, \varphi_\nu)$, and the Pontryagin function

$H[\lambda](x,u,t) = \psi[f_0(x,t) + F(x,t)u] = (\psi, f_0(x)) + \sum u_i(\psi, f_i(x))$,

such that the following conditions hold:

a) normalization condition: $|\alpha| + |\beta| = 1$ (here $|\cdot|$ is an arbitrary norm in the finite-dimensional space),

b) complementary slackness: $\alpha_i \varphi_i(\hat{p}) = 0, \quad i = 1, ..., \nu,$
(in the sequel we assume, without loss of generality, that $\varphi_i(\hat{p}) = 0, \quad \forall\, i = 1, ..., \nu,$ i.e. all indices i are active, and we denote $I = \{0, 1, ..., \nu\}$),

c) adjoint, or costate equations: $\dot{\psi} = -H_x[\lambda], \qquad \dot{H}[\lambda] = H_t[\lambda],$

d) transversality conditions: $\psi(0) = l'_{x_0}[\lambda], \quad \psi(T) = -l'_{x_T}[\lambda],$

e) maximality condition: for all t

$$\max_{u \in U} H[\lambda](\hat{x}(t), u, t) = H[\lambda](\hat{x}(t), \hat{u}(t), t). \tag{10}$$

The set of all $\lambda = (\alpha, \beta, \psi)$, satisfying conditions (a)–(e) for the trajectory $\hat{w}$, we denote by $\Lambda(\hat{w})$, or, having in mind that the trajectory $\hat{w}$ will be the same throughout the paper, simply by Λ. Obviously, Λ is a finite-dimensional compact set, and generally it may consist of more than a single point. (In particular case, when Λ consists of a single point, we will write $\Lambda = \{\cdot\}$.) We say that the trajectory $\hat{w}$ is stationary, or extremal, if $\Lambda(\hat{w})$ is nonempty. The MP guarantees that if $\hat{w}$ is a Π-minimum point in Problem A, then the set $\Lambda(\hat{w})$ is nonempty, i.e. $\hat{w}$ is an extremal.

It is well known, however, that the MP is only a necessary, but not a sufficient condition (like any other first order necessary condition for general nonconvex problems): its fulfillment does not guarantee even a weak minimum at $\hat{w}$. Especially it is true for our Problem A . Because of this, many authors made investigations on higher order (mostly "second order") conditions of a local minimum (for some particular statements of Problem A) since early 1960-s; we just mention here Kelley, Kopp, Moyer, Bryson, Robbins, Goh, Vapnyarsky, Speyer, Jacobson, Bell, McDanell, Powers, Gabasov, Kirillova, Krener, Agrachiov, Gamkrelidze, Milyutin, Dmitruk, Knobloch, Zelikin, Gurman, Dykhta, Lamnabhi-Lagarrigue, Stefani, Sarychev, and others, see [4, 12], [17] – [39], and references therein. These more fine conditions require a more fine specification of examined extremals. There are two essentially different classes of extremals: singular and nonsingular extremals. The first one of them includes, as the most pronounced case, the class of totally singular extremals.

Definition 2.1. An extremal $\hat{w}$ is called *totally singular,* if for any $\lambda \in \Lambda(\hat{w})$, and for any t the Pontryagin function $H[\lambda](\hat{x}(t), u, t)$ does not depend on $u \in U$, i.e. takes the same value $H[\lambda](\hat{x}(t), \hat{u}(t), t)$ for all $u \in U$.

Since H is linear in u, and U has nonempty interior, this means simply, that for all t

$$H_u[\lambda](\hat{x},\hat{u},t) = \psi(t)F(\hat{x}(t),t) = 0, \tag{11}$$

or, equivalently,

$$(\psi(t), f_i(\hat{x}(t),t)) = 0 \quad \forall\, i = 1, ...k.$$

Remark 2.1. Perhaps, it would be more proper to say in this case, that the constraint $u \in U$ (or the set U) is totally singular for the extremal $\hat{w}$ (rather than the extremal $\hat{w}$ itself is totally singular), because U does not actually enter the MP at all.

Remark 2.2. Note that if U is totally singular, then any other U', containing U, is totally singular as well, and the set $\Lambda(\hat{w})$ is one and the same for all such sets U', and coincides with the set $\Lambda(\hat{w})$ for $U = R^k$, i.e. for Problem A with unconstrained control. However, if one takes U' containing $\hat{u}(t)$, but not containing U, it may happen to exist $\lambda \in \Lambda(\hat{w}, U')$, for which (11) fails.

If there exists at least one $\lambda \in \Lambda$, which does not satisfy (11), then $\hat{w}$ is not totally singular. The most distinct case is when there exists $\lambda \in \Lambda$, for which $\forall t$ $H[\lambda](\hat{x}(t), u, t)$ attains its maximum over U only at the single point $\hat{u}(t)$. In this case we call $\hat{w}$ *strictly nonsingular.* In the intermediate case, when $\hat{w}$ is neither totally singular, nor strictly nonsingular, it is of a mixed type. This case practically has not been studied in the framework of higher order conditions of a local minimum. (Only conditions at junction times of singular and nonsingular subarcs have been analyzed until now, see e.g. [40].) The main efforts have been spent to investigate the two principal cases of totally singular and strictly nonsingular extremals. The higher order conditions for these two cases turned out to be essentially different. In this paper we consider only the totally singular case. Note that in the classical calculus of variations (CCV) this case had not been studied, because all considerations in CCV were made under the assumption of strong Legendre condition: $-H_{uu}[\lambda] \geq \text{const} > 0$, whereas for totally singular extremals this coefficient is identically zero.

Now let $\hat{w}$ be a totally singular extremal. We assume that the control $\hat{u}(t)$ is continuous. In view of equations (11), this is rather a mild requirement, because in a generic case it is possible, by differentiating these equations two times in t, to express $\hat{u}(t)$ in terms of $\psi(t)$ and $\hat{x}(t)$ (see e.g. [17, 19, 20]), whence $\hat{u}(t)$ is in fact Lipschitzian.

Thus, we have taken two assumptions about the examined extremal:

Assumption 1. $\hat{w} = (\hat{x}, \hat{u})$ is a totally singular extremal.

Assumption 2. The control $\hat{u}(t)$ is continuous.

Now we impose an assumption on the character of contact of the control $\hat{u}(t)$ with the boundary ∂U of the admissible control set U.

Assumption 3. U is a closed polyhedral set (may be unbounded), and there exists a face U_0 of U, such that $\forall t \quad \hat{u}(t) \in \operatorname{reint} U_0$. (By *reint* we denote the relative interior of the convex set.)

In particular case, $U_0 = U$ is allowed, which means that $\forall t \;\; \hat{u}(t) \in \operatorname{int}\, U$. If U_0 is a proper face of U, then $\hat{u}(t) \in \partial\, U$, and we call it the simplest case of boundary control.

Assumption 3 can be weakened to the following one.

Assumption 3'. $U = M \cap D$, where M satisfies Assumption 3, and D is an arbitrary convex set, such that $\forall t \;\; \hat{u}(t) \in \operatorname{int} D$.

3 Quadratic order of estimation

As was said already, having got first order conditions for a local minimum, it is natural to move to study "second order" conditions. But here we have to define preciser, what will be meant by "second order" conditions. In finite-dimensional problems it is quite clear: they are conditions of the order $|\delta x|^2$, i.e. when all considerations are made modulo $o(|\delta x|^2)$. These conditions possess the following important properties: a) both necessary and sufficient conditions withstand perturbations of all data functions in the problem within $o(|\delta x|^2)$,
b) the sufficient condition consists of a lower bound of second variations by $|\delta x|^2$,
c) if a point satisfies the necessary condition, one can make an arbitrary small perturbation of the data functions in the C^2-norm, after which this point satisfies the sufficient condition.

However, in infinite-dimensional problems, such as CCV and optimal control problems, the situation is not so simple, and it is not clear a'priori, what is the "second order". If one undertake the straightforward generalization of the finite-dimensional ideology, i.e. perform all considerations w.r.t. $\gamma_0(\delta w) = ||\delta w||^2, \;\; \delta w = (\delta x, \delta u)$, then unfortunately one would never obtain the fulfillment of sufficient conditions (apart from the case of Hilbert space, which is not characteristic neither for CCV, nor for optimal control): it is well known, that a continuous quadratic form can be bounded from below by the square of the norm only if the space is isomorphic to a Hilbert space. The functional γ_0 can be said to be "too rough" to estimate second variations in these classes of problems. (Note by the way, that in CCV the term "second order conditions" was never used. In those times the authors said "conditions, based on the second variation", see e.g. [1]. Obviously, they understood that the term "second order conditions" is not a proper one for problems in functional spaces.)

Thus, the question arises: *what a functional should be taken instead of* $\gamma_0(\delta w)$, in such a way that, roughly speaking, it would preserve the abovementioned properties of γ_0 ? The answer depends on the class of problems being considered, and may also depend on the type of examined trajectory. It was discovered, due to a series of deep works by A.A.Milyutin and his scientific school [5] – [9], [31] – [38], that for any properly defined class of extremal problems there should exist an estimating functional $\gamma(\delta w)$, positive outside of zero (which we call an order of estimation, or simply an order), characteristic to this class. (See precise definitions and details in [5].) For general optimal control problems, nonlinear in the control (e.g. for CCV), in case when $\hat{u}(t)$ is continuous, such an order is:

$$\gamma_{clas}(\delta w) = |\delta x(0)| + \int |\delta u(t)|^2 dt$$

(here and throughout the paper all integrals without limits are taken over the whole interval $[0,T]$). The corresponding necessary and sufficient conditions (in particular, those in CCV) possess the above properties a) – c) [5]. However, it does not suit to our Problem A: from the very outset one can say, that, since the control comes just linearly in the problem, the second variations in this problem will never be bounded from below by this $\gamma_{clas}(\delta w)$, because they never contain the term with δu^2.

It turns out that the proper quadratic functional of estimation for Problem A is:

$$\gamma(\delta w) = |\ \delta x(0)\ |^2 + |\ \delta y(T)\ |^2 + \int |\ \delta y(t)\ |^2\, dt, \tag{12}$$

where

$$\delta \dot{y} = \delta u, \qquad \delta y(0) = 0.$$

Note that the control variation δu does not come as such in the quadratic order (12); it comes only through the variation of a new state variable y, where $\dot{y} = u,\ y(0) = 0$. (This relation between y and u will be preserved throughout the paper). Here we give conditions of this order γ.

4 The second and third variations, and the critical cone

For any $\lambda \in \Lambda$ consider the corresponding Lagrange function

$$\Phi[\lambda](w) = \; l[\lambda](p) + \int ((\psi, \dot{x}) - H[\lambda](x, u, t))\, dt$$

(recall that $l[\lambda]$ and $H[\lambda]$ were defined in the beginning of Sec.2), and consider also the half of its second variation at $\hat{w}$ - the quadratic functional

$$\Omega[\lambda](\bar{w}) = \frac{1}{2}(l''[\lambda]\bar{p}, \bar{p}) - \int \left((\frac{1}{2}H_{xx}[\lambda]\, \bar{x}, \bar{x}) + (\bar{x}, H_{xu}[\lambda]\, \bar{u}) \right) dt. \tag{13}$$

Define the matrices $A(t) = f_0'(\hat{x}(t), t) + F'(\hat{x}(t), t)\hat{u}$, $\quad B(t) = F(\hat{x}(t), t)$, and the tensor $R(t) = F'(\hat{x}(t), t)$ (by f_i' we denote the derivative of f_i w.r.t. x) in such a way that for $x = \hat{x} + \delta x$ and $u = \hat{u} + \delta u$ equation (1) takes the form:

$$\delta\dot{x} = A(t)\delta x + B(t)\delta u + (R(t)\delta x, \delta u) \; + \text{ higher order terms.} \tag{14}$$

Denote by $\mathcal{K}$ the so-called critical cone for the problem (1)–(4) with the free control. It consists of all variations $\bar{w} = (\bar{x}, \bar{u})$ in W, such that

$$\varphi'(0)\bar{p} \le 0, \qquad K'(0)\bar{p} = 0, \tag{15}$$

and

$$\dot{\bar{x}} = A(t)\bar{x} + B(t)\bar{u}. \tag{16}$$

The cone of critical variations for the "full" Problem A is $\mathcal{K} \cap \mathcal{N}$, where $\mathcal{N} = \{\bar{w}|\; \bar{u}(t) \in N \; a.e.\}$, and $N = con(U - \hat{u}(t)) = con(M - \hat{u}(t))$ is the local (pointwise) tangent cone for the polyhedron M at the point $\hat{u}(t)$. (By $con\, M$ we denote the conical hull $\bigcup \{\alpha M \mid \alpha > 0\}$ of the set M.) Due to Assumption 3 (or 3') N does not depend on t.

We also introduce the cubic functional

$$\rho[\lambda](\bar{w}) = \int \left[-(\frac{1}{2}H_{uxx}[\lambda]\bar{x}, \bar{x}, \bar{u}) + (H_{xu}[\lambda]\bar{y}, (R(t)\bar{x}, \bar{u})) \right] dt. \tag{17}$$

It is one sixth of the third variation of Lagrange function at $\hat{w}$ (in case when all $f_i(x,t)$ are C^3–smooth in x) on equation (1) to within $o(\gamma)$ on Pontryagin sequences, see [32, 33] for details.

It is convenient to consider functionals (13) and (17) in slightly transformed variables. Namely, there is a convenient change of variables - the so-called Goh transformation [18, 19]: $(\bar{x}, \bar{u}) \longmapsto (\bar{\xi}, \bar{y}, \bar{u})$, where $\bar{\xi} = \bar{x} - B\bar{y}$,

$$\dot{\bar{y}} = \bar{u}, \qquad \bar{y}(0) = 0, \tag{18}$$

and hence

$$\dot{\bar{\xi}} = A(t)\bar{\xi} + B_1(t)\bar{y}, \qquad B_1 = AB - \dot{B}. \tag{19}$$

This transformation is convenient because, instead of the state variable $\bar{x}$, related to $\bar{u}$ by a general linear system (16), there are now two state variables $\bar{\xi}, \bar{y}$, such that $\bar{u}$ does not come into equation (19) for $\bar{\xi}$, and comes, in the simplest way, only into equation (18) for $\bar{y}$.

In these new variables, the second variation (13) takes the form:

$$\Omega[\lambda](\bar{\xi}, \bar{y}, \bar{u}) \; = \; g[\lambda](\bar{\xi}_0, \bar{\xi}_T, \bar{y}_T) + \tag{20}$$
$$+ \int ((D[\lambda](t)\bar{\xi}, \bar{\xi}) + (P[\lambda](t)\bar{\xi}, \bar{y}) + (Q[\lambda](t)\bar{y}, \bar{y}) + (V[\lambda](t)\bar{y}, \bar{u}))dt,$$

where $g[\lambda]$ is a terminal quadratic form, $Q[\lambda](t)$ is a symmetric and $V[\lambda](t)$ is a skew-symmetric Lipschitz matrices. The term $(G[\lambda]\bar{\xi}, \bar{u})$ has been taken by parts in view of (19) and (18).

Putting in (17) $\bar{x} = \bar{\xi} + B\bar{y}$, reduce $\rho[\lambda]$ in the new variables to the form: $\rho[\lambda] = \theta[\lambda] + \eta[\lambda]$, where

$$\theta[\lambda](\bar{\xi}, \bar{y}, \bar{u}) = \int ((T_1[\lambda](t)\bar{\xi}, \bar{\xi}, \bar{u}) + (T_2[\lambda](t)\bar{\xi}, \bar{y}, \bar{u}))dt,$$

$$\eta[\lambda](\bar{\xi}, \bar{y}, \bar{u}) = \int (\mathcal{E}[\lambda](t)\bar{y}, \bar{y}, \bar{u}))dt, \tag{21}$$

the tensor $\mathcal{E}[\lambda]$ is obtained from (17) by substituting $B\bar{y}$ instead of $\bar{x}$. It can be easily shown (see e.g. [33]) that on any Pontryagin sequence we have $\theta[\lambda] = o(\gamma)$, so the only essential term in $\rho[\lambda]$ is $\eta[\lambda]$.

The cone $\mathcal{K}$ in the new variables is given by equations (18), (19) and terminal relations (15), in which one should put $\bar{p} = (\bar{x}_0 = \bar{\xi}_0,\ \bar{x}_T = \bar{\xi}_T + B_1(T)\bar{y}_T)$.

At last, introduce some notations concerning the local cone N. Let H_2 be the maximal subspace in N, and H_1 be its complement, i.e. $R^k = H_1 \oplus H_2$. Then $N_1 = N \cap H_1$ is a pointed cone in the subspace H_1, and $N = N_1 \oplus H_2$. If one consider H_1, H_2 to be coordinate subspaces, then any $u \in N$ can be uniquely represented in the form $u = (u_1, u_2)$, where $u_1 \in N_1,\ u_2 \in H_2$.

Now we are ready to formulate conditions of the order γ for the presence of weak and Π- minima at $\hat{w}$.

5 Conditions for a weak minimum

We begin with a general description of the quadratic order conditions. For both weak and Pontryagin minimum these conditions are of the same form. To be precise, for any $a \in R$ we define a subset $M_a \subset \Lambda$, nonincreasing while a increases, and define the corresponding functional

$$\Omega[M_a](\bar{w}) = \sup_{\lambda \in M_a} \Omega[\lambda](\bar{w}) \tag{22}$$

(the sup over $\emptyset$, as usually, being equal to $+\infty$), and then conditions of minimality are as follows:

necessary condition:

$$\Omega[M_0](\bar{w}) \geq 0 \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N},$$

sufficient condition: $\exists\, a > 0$, such that

$$\Omega[M_a](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}.$$

The difference between conditions for a weak minimum and conditions for a Π-minimum is only in the definition of the set M_a. This set consists of all $\lambda \in \Lambda$, for which the Lagrange function satisfies some special pointwise conditions. (Any pointwise conditions we call *conditions of Legendre type*). For a $\Pi-$minimum these pointwise conditions are more restrictive, hence the set M_a is smaller, and the corresponding necessary and sufficient conditions are stronger.

Let us now formulate these pointwise conditions. We begin with a weak minimum. Here one can take as M_a the entire Λ, but it is possible to choose a more narrow set, thus giving more strong necessary conditions, and more simple sufficient ones.

For any $a \in R$ denote by $G_a(\Lambda)$ the set of all $\lambda \in \Lambda$, such that the corresponding second variation $\Omega[\lambda]$ satisfies the following three conditions: $\forall t \in [0,T]$, $\forall\, \bar{h} \in H_1$, and $\forall\, \bar{u}, \bar{v} \in H_2$

$$\begin{aligned} i)&\ \ (V[\lambda](t)\bar{u}, \bar{h}) = 0, \\ ii)&\ \ (V[\lambda](t)\bar{u}, \bar{v}) = 0, \\ iii)&\ \ (Q[\lambda](t)\bar{u}, \bar{u}) \geq a|\bar{u}|^2. \end{aligned} \tag{23}$$

We introduce also the set $G_+(\Lambda) = \bigcup_{a>0} G_a(\Lambda)$.

Remark 5.1. For the case when $\Lambda = \{\cdot\}$, and $\hat{u}(t) \in \operatorname{int} U$ (i.e. $H_2 = R^k$), conditions *(ii)* and *(iii)* with $a = 0$ were obtained by B.S.Goh [18] as necessary conditions for a weak minimum. Note that if *(ii)* holds, the quadratic form (20) does not contain the control u, and then, taking into account (19), the variable y can be regarded as *a new control,* whence condition *(iii)* is just the classical

Legendre condition w.r.t. this new control y. Condition *(i)* is due to the author [38].

Before the formulation of the conditions for a weak minimum, let us recall the following

Definition 5.1. We say that the equality constraints (1) and (4) satisfy Lyusternik condition at $\hat{w}$, or that they are mutually nondegenerate at $\hat{w}$, if the operator

$$g: \quad w = (x, u) \longmapsto (\dot{x} - f_0(x,t) - F(x,t)u,\ K(x_0, x_T)),$$

mapping W to $Z = L_1^n[0,T] \times R^{\dim K}$, has a surjective derivative at $\hat{w}$, i.e. $g'(\hat{w})W = Z$.

Theorem 5.1. *i) Let $\hat{w}$ be a weak minimum point in Problem A. If $U_0 \neq U$, i.e. $\hat{u}(t) \in \partial U$, suppose also that the equality constraints (1) and (4) satisfy Lyusternik condition at $\hat{w}$. Then $G_0(\Lambda)$ is nonempty, and*

$$\Omega[G_0(\Lambda)](\bar{w}) \geq 0 \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{24}$$

ii) Let $G_0(\Lambda)$ be nonempty, and for some $a > 0$

$$\Omega[G_0(\Lambda)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{25}$$

Then $\hat{w}$ is a strict weak minimum point in Problem A.

iii) From (25) it follows, that $G_a(\Lambda)$ is nonempty, and

$$\Omega[G_a(\Lambda)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{26}$$

The proof is based on a general theory of higher order conditions for a local minimum in extremum problems with constraints [5]. For the case $U_0 = U$, i.e. $\hat{u} \in \operatorname{int} U$, it is given in [4], [31], for the case $\hat{u} \in \partial U$ it is due to A.A.Milyutin, see [37]. Note that the nondegeneracy of the equality constraints is required only in the necessary conditions, and only in the case of boundary control.

Definition 5.2. We say that $\hat{w}$ is a weak γ-minimum point in Problem A, if condition (25) or (26) is fulfilled for some $a > 0$.

We also give two equivalent formulations of these conditions. Recall that U satisfies Assumption 3$'$. Let the polyhedron $M \subset R^k$ be given by the inequalities:

$$(a_s, u) + b_s \leq 0, \quad s = 1, ..., \hat{s}. \tag{27}$$

Denote by $S_0 = S_0(\hat{u})$ the set of all active indices s for $\hat{u}(t)$, i.e. such s, that the inequality turns into equality. Again due to Assumption 3$'$ it does not depend on t. Note by the way, that the cone $N = con(U - \hat{u})$ is given by the inequalities $(a_s, \bar{u}) \leq 0, \quad s \in S_0$.

Now let us introduce the so-called violation function:

$$\sigma(w) = \sum_{i \in I} \varphi_i^+(p) + |K(p)| +$$

$$+ \int |\dot{x} - f(x,t) - F(x,t)u| dt + \sum_{s \in S_0} \text{ess sup } [(a_s, u(t)) + b_s]^+.$$

(We use the conventional notation $z^+ = \max\{0, z\}$.)

Theorem 5.2. *i) Conditions (25) and(26) are equivalent to the following one: there exists a neighborhood of $\hat{w}$ in W, in which the violation function has the lower bound:*

$$\sigma(w) \geq a'\gamma(w - \hat{w}), \qquad a' > 0. \tag{28}$$

ii) Suppose that equality constraints (1) and (4) satisfy Lyusternik condition at $\hat{w}$. Then conditions (25) and (26) are equivalent to the following one: $\exists\, a' > 0$, such that inequality (28) holds for any w from a neighborhood of $\hat{w}$ in W, satisfying equations (1) and (4) with $u \in U$. (Note that for such w the σ reduces to the first term only).

Condition *(i)* is proved in the general theory [5], the proof of *(ii)* is given in Appendix.

Let us dwell for a moment on condition (28). It obviously implies that $\hat{w}$ is a strict weak minimum point, thus it is a sufficient condition for a weak minimum. The equivalence of (25) and (26) to (28) has, at first look, rather an abstract character. However, it turned out to be quite an effective tool of investigation. (It will be essentially used in a forthcoming paper, devoted to application of the general conditions to abnormal sub-Riemannian geodesics). It worth noting also, that condition (28) has a "nonvariational" character: it does not include neither critical variations, nor Lagrange multipliers; it includes just the functions from the statement of Problem A as they are, not having been undergone any differentiations or approximations.

Checking conditions (24) and (25), i.e. checking the functional $\Omega[G_0(\Lambda)](\bar{w})$ (the maximum of a family of quadratic forms) for the nonnegative or positive definiteness on $\mathcal{K} \cap \mathcal{N}$, is in itself rather a nontrivial problem. It is known, that for a *single quadratic form* the question of its nonnegative or positive definiteness on a *subspace* can be solved by means of the classical Jacobi theory, which gives an answer in terms of the absence of conjugate (or focal) points. For a quadratic form on a *cone*, the more so for a *maximum of quadratic forms*, this question is much more difficult, because here the Euler-Jacobi equation is no longer linear (the difference of two solutions is not a solution). However, for this nonclassical situation there is an analogue of Jacobi theory [14], [15], [16], which establishes the equivalence of the positive definiteness of Ω to the absence of conjugate (focal) points in [0,T]. But for the nonnegativity of Ω, generally, only the one-way implication is true: $\Omega \geq 0$ implies that there are no conjugate points in (0,T). In the case of a single quadratic form, being considered on a polyhedral cone (hence the local cone must be absent: $N = R^k$, only terminal inequalities are allowed), the reverse implication is true as well [16], and so the Jacobi theory for this case is in a sense complete. In the general case there are counterexamples to this reverse implication [14], [15].

6 Sufficient conditions for small time intervals

There is an important case, when the above question of checking the quadratic functional for the positive definiteness has a simple solution. It is the case when the length of time interval is small enough. Recall that in CCV the positive definiteness of a quadratic functional on small time intervals is guaranteed by the strengthened Legendre condition, provided that at least one of the endpoints in the initial problem is fixed, i.e. one of the endpoint variations equals to zero ($\bar{x}(0) = 0$ or $\bar{x}(T) = 0$). A similar fact holds true for a quadratic functional of the form (20) on an arbitrary cone $\mathcal{K}$, contained in the subspace given by equations (19), (18).

6.1 A general theorem

For any interval $\Delta = [t_0, t_1] \subset [0, T]$ let us consider a quadratic functional of the type

$$\Omega(\bar{\xi}, \bar{y}, \bar{u}) = g(\bar{\xi}_0, \bar{\xi}_1, \bar{y}_1) + \tag{29}$$

$$+ \int_\Delta \left((D(t)\bar{\xi}, \bar{\xi}) + (P(t)\bar{\xi}, \bar{y}) + (Q(t)\bar{y}, \bar{y}) + (V(t)\bar{y}, \bar{u}) \right) dt,$$

where g is a terminal quadratic form, and matrices D, P, Q, V are defined and uniformly bounded on $[0, T]$. We also consider the subspace $\mathcal{L}_\Delta$, consisting of all $\bar{w} = (\bar{\xi}, \bar{y}, \bar{u})$ satisfying (19) and (18) a.e. on Δ, and the cone $\mathcal{N} = \{\bar{w} \mid \bar{u}(t) \in N \ \ a.e. \text{ on } \Delta\}$.

Theorem 6.1. *Suppose that* $\exists\, C,\ \delta_0 > 0$ *such that* $\forall \Delta,\ |\Delta| \leq \delta_0$, *the following estimates hold on* $\mathcal{L}_\Delta$*:*

$$||\bar{\xi}||_\infty \leq C||\bar{y}||_1, \qquad |\bar{y}(t_1)| \leq C||\bar{y}||_1, \tag{30}$$

and let $N = N_1 \oplus H_2$, *where* N_1 *is a pointed cone in a subspace* H_1, *and* $H_1 \oplus H_2 = R^k$. *Suppose further that quadratic form (29) satisfies on [0, T] the above three* *conditions (23) with $a > 0$. Then there exists $\delta > 0$, depending on the numbers C, δ_0, a, and on the coefficients of Ω, such that if $\Delta \subset [0,T]$ and $|\Delta| \leq \delta$, then*

$$\Omega(\bar{w}) \geq \frac{a}{2}\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{L}_\Delta \cap \mathcal{N}_\Delta.$$

This theorem in fact claims, that under conditions (23) with $a > 0$, for all small enough Δ, the main term in Ω is $\int (Q_{22}\bar{y}_2, \bar{y}_2)dt$.

Proof. (We omit the bars over the variables.) Take $u = (u_1, u_2)$, $u_1 \in H_1$, $u_2 \in H_2$ on Δ, and $y = (y_1, y_2)$, $\dot{y} = u$, $y(t_0) = 0$. Like in the classical situation, we will evaluate all the terms in Ω, taking into account that Δ is small. Recall the known estimate for the functions on an interval Δ with $|\Delta| \leq \delta$:

$$||y||_1 \leq \sqrt{\delta} \cdot ||y||_2. \tag{31}$$

Denote $\gamma_1(y) = \int_\Delta y_1^2 dt$, $\gamma_2(y) = \int_\Delta y_2^2 dt$, and let us compare $\gamma_2(y)$ with our "full"

$$\gamma(w) = |\ \xi(t_0)\ |^2 + |\ y(t_1)\ |^2 + \int_\Delta (y_1^2 + y_2^2)dt,$$

when $w = (\xi, y, u) \in \mathcal{L}_\Delta$. Since N_1 is a pointed cone, and $\dot{y} = u \in N_1$, $y(t_0) = 0$, then y_1 is monotone nondecreasing w.r.t. N_1, hence for some constant b

$$||y_1||_\infty \leq b|y_1(t_1)|\ , \tag{32}$$

and due to (30), (31)

$$||y_1||_\infty \leq bC||y||_1 \leq bC\sqrt{\delta}||y||_2,$$

therefore $\gamma_1(y) \leq \delta||y_1||_\infty^2 \leq b^2C^2\delta^2||y||_2^2 = (b^2C^2\delta^2)(\gamma_1 + \gamma_2)$, which yields that for δ small enough $\gamma_1(y) \leq 2(b^2C^2\delta^2)\gamma_2(y)$.

From here, (30) and (31) it follows that $\gamma(y)$ is of the same order as $\gamma_2(y)$; moreover, for $\delta > 0$ small enough

$$\gamma_2(y) \leq \gamma(y) \leq \frac{3}{2}\gamma_2(y). \tag{33}$$

With (30) and (31) this gives

$$||\xi||_\infty \leq \text{const} \cdot \sqrt{\delta}\sqrt{\gamma_2(y)}, \qquad |y(t_1)| \leq \text{const} \cdot \sqrt{\delta}\sqrt{\gamma_2(y)}, \tag{34}$$

whence the outside term in Ω is $\leq$ const$\cdot\delta \cdot \gamma_2(y)$.

From (34) it follows also that the integrand terms in Ω, containing ξ, have the same estimate. Due to (32) and (34) we get $||y_1||_\infty \leq$ const$\cdot\sqrt{\delta}\sqrt{\gamma_2(y)}$, whence the integrand terms, containing (y_1, y_1) and (y_1, y_2), are $\leq$ const$\cdot\delta \cdot \gamma_2(y)$ too. Recall, that due to (23) we have

$$\int_\Delta (Q_{22}y_2, y_2)\,dt \ge a\cdot\gamma_2(y). \qquad (35)$$

Thus, it remains to evaluate only the integrand term $(V_{11}y_1, u_1)$. Since $u_1\in N_1$, and N_1 is a pointed cone, there exists a vector $l\in R^{k_1}$, such that $|u_1|\le$ const (l,u_1) $\forall\, u_1\in N_1$ (actually, here one may take any $l\in \text{int } N_1^*$). Then

$$\int_\Delta |(V_{11}y_1, u_1)|\,dt \le \int_\Delta |V_{11}|\cdot|y_1|\cdot|u_1|\,dt \le \text{const}\cdot||V||_\infty\cdot||y_1||_\infty\cdot\int_\Delta (l,u_1)\,dt \le$$

$$\le \text{const}\cdot||y_1||_\infty\cdot(l, y_1(t_1)) \le \text{const}\cdot|y_1(t_1)|^2$$

(the last inequality is due to (32)), and taking into account (34), we continue $\le$ const$\cdot\delta\cdot\gamma_2(y)$.

Thus, $\Omega = \int (Q_{22}y_2, y_2)dt + \mu(\xi, y, u)$, where $|\mu|\le \text{const}\cdot\delta\cdot\gamma_2(y)$.
If δ is small enough, then $|\mu|\le \frac{1}{4}a\gamma_2(y)$, and in view of (33) and (35):

$$\Omega \ \ge\ \frac{3}{4}a\,\gamma_2(y)\ \ge\ a\,\frac{3}{4}\frac{2}{3}\,\gamma(y)\ \ge\ \frac{a}{2}\,\gamma(y), \qquad \text{q.e.d.}$$

6.2 A special case

Consider here an important special case, in which the above assumptions about Ω and estimates (30) do hold. Let system (1) be of the form:

$$\dot x = z f_0(x) + \sum_{i=1}^{k} u_i f_i(x), \qquad \dot z = 0, \qquad (36)$$

where $x\in R^n$ and $z\in R^1$ are the state variables, $\mathcal{D}_{x,z,t} = \mathcal{D}_x\times R^1\times[0,T]$, the vectors $f_i(x)$, $i = 0,1,...,k$ are C^2-smooth and linearly independent at any point $x\in\mathcal{D}_x$. Take a trajectory $\hat w = (\hat x(t),\ \hat z = 1,\ \hat u = 0)$, $t\in[0,T]$, and for any $\Delta = [t_0,t_1]\subset[0,T]$ let us consider the Problem A_Δ, associated with this trajectory, with the terminal constraints $x(t_0) = \hat x(t_0)$, $x(t_1) = \hat x(t_1)$, with $z(t_0)$ and $z(t_1)$ being free, the control u being free too, and with the cost functional $J = z(t_0)\longrightarrow \min$. Begin with a simple

Proposition 6.1. *If $\hat w$ satisfies MP in Problem $A_{[0,T]}$, then it satisfies MP in Problem A_Δ for any $\Delta\subset[0,T]$.*

Proof. The MP for Problem A_Δ means that there exists a nonzero n-vector function $\psi_x(t)$ (the costate variable, corresponding to x), such that for $H = \psi_x(zf_0(x) + \sum u_i f_i(x))$ the following relations hold on Δ along $\hat w(t)$:

$$\dot\psi_x = -H_x, \qquad H_u = 0, \qquad H = \text{const}\ \ge 0.$$

(The costate variable ψ_z, corresponding to z, is a scalar linear function with $\psi_z(t_1) = 0$ and $\dot{\psi}_z = -H$, so it is determined by ψ_x.) If such ψ_x exists for $[0,T]$, it obviously fits for any $\Delta \subset [0,T]$.

Since there are no inequality constraints in our problem, the critical cone here is, in fact, the subspace $\mathcal{L}_\Delta$, consisting of all $\bar{w} = (\bar{x}, \bar{z}, \bar{u})$, satisfying the linear equations:

$$\dot{\bar{x}} = \bar{z} f_0(\hat{x}) + f_0'(\hat{x})\bar{x} + \sum \bar{u}_i f_i(\hat{x}), \tag{37}$$

$$\dot{\bar{z}} = 0, \qquad \bar{x}(t_0) = \bar{x}(t_1) = 0.$$

Passing to the Goh variables $(\bar{z}, \bar{\xi}, \bar{y}, \bar{u})$, where

$$\bar{x} = \bar{\xi} + \sum \bar{y}_i f_i(\hat{x}),$$

$$\dot{\bar{y}}_i = \bar{u}_i, \quad \bar{y}_i(t_0) = 0, \quad i = 1, ..., k, \tag{38}$$

we get instead of (37):

$$\dot{\bar{\xi}} = \bar{z} f_0 + f_0' \bar{\xi} + \sum \bar{y}_j [f_0, f_j], \qquad \bar{\xi}(t_0) = 0 \tag{39}$$

(here $[f,g] = f'g - g'f$ are Lie brackets), and the terminal equality $\bar{x}(t_1) = 0$ becomes:

$$\bar{\xi}(t_1) + \sum \bar{y}_i(t_1) f_i(\hat{x}(t_1)) = 0, \tag{40}$$

so that now we may consider $\mathcal{L}_\Delta$ as the set of all $(\bar{z}, \bar{\xi}, \bar{y}, \bar{u})$, satisfying (39), (40) and (38). Since the vectors $f_i(\hat{x}(t))$ are linearly independent, and continuous in t, there exists a number C_0 (independent of t_1), such that (40) implies

$$|\bar{y}(t_1)| \le C_0 \left|\bar{\xi}(t_1)\right| . \tag{41}$$

Besides, (39) obviously implies that for some C_1 (independent of Δ)

$$||\bar{\xi}||_\infty \le C_1 \left(|\bar{z}| \cdot |\Delta| + \int_\Delta |\bar{y}| \, dt \right). \tag{42}$$

Let us show that $\exists\, C_2$, such that

$$|\bar{z}| \le C_2 \frac{1}{|\Delta|} ||\bar{y}||_1 \quad \text{on } \mathcal{L}_\Delta. \tag{43}$$

In order to do this, consider the space $W_\Delta = R \times L_1^k(\Delta) \times R^k$ with elements $\bar{w} = (\bar{z}, \bar{y}, \bar{h})$, and consider the subspace $\mathcal{M}_\Delta \subset W_\Delta$, defined by the equality

$$\bar{\xi}(t_1) + \sum_{i=1}^k \bar{h}_i \hat{f}_i(t_1) = 0, \tag{44}$$

where $\bar\xi$ is the solution to (39) for $\bar z, \bar y,$ and for brevity we write $\hat f_i(t) = f_i(\hat x(t))$. There is a natural injection $\mathcal{L}_\Delta \longrightarrow \mathcal{M}_\Delta,\;\; (\bar z, \bar\xi, \bar y, \bar u) \longmapsto (\bar z, \bar y, \bar h)$ (we just ignore relations (38) and put $\bar h = \bar y(t_1)$), so that $\mathcal{M}_\Delta$ can be considered as an extension of $\mathcal{L}_\Delta$. We will show that estimate (43) holds on $\mathcal{M}_\Delta$.

First we establish the following abstract fact.

Lemma 6.1. *Let Y and H be Banach spaces, $W = R \times Y \times H$, and $\mathcal{M}$ be a subspace in W. Suppose that $\exists\, \rho > 0$ such that for any $y \in Y$, $||y|| < \rho$, for any $h \in H$ the point $(1, y, h) \notin \mathcal{M}$. Then $\forall w = (z, y, h) \in \mathcal{M} \quad |z| \le \frac{1}{\rho}||y||$.*

Proof. Take any $w = (z, y, h) \in \mathcal{M}$. If $z = 0$, the desired estimate is fulfilled trivially. If $z \ne 0$, then $w' = (1, y/z, h/z) \in \mathcal{M}$, and hence $||y/z|| \ge \rho$ (otherwise $w' \notin \mathcal{M}$), which gives $\rho|z| \le ||y||$, q.e.d.

Let us return to our space W_Δ. Here $Y = Y_\Delta = L_1^k(\Delta)$, $H = R^k$, and $\mathcal{M}$ is the above $\mathcal{M}_\Delta$. In the space Y_Δ we will consider the norm $||y|| = \frac{1}{|\Delta|}||y||_1$, and we have to check the premise of Lemma 6.1.

First, we show that $\forall \bar h$ the vector $\bar w = (\bar z = 1,\; \bar y = 0,\; \bar h) \notin \mathcal{M}_\Delta$. To prove this, we have only to check that (44) does not hold. Note that for $\bar z = 1$, $\bar y = 0$ equation (39) reads: $\dot{\bar\xi} = f_0 + f_0'\bar\xi, \quad \bar\xi(t_0) = 0$, and it has the explicit solution $\bar\xi(t) = (t - t_0) f_0(\hat x(t))$ (this was noticed by A.A.Milyutin), hence its terminal value on Δ is $\bar\xi(t_1) = (t_1 - t_0) f_0(\hat x(t_1))$. Since $\hat f_i(t)$, $i = 0, 1, ..., k$ are linearly independent $\forall t \in [0, T]$, then for any $\bar h \in R^k$

$$|\Delta| \cdot \hat f_0(t_1) + \sum_{i=1}^k \bar h_i \hat f_i(t_1) \ne 0, \tag{45}$$

and so $\bar w = (1, 0, \bar h) \notin \mathcal{M}_\Delta$.

Now let us check this condition for small nonzero $\bar y$-s.

Lemma 6.2. *$\exists\, \rho > 0$ such that $\forall \Delta$, for any $\bar y \in Y_\Delta$, $||\bar y|| \le \rho$, for any $\bar h \in R^k$, the point $\bar w = (1, \bar y, \bar h) \notin \mathcal{M}_\Delta$.*

Proof. Denote $L(t) = \text{Lin}\, \{\hat f_i(t),\; i = 1, ..., k\}$. (We write *Lin* for the linear hull of the set.) Equality (44) means that $\bar\xi(t_1) \in L(t_1)$, and we have to show that this is not so. Since the vectors $\hat f_i(t)$, $i = 0, 1, ..., k$ are linearly independent $\forall t \in [0, T]$ and continuous, $\exists\, \varepsilon > 0$ such that

$$\forall t \in [0, T] \qquad B_\varepsilon(\hat f_0(t)) \cap L(t) = \emptyset. \tag{46}$$

Given this ε, $\exists\, \rho > 0$, such that $\forall \Delta$, if $||\bar y||_1 \le |\Delta| \cdot \rho$, then for $\bar\xi(t_1)$, corresponding to $\bar z = 1$ and this $\bar y$, we have

$$\left|\bar\xi(t_1) - |\Delta| \cdot \hat f_0(t_1)\right| \le \text{const} \cdot \int_\Delta |\bar y|\, d\tau \le \text{const} \cdot |\Delta| \cdot \rho \le |\Delta| \cdot \varepsilon,$$

i.e. $\bar{\xi}(t_1) \in B_{|\Delta|\varepsilon}(|\Delta| \cdot \hat{f}_0(t_1)) = |\Delta| \cdot B_\varepsilon(\hat{f}_0(t_1))$. In view of (46), for any Δ and any $\bar{y} \in Y_\Delta$, such that $||\bar{y}|| = \frac{1}{|\Delta|}||\bar{y}||_1 \le \rho$, we have $\bar{\xi}(t_1) \notin L(t_1)$, and hence $\forall \bar{h} \in R^k$ the point $\bar{w} = (1, \bar{y}, \bar{h}) \notin \mathcal{M}_\Delta$. Lemma 6.2 is proved.

Thus, the premise of Lemma 6.1 is fulfilled with a $\rho > 0$ independent of Δ. By this Lemma we get the desired estimate

$$|\bar{z}| \le C_2 \frac{1}{|\Delta|} ||\bar{y}||_1 \qquad \text{on } \mathcal{M}_\Delta \supset \mathcal{L}_\Delta \quad \text{with } C_2 = 1/\rho.$$

Finally, (42) and (43) give $||\bar{\xi}||_\infty \le \text{const } ||\bar{y}||_1$, and in view of (41) we get $|\bar{y}(t_1)| \le \text{const } ||\bar{y}||_1$ on $\mathcal{L}_\Delta$, thus estimates (30) hold, and so Theorem 6.1 applies to the considered special case.

Remark 6.1. Estimates (30) are fulfilled also for the critical subspace, corresponding to the system

$$\dot{x} = z\left(f_0(x) + \sum_{i=1}^{k} u_i f_i(x)\right), \qquad \dot{z} = 0, \tag{47}$$

at the same reference trajectory $\hat{w} = (\hat{x}(t),\ \hat{z} = 1,\ \hat{u} = 0)$, because its linearization has the same form (37) (the difference between systems (36) and (47) is only in their Ω-s).

7 Conditions for a Π-minimum

Let us now pass to conditions for a $\Pi-$minimum. As was said already, they differ from the conditions for a weak minimum by an additional pointwise condition. (Recall that we regard any pointwise condition as a condition of Legendre type.) This new condition of Legendre type involves the third variation of Lagrange function, or, more precisely, the cubic functional (17), and also the admissible control set U. Recall that we denote by U_0 the minimal face of U, containing $\hat{u}(t)$, and that due to Assumption $3'$ this face is the same for all t. For any $\lambda \in \Lambda$ and any fixed $t_* \in [0, T]$ define the functional

$$L[\lambda, t_*](\bar{y}) = \int_0^1 ((Q[\lambda](t_*)\bar{y}, \bar{y}) + (\mathcal{E}[\lambda](t_*)\bar{y}, \bar{y}, \bar{u}))d\tau, \tag{48}$$

which is to be considered for all absolutely continuous functions $\bar{y}(\tau)$ on $[0, 1]$, such that $\bar{y}(0) = \bar{y}(1) = 0$ (we will call such functions cycles). Here $Q[\lambda]$ is the matrix from the second variation (20) of Lagrange function, and $\mathcal{E}[\lambda]$ is the tensor from the third variation (21), both frozen at the point t_*.

Definition 7.1. We say that λ satisfies the $\Pi-$Legendre condition with a parameter $a \in R$, if for every t_*, and every cycle $\bar{y}(\tau)$, such that

$$\dot{\bar{y}}(\tau) = \bar{u}(\tau) \in U - \hat{u}(t_*), \tag{49}$$

the following inequality holds:

$$L[\lambda, t_*](\bar{y}) \geq a \int_0^1 (\bar{y}, \bar{y})\, d\tau. \tag{50}$$

(Relation (49) means that the control set is also frozen at the point t_*.)
Denote by $E_a(\Lambda)$ the set of all $\lambda \in G_a(\Lambda)$ satisfying this condition.

Theorem 7.1. *i) Let $\hat{w}$ be a Π-minimum point in Problem A. If $U_0 \neq U$, i.e. $\hat{u}(t) \in \partial U$, suppose also that the equality constraints (1) and (4) satisfy Lyusternik condition at $\hat{w}$. Then $E_0(\Lambda)$ is nonempty, and*

$$\Omega[E_0(\Lambda)](\bar{w}) \geq 0 \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{51}$$

ii) Let for some $a > 0$ the set $E_a(\Lambda)$ be nonempty, and

$$\Omega[E_a(\Lambda)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{52}$$

Then $\hat{w}$ is a strict Π-minimum point in Problem A.

The proof again is based on the general theory [5]. For the case $U_0 = U$, i.e. $\hat{u} \in \text{int}\ U$, it is given in [33], [34] (see also [39]); for the case $\hat{u} \in \partial U$ it is given in [37].

Consider inequality (52). Theorem 6.1 readily implies the following property for small intervals.

Lemma 7.1. *Consider $\Omega[\lambda]$ and $\eta[\lambda]$ on any interval $\Delta \subset [0, T]$, and suppose that for all small enough Δ estimates (30) hold. If for some $a > 0$ the set $E_a(\Lambda)$ is nonempty, then for all small enough Δ inequality (52) is fulfilled with $a/2$.*

Proof. If $\lambda_0 \in E_a(\Lambda)$, then by definition $\lambda_0 \in G_a(\Lambda)$, i.e. (23) is fulfilled with this $a > 0$. By Theorem 6.1 for all small enough Δ we have $\Omega[\lambda_0](\bar{w}) \geq \frac{a}{2}\gamma(\bar{w}) \quad \forall \bar{w} \in \mathcal{K}_\Delta \cap \mathcal{N}_\Delta$, and since $E_{a/2}(\Lambda) \supset E_a(\Lambda) \ni \lambda_0$, we get $\Omega[E_{a/2}(\Lambda)](\bar{w}) \geq \Omega[\lambda_0](\bar{w}) \geq \frac{a}{2}\gamma(\bar{w}) \quad \forall\, \bar{w} \in \mathcal{K}_\Delta \cap \mathcal{N}_\Delta$, q.e.d.

For large intervals, as was said already, in order to check (52) one should use a theory of Jacobi type (see e.g. [15, 16]).

Definition 7.2. We say that $\hat{w}$ is a $\Pi - \gamma$-minimum point in Problem A, if condition (52) is fulfilled for some $a > 0$.

As before, we give two equivalent formulations of this condition. Denote by $\Pi(U,\hat{w})$ the set of all Pontryagin sequences, converging to $\hat{w}$, such that

$$\text{ess sup } dist\,(u(t),U) \to 0.$$

Theorem 7.2. *i) Condition (52) with $a>0$ is equivalent to the following one: there exists $a'>0$, such that for any sequence $\{w_m\}\in\Pi(U,\hat{w})$, for all large enough m, the violation function has the lower bound:*

$$\sigma(w_m) \geq a'\gamma(w_m-\hat{w})\ . \tag{53}$$

ii) Suppose that equality constraints (1) and (4) satisfy Lyusternik condition at $\hat{w}$. Then condition (52) with $a>0$ is equivalent to the following one: there exists $a'>0$, such that if $\{w_m\}$ is a Pontryagin sequence, converging to $\hat{w}$, and satisfying equations (1) and (4) with $u_m\in U$, then inequality (53) holds for all large enough m. (Note that for such w_m the σ reduces to the first term only.)

Assertion *(i)* is proved in the general theory [5], the proof of *(ii)* is given in Appendix. Before going further, let us make some remarks about the $\Pi-$Legendre condition (50).

Remark 7.1. From (49) and $\bar{y}(0)=\bar{y}(1)=0$ it follows obviously that actually $\bar{u}(\tau)\in U_0-\hat{u}(t_*)\subset H_2$, and $\bar{u}_1=\bar{y}_1=0$, hence, for every cycle $\bar{y}$ only the second component $\bar{y}_2$ may be nonzero. Therefore, we can put in (49) $\bar{u}(\tau)\in U_0-\hat{u}(t_*)$, and so condition (50) in fact concerns the minimal face U_0 of U, containing $\hat{u}(t_*)$ (and hence containing all $\hat{u}(t)$).

Remark 7.2. Condition (50) is invariant w.r.t. the interval of integration: taking $u'(\tau)=\bar{u}(s\tau)$, $y'(\tau)=\bar{y}(s\tau)/s$, by an easy calculation we get the same inequality (50) for interval $[0,1/s]$, so we may choose to consider $[0,1]$.

Remark 7.3. Condition (50), though being seemingly of an integral form, has a pointwise character, i.e. it can be verified at each point t_* separately.

To verify condition (50), one have in fact to solve an auxiliary extremal problem: to find the maximal a, for which this inequality holds true for any cycle satisfying (49). This is rather an unusual and difficult problem. At present we can solve this problem completely for three cases of the control set U: a) the whole space, b) an arbitrary stripe of codimension 1, c) an arbitrary ellipse in the plane (see [12, 38]). For a general case some theoretical results are obtained in [13].

Here, in view of our further purposes, we consider only the most simple case (a), in which u is unconstrained. Define the differential 1-form

$$\omega[\lambda](t_*) = (\mathcal{E}[\lambda](t_*)y,y,dy) = \sum_{ijs}\mathcal{E}_{ijs}[\lambda](t_*)y^iy^j\,dy^s. \tag{54}$$

Theorem 7.3 [32, 33]. *Condition (50) holds for a given* a *iff* $Q[\lambda](t_*) \geq a$ *(Goh condition for a weak minimum), and* $\omega[\lambda](t_*)$ *is closed, i.e.*

$$d\omega[\lambda](t_*) = \sum_{ijs} \mathcal{E}_{ijs}[\lambda](t_*)(y^i dy^j + y^j dy^i) \wedge dy^s = 0, \tag{55}$$

the differential is taken w.r.t. y, *and* λ, t_* *are just parameters.*

Thus, in this case condition (50) decomposes onto conditions concerning the quadratic and cubic parts separately, the last being an additional optimality condition (55) of equality type.

Similar conditions take place in the case when $U = U_1 \oplus H_2$, $R^n = H_1 \oplus H_2$, and U_1 is a polyhedron in H_1, having $\hat{u}_1(t) \equiv \hat{u}_1$ as a vertex. (Here the above face $U_0 = H_2$.) In view of Remark 7.1, condition (50) here concerns only the cycles lying in H_2, hence (50) means that on the subspace H_2 both $Q[\lambda](t_*) \geq a$ and (55) must hold.

Note also that in this case the following property holds [31, 15].

Theorem 7.4. *Suppose that* $E_0(\Lambda)$ *is nonempty, and for some* $a > 0$

$$\Omega[E_0(\Lambda)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}.$$

Then $E_a(\Lambda)$ *is nonempty (with the same* a*), and still*

$$\Omega[E_a(\Lambda)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}.$$

The reverse implication holds trivially due to inclusion $E_a \subset E_0$. In the general case, when the face U_0 is not a subspace, Theorem 7.4 fails to hold.

Let us now consider another case, when the minimal face U_0 is "small", and establish a relation between conditions $Q[\lambda](t_*) \geq a$ and (50) in this case. As before, we have $R^n = H_1 \oplus H_2$, $H_2 = \text{Lin}\,(U_0 - \hat{u}(t))$, $N = \text{con}\,(U - \hat{u}(t)) = N_1 \oplus H_2$, $N_1 = N \cap H_1$.

Lemma 7.2. *(i) Let for* t_* *condition (50) holds. Then* $Q[\lambda](t_*) \geq a$ *on* H_2.

(ii) Suppose that $Q[\lambda](t_*) \geq a$ *on* H_2. *Then* $\forall a' < a$ $\exists \delta > 0$, *such that if* U_0 *is contained in the* δ*-neighbourhood of* $\hat{u}(t_*)$, *then (50) is fulfilled with* a'. *The number* δ *depends only on* $|a - a'|$ *and* $||\mathcal{E}[\lambda](t_*)||$.

Proof. Assertion *(i)* follows from the fact that one can take as a cycle a function $\bar{y}(\tau)$ passing with a constant velocity along an arbitrary vector $v \in H_2$ and back to zero. On such a cycle any point $\bar{y}$ is passed forward and back with the same velocity, hence the cubic term in (48) obviously vanishes, and the quadratic term reduces to $(Q[\lambda](t_*)v, v)$, which should be $\geq a(v,v)$.

Assertion *(ii)* follows from the simple estimate:

$$\int |(\mathcal{E}[\lambda](t_*)\bar{y}, \bar{y}, \bar{u})| d\tau \leq ||\mathcal{E}[\lambda](t_*)|| \cdot ||\bar{u}||_\infty \int (\bar{y}, \bar{y}) d\tau.$$

If $||\bar{u}||_\infty \leq \delta$, and $\delta||\mathcal{E}|| \leq |a - a'|$, then the cubic term, being added to the quadratic term, can worsen the constant a not lower than to a'.

Assertion *(ii)* means, that the smaller in size is the face U_0 containing $\hat{u}(t)$, the smaller contribution in $L[\lambda]$ gives the cubic term. If $Q[\lambda](t_*) \geq a > 0$ on H_2, and the face U_0 is small enough in size, then (50) is also fulfilled, perhaps with a smaller, still positive a'. Note that here the size of the entire U does not matter, it may be arbitrarily large, only the size of the face U_0 is important. This observation will be used in the next section.

8 Sufficient conditions for a strong minimum in case of a strictly convex set U

Let us now consider the case, when the control set U in Problem A is a strictly convex compactum with a smooth boundary ∂U, and $\hat{u}(t) \equiv \hat{u} \in \partial U$ on $[0, T]$. Since in this case $\partial U = ex\, U$, then, supposing also that $\forall t$ $\operatorname{rank} F(\hat{x}(t), t) = k$, by Proposition 1.1 we have that the Π-minimum implies here the strong minimum (the reverse implication always holds trivially), and so any sufficient conditions for the Π-minimum are at the same time sufficient conditions for the strong minimum. In view of this property, we are interested to obtain sufficient conditions for the Π-minimum in this case.

Remark 8.1. All considerations in this section are in fact valid for a more general case, when U is a convex compact set (not necessarily strictly convex), $\hat{u}(t) \equiv \hat{u} \in \partial U$, there is a unique outward normal $\eta \in R^k$ for U at $\hat{u}$, and $\operatorname{Argmax}(\eta, U) = \{\hat{u}\}$, i.e. $\hat{u}$ is a unique maximum point of the linear functional η on U. For example, if the set U is given by the inequality $\varphi(u) \leq 1$, where φ is a smooth strictly convex sublinear functional, these assumptions obviously are fulfilled for any $\hat{u} \in \partial U$.

Let us try to apply the abovestated theory to this case. Formally, this case does not fit in this theory, because U does not satisfy Assumption 3′ (it is not a polyhedron in a vicinity of $\hat{u}$). However, this theory applies, if we take the following approach. Since we are interested in *sufficient* conditions, then instead of U we may consider any polyhedron U' containing U, and obtain sufficient conditions for Problem A with the control set U' : obviously, these conditions will be as well sufficient for Problem A with the set U. What is important is that the choice of the polyhedron U' is in our disposal.

Let us fix any $\varepsilon > 0$, and take as U' an arbitrary polyhedron containing U, such that $\hat{u} \in \partial U'$, and the corresponding face U_0' (containing $\hat{u}$) has diameter $\le \varepsilon$. Due to the smoothness of ∂U, $\dim U_0' = k-1$, $\hat{u} \in \operatorname{reint} U_0'$, and the affine hull $\operatorname{Aff} U_0'$ is a (unique) tangent hyperplane for U at $\hat{u}$.

Remark 8.2. There are plenty of ways to construct such a polyhedron. Here we point out two of them.

1) Take any $\delta > 0$, and pick a finite δ-net $\Theta \subset \partial U$ containing $\hat{u}$ as an element. For each $u \in \Theta$ take any support hyperplane for U at u, and consider the corresponding closed halfspace $P(u)$ containing U. Then $U' = \bigcap_{u\in\Theta} P(u)$ satisfies the above requirements with some $\varepsilon(\delta)$ tending to zero as $\delta \to 0$.

2) Let P be an arbitrary polyhedron containing U, such that $\hat{u} \in \partial P$, let η be the outward normal to U (and then to P) at $\hat{u}$, and L be the tangent hyperplane at $\hat{u}$. Choose an arbitrary polyhedral norm in R^k, and for any $\varepsilon > 0$ denote by $B_\varepsilon(\hat{u})$ the ε-ball around $\hat{u}$.

Given an $\varepsilon > 0$, denote by K_δ, $\delta > 0$ the cone generated by $B_\varepsilon(\hat{u}) \cap L$ with vertex at $\hat{u} + \delta\eta$. Since $L \cap U = \{\hat{u}\}$, then $K_\delta \supset U$ for some $\delta > 0$. (The existence of such $\delta > 0$ is obvious for any 2-dimensional cutset of U containing η, and then, by virtue of the arguments of compactness, for the entire U.)

Take, for the certainty, δ be the maximal of those δ-s, and denote it by $\delta(\varepsilon)$. Then the set $U' = P \cap K_{\delta(\varepsilon)}$ will do. (For small $\varepsilon > 0$, its face containing $\hat{u}$, is $U_0' = B_\varepsilon(\hat{u}) \cap L$, which diameter is $\le \varepsilon$.)

Chosen any method of constructing the polyhedron U', we will consider that $\forall \varepsilon > 0$ we have a polyhedron U_ε with the above properties. Note that generally $\bigcap_{\varepsilon>0} U_\varepsilon \neq U$ (e.g. in the above second method), and in fact we do not need that U_ε would be close to U; we only need that the size of the face U_ε^0 in U_ε (corresponding to the normal η) would tend to zero. We do not provide here pictures, since the reader can easily draw them himself.

Problem A with the control set U_ε (instead of U) will be denoted by Problem A_ε. We want to obtain sufficient conditions for Π-minimum in Problem A_ε. But now, since U_ε is a polyhedron, we can use the above theory, and formulate sufficient conditions for $\Pi-\gamma$-minimum in Problem A_ε. In order to do this, we need to inspect all objects of this theory for Problem A_ε.

As was noted in Remark 2.2 (and it is very important for us), due to the singularity of $\hat{w}$, the set $\Lambda(\hat{w})$ (and hence the families $\Omega[\lambda]$ and $\rho[\lambda]$) do not depend on the choice of the control set U_ε, since the last contains the initial set U. Besides, for all U_ε the tangent (at $\hat{u}$) cone N is one and the same: $N = \{\bar{u} \mid (\eta, u) \leq 0\}$, where η is the outward normal to U at $\hat{u}$ (which is assumed to be unique), the maximal subspace in N is $H_2 = \{\bar{u} \mid (\eta, u) = 0\}$, its complement $H_1 = R\eta$, and $N_1 = -R_+\eta$.

All this implies, that $\forall a \in R$ the set $G_a(\Lambda)$ is one and the same for all Problems A_ε, $\varepsilon > 0$. It consists of all $\lambda \in \Lambda$, such that quadratic form $\Omega[\lambda]$ satisfies conditions:

$$i)\ \ V[\lambda](t) = 0 \tag{56}$$

(here both conditions *(i)* and *(ii)* of (23) are taken into account, as well as the fact that $\dim H_1 = 1$), and

$$ii)\ \ Q[\lambda](t) \geq a \quad \text{on} \quad H_2. \tag{57}$$

As to the set $E_a(\Lambda)$, involving in its definition the face of U_ε containing $\hat{u}$, here the situation is as follows. Denote this set for Problem A_ε by $E_a(\Lambda, U_\varepsilon)$. Let us fix an $a > 0$, and take a $\lambda \in G_a(\Lambda)$. Due to Lemma 7.2, $\exists\, \varepsilon_0 > 0$, such that $\lambda \in E_{a/2}(\Lambda, U_\varepsilon)$ for all $\varepsilon \leq \varepsilon_0$. The number ε_0 depends only on a and $||\mathcal{E}[\lambda](t)||_\infty$. Since the tensor $\mathcal{E}[\lambda](t)$ is linear in λ, and the set Λ is bounded, then $\max\{||\mathcal{E}[\lambda]||_\infty,\ \lambda \in \Lambda\}$ is a finite number, hence we can consider that ε_0 depends only on a, and does not depend on λ.

Thus, $\forall \varepsilon \leq \varepsilon_0(a)$ we have: if $\lambda \in G_a(\Lambda)$, then $\lambda \in E_{a/2}(\Lambda, U_\varepsilon)$, i.e.

$$G_a(\Lambda) \subset E_{a/2}(\Lambda, U_\varepsilon). \tag{58}$$

The critical cone $\mathcal{K}$ for all Problems A_ε with the free control (i.e. without taking into account the constraint $u \in U_\varepsilon$) is one and the same by its definition, because all these problems differ only in the sets U_ε.

All this obviously implies, that if the inequality

$$\Omega[G_a(\Lambda)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N} \tag{59}$$

holds, then the inequality

$$\Omega[E_{a/2}(\Lambda, U_\varepsilon)](\bar{w}) \geq a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N} \tag{60}$$

holds too (and the more so the last one holds, if in its right hand side $a\gamma$ is replaced by $a\gamma/2$). Recall now, that by Theorem 5.1 inequality (59) is a sufficient condition for the weak γ-minimum in Problem A_ε, and by Theorem 7.1 inequality (60) is a sufficient condition for $\Pi - \gamma$-minimum in Problem A_ε. Thus, we have got the following chain of implications:

weak $\gamma-$minimum in Problem A_ε $\Longrightarrow$ $\Pi-\gamma-$minimum in Problem A_ε $\Longrightarrow$
strict $\Pi-$minimum in Problem A_ε $\Longrightarrow$
(due to inclusion $U_\varepsilon \supset U$) strict $\Pi-$minimum in Problem A $\Longrightarrow$
(due to Proposition 1.1) strict strong minimum in Problem A.

Finally, in order to complete the picture, we add one more step. Consider the halfspace $U_* = \{u \mid (\eta, u) \le (\eta, \hat{u})\}$ containing U, bounded by the tangent hyperplane to U at $\hat{u}$. Problem A with U_* we denote by A_*. Obviously, for this problem the local cone is still the same as for Problems A_ε (i.e. N). Therefore, $G_a(\Lambda, U_*) = G_a(\Lambda)$ for all a, whence the weak $\gamma-$minimum in Problem A_* simply coincides with the weak $\gamma-$minimum in any Problem A_ε, $\varepsilon > 0$. Note by the way, that on the level of weak minimum not only $\gamma-$sufficient conditions in all these problems coincide: in fact, the problems themselves do not differ one from another, because in a neighborhood of $\hat{u}$ (depending on ε) the sets U_* and U_ε simply coincide.

These arguments and the above chain of implications yield the following sufficient conditions for the strong minimum in Problem A.

Theorem 8.1. *Suppose that $\hat{w}$ satisfies $\gamma-$sufficient conditions for the weak minimum in Problem A_*, i.e. the set $G_0(\Lambda)$ is nonempty, and for some $a > 0$*

$$\Omega[G_0(\Lambda)](\bar{w}) \ge a\gamma(\bar{w}) \qquad \forall \bar{w} \in \mathcal{K} \cap \mathcal{N}. \tag{61}$$

Then $\hat{w}$ is a strict strong minimum point in the initial Problem A.

Moreover, for some $\varepsilon, C > 0$, on the set $||x - \hat{x}||_\infty < \varepsilon$, for all $w = (x, u)$, $u \in U$, satisfying equation (1), the following inequality holds:

$$\sum_{i \in I} \varphi_i^+(p) + |K(p)| \ge C\gamma(w - \hat{w}).$$

Proof. By Theorem 5.1 inequality (61) is equivalent to (26) = (59), and, as was shown, this implies (60) for some $\varepsilon > 0$, which means that $\hat{w}$ is a $\Pi - \gamma$-minimum point in Problem A_ε, and which is a sufficient condition for the strict Π-minimum in Problem A_ε, and hence for the strict Π-minimum in Problem A. In our case the last is equivalent to the strict strong minimum, so the first assertion is proved.

Now let us pass to the second assertion. Due to Theorem 7.2, from (60) it follows that $\exists\, C > 0$, such that for any sequence $w_m = (x_m, u_m)$, $u_m \in U$, satisfying (1) and such that

$$||x_m - \hat{x}||_\infty \to 0, \qquad ||u_m - \hat{u}||_1 \to 0, \tag{62}$$

we have

$$\sigma(w_m) = \sum_{i \in I} \varphi_i^+(p_m) + |K(p_m)| \ge C \cdot \gamma(w_m - \hat{w}). \tag{63}$$

Due to Lemma 1.1, the first condition in (62) implies (for our U, and $\hat{u} \in \partial U$) the second one, so the last one can be omitted. Then the resulting assertion obviously implies (in fact is equivalent to) the required second assertion of the theorem.

From here and Theorem 6.1 we obtain the following theorem for small time intervals. Consider a special case of Problem A as in Sec. 6.2, for system (47) with the constraint $u \in U$, and with the cost $J = z(t_0) \to \min$. (This corresponds to the problem of time-optimality for system (47) with $z = 1$.)

Theorem 8.2. *Suppose that* $\exists\, \delta_0 > 0$, *such that* $\forall \Delta$, $|\Delta| \le \delta_0$, *the set* $G_+(\Lambda(A_*(\Delta)))$ *is nonempty, i.e. there exists a costate function* $\psi(t)$, *satisfying on* Δ *the costate equation and conditions (56) and (57) with an* $a > 0$. *Then there exists* $\delta_1 > 0$, *such that* $\forall \Delta$, $|\Delta| \le \delta_1$ *trajectory* $\hat{w}_\Delta$ *is a strict weak minimum point in Problem* $A_*(\Delta)$, *and a strict global minimum point in Problem* A_Δ.

Note that here we do not require that the set $G_+(\Lambda(A_*[0,T]))$ should be nonempty, or even that the Lagrange set $\Lambda(A_*[0,T])$ itself should be nonempty.

In the proof we will need the following three lemmas.

Lemma 8.1. *Suppose that* $\exists\, \delta_0 > 0$, *such that* $\forall\, \Delta_0 \subset [0,T]$, $|\Delta_0| \le \delta_0$, $\exists\, \delta > 0$, *such that if* $\Delta \subset \Delta_0$, $|\Delta| \le \delta$, *then some Property* $P(\Delta)$ *holds. Then* $\exists\, \delta_1 > 0$ *such that if* $\Delta \subset [0,T]$ *and* $|\Delta| \le \delta_1$, *then* $P(\Delta)$ *holds.*

The proof follows from simple arguments of compactness.

Lemma 8.2. *Suppose for some* $\Delta_0 \subset [0,T]$ *the* $\hat{w}$ *is a strict strong minimum point in Problem* A_{Δ_0}. *Then* $\exists\, \delta, \varepsilon > 0$, *such that if* $\Delta \subset \Delta_0$, $|\Delta| \le \delta$, *a point* $w = (z,x,u)$ *is admissible in Problem* A_Δ, *and* $|z-1| < \varepsilon$, *then* $z > 1$.

Proof. The strict strong minimum in Problem A_{Δ_0} means that $\exists\, \varepsilon_0 > 0$, such that if $w = (z,x,u)$ is admissible in Problem A_{Δ_0}, and $|z-1| < \varepsilon_0$, $||x-\hat{x}||_\infty < \varepsilon_0$ on Δ_0, then $z > 1$.

Consider the operator $L_\infty(\Delta_0) \times L_1(\Delta_0) \to C(\Delta_0)$, mapping $(z,u) \mapsto x$ according to system (47) with the fixed value $x = \hat{x}$ at the left end of Δ_0. Since this operator is continuous at $(\hat{z} = 1,\ \hat{u} = 0)$, and maps it to $\hat{x}$, $\exists\, \varepsilon_1 > 0$, such that if $||z(t)-1||_\infty < \varepsilon_1$, $||u||_1 < \varepsilon_1$, then $||x-\hat{x}||_\infty < \varepsilon_0$ on Δ_0. Take this $\varepsilon_1 \le \varepsilon_0$. We claim that the following Property F holds:

$\exists\, \varepsilon_2 > 0$, *such that if* $(z(t),x,u)$ *satisfies (47) on* Δ_0 *with* $u \in U$, $x = \hat{x}$ *at* $\partial\,\Delta_0$, $1-\varepsilon_2 \le z(t) \le 1$, *and* $||u||_1 < \varepsilon_2$, *then there exists a triple* (z',x',u'), *admissible in Problem* A_{Δ_0}, *such that* $1-\varepsilon_0 \le z' \le 1$, *and* $||x'-\hat{x}||_\infty < \varepsilon_0$.

Indeed, let us reparametrize the given x, u, putting

$$q = \frac{1}{|\Delta_0|} \int_{\Delta_0} z(t)\, dt, \qquad \text{and} \quad d\tau = \frac{1}{q} z(t)\, dt,$$

and define $z' = q$, $x'(\tau) = x(t(\tau))$, $u'(\tau) = u(t(\tau))$. Obviously, the new triple (z', x', u') still satisfies (47), and since z' is constant, this triple is admissible in Problem A_{Δ_0}; besides, since z' is the mean value of $z(t)$, we have $1-\varepsilon_2 \le z' \le 1$. If ε_2 is small enough, then $|z'-1| < \varepsilon_1 \le \varepsilon_0$, and $||u'||_1 < \varepsilon_1$, and hence, due to the continuity of the above operator, $||x' - \hat{x}||_\infty < \varepsilon_0$. Property F is proved.

Now, since U is bounded, $\exists\, \delta_2 > 0$, such that if $|\Delta| < \delta_2$, then $||u||_1 < \varepsilon_2$. We claim that for these δ_2 and ε_2 the assertion of Lemma holds.

Take any $\Delta \subset \Delta_0$, $|\Delta| \le \delta_2$, and let a point $w = (z, x, u)$ be admissible in Problem A_Δ with $|z-1| < \varepsilon_2$. We have to show that $z > 1$. Suppose that $z \le 1$. By the definition of δ_2 we have $||u||_1 < \varepsilon_2$. Expand $x(t)$ on the entire Δ_0, putting $x = \hat{x}$ outside of Δ (this is possible, because $x = \hat{x}$ at both ends of Δ), and expand $u(t)$, putting $u = 0$ outside of Δ, and define $z(t) = z$ on Δ and $z(t) = 1$ outside of Δ. The triple $(z(t), x(t), u(t))$ satisfies the conditions of Property F, and hence there exists a triple (z', x', u'), admissible in Problem A_{Δ_0}, such that $1 - \varepsilon_0 \le z' \le 1$, $||x' - \hat{x}||_\infty < \varepsilon_0$ and $z' \le 1$. But this contradicts the strict strong minimum at $\hat{w}$ in Problem A_{Δ_0}, and hence $z > 1$. Lemma 8.2 is proved.

Lemma 8.3. *Suppose that for some $\Delta_0 \subset [0,T]$ the $\hat{w}$ is a strict strong minimum point in Problem A_{Δ_0}. Then $\exists\, \delta > 0$, such that if $\Delta \subset \Delta_0$, $|\Delta| \le \delta$, then $\hat{w}$ is a strict global minimum point in Problem A_{Δ_0}.*

Proof. By Lemma 8.2 we have the following Property B: $\exists\, \delta,\, \varepsilon > 0$, *such that if* $\Delta \subset \Delta_0$, $|\Delta| \le \delta$, *a point* $w = (z, x, u)$ *is admissible in Problem* A_Δ, *and* $|z-1| < \varepsilon$, *then* $z > 1$. Fix this $\delta > 0$, and let us show that we can get rid of the requirement $|z-1| \le \varepsilon$, i.e. that the following Property C holds: *if* $|\Delta| \le \delta$, *and the point* $w = (z, x, u)$ *is admissible in Problem* A_Δ, *then* $z > 1$.

Indeed, if it is not true, then $\exists\, \Delta$, $|\Delta| \le \delta$, and $w = (z, x, u)$, admissible in Problem A_Δ, such that $z \le 1$. Taking a linear change of time: $d\tau = z \cdot dt$, we get $x'(\tau)$ on an interval Δ', with $|\Delta'| = z \cdot |\Delta| \le \delta$, satisfying equation (47) with $z = 1$, which contradicts Property B. Thus, Property C is proved, and it is just the strict global minimum in Problem A_Δ for any $\Delta \subset [0,T]$ with $|\Delta| \le \delta$, q.e.d.

Proof of Theorem 8.2. Take any $\Delta_0 \subset [0,T]$, $|\Delta_0| \le \delta_0$, and a $\lambda \in G_+(\Lambda(A_*(\Delta_0)))$. By definition λ satisfies conditions (56) and (57) on Δ_0 with some $a = a(\Delta_0) > 0$. Due to Theorem 6.1, $\exists\,\delta > 0$ such that if $\Delta \subset \Delta_0$, and $|\Delta| \le \delta$, then

$$\Omega[\lambda](\bar{w}) \ge \frac{a}{2}\gamma(\bar{w}) \qquad \forall\,\bar{w} \in \mathcal{L}_\Delta \cap \mathcal{N}_\Delta\,,$$

i.e. $\hat{w}$ is a weak γ-minimum point in Problem $A_*(\Delta)$. By Theorem 8.1 $\hat{w}$ is a strict strong minimum point in Problem A_Δ. From here by Lemma 8.1 we get that $\exists\,\delta_1 > 0$, such that if $\Delta \subset [0,T]$ and $|\Delta| \le \delta_1$, then $\hat{w}$ is a strict strong minimum point in Problem A_Δ. For this δ_1 we have by Lemma 8.3, that $\forall\,\Delta \subset [0,T]$, $|\Delta| \le \delta_1$, $\exists\,\delta > 0$ such that if $\Delta' \subset \Delta$ and $|\Delta'| \le \delta$, then $\hat{w}$ is a strict global minimum point in Problem $A_{\Delta'}$. Using again Lemma 8.1, we obtain that $\exists\,\delta_2 > 0$, such that if $\Delta \subset [0,T]$ and $|\Delta| \le \delta_2$, then $\hat{w}$ is a strict global minimum point in Problem A_Δ. Theorem 8.2 is proved.

9 Acknowledgments

The author is grateful to Prof. A.A.Milyutin for numerous discussions and valuable suggestions. The author thanks the American Mathematical Society for the invitation to the 1997 Summer Research Institute at the University of Colorado, Boulder, where these results were presented, and the anonymous referee for useful comments.

APPENDIX

Here we prove assertions *(ii)* of Theorems 5.2 and 7.2. Begin with some known abstract facts. In a Banach space X consider a system of linear equalities and inequalities:

$$\Phi_i(x) \le 0, \;\; i \in I, \qquad G\,x = 0, \tag{64}$$

where $\forall\,i$ Φ_i is a sublinear functional, such that the set $\Phi_i(x) < 0$ is nonempty, and G is a linear operator from X onto a Banach space Y. In [2] A.Ja.Dubovitskii and A.A.Milyutin proved that the system with strict inequalities

$$\Phi_i(x) < 0, \;\; i \in I, \qquad G\,x = 0,$$

has no solution if and only if there exist $\alpha_i \ge 0$, $x_i^* \in \partial\Phi_i$, $i \in I$, and $y^* \in Y$, such that $\sum \alpha_i + ||y^*|| = 1$, and

$$\sum \alpha_i x_i^* + y^* G = 0. \tag{65}$$

(Here $\partial\Phi_i$ is the subdifferential in the sense of convex analysis.) We denote by Λ the set of all tuples $\lambda = (\{(\alpha_i,\, x_i^*),\; i \in I\},\; y^*)$, satisfying these conditions. Suppose that Λ is nonempty.

Definition. We say that an index i_0 (or the constraint $\Phi_{i_0}(x) \le 0$) is singular in system (64), if $\forall\,\lambda \in \Lambda$ the multiplier $\alpha_{i_0} = 0$. Otherwise we say that i_0 is nonsingular.

The following relation between this "dual" and a "primal" characterizations of singular constraints holds.

Lemma A1. *a) Suppose that $\exists\,x$ satisfying (64), such that $\Phi_{i_0}(x) < 0$. Then the index i_0 is singular.*

b) Let I_0 be the set of all singular indices, and suppose that $\forall\,j \notin I_0$ the functional Φ_j is linear. Then $\exists\,x$ satisfying (64), such that $\forall\,i \in I_0 \quad \Phi_i(x) < 0$.

Proof. *a)* Suppose that i_0 is nonsingular, i.e. $\exists\,\lambda \in \Lambda$ with $\alpha_{i_0} > 0$. From $x_i^* \in \partial\Phi_i$ we have $(x_i^*, x) \le \Phi_i(x) \le 0 \;\; \forall\, i$, and < 0 for i_0, whence this x violates equation (65), a contradiction.

b) Denote by $I_1 = I \setminus I_0$ the set of all nonsingular indices, and suppose there is no such x, i.e. the cones $K_i = \{x \mid \Phi_i(x) < 0\}$, $i \in I_0$, and $C = \{x \mid \Phi_j(x) \le 0,\; \forall\, j \in I_1,\; Gx = 0\}$ do not intersect. By the Dubovitskii–Milyutin theorem [2] there exists a nontrivial collection of $p_i \in K_i^*$ and $q \in C^*$, such that $\sum p_i + q = 0$.

For each $i \in I_0$, since the set $\Phi_i(x) < 0$ is nonempty, we have $p_i = -\alpha_i x_i^*$, where $\alpha_i \ge 0$ and $x_i^* \in \partial\Phi_i$. The collection of α_i is nontrivial, because otherwise all $p_i = 0$, and then also $q = 0$, a contradiction. Now, since for each $j \in I_1$ we assume $\Phi_j(x) = (l_j, x)$, where $l_j \in X^*$, then by the Farkas lemma $q = -\sum \alpha_j l_j - y^* G$, where $\alpha_j \ge 0$ and $y^* \in Y$. Summarizing these expressions for p_i and q, we get

$$\sum_{I_0} \alpha_i\, x_i^* + \sum_{I_1} \alpha_j\, l_j + y^* G = 0.$$

The collection $(\alpha_i,\, \alpha_j,\, y^*)$, properly normalized, belongs to Λ, and since $\sum_{I_0} \alpha_i > 0$, there is an $i_0 \in I_0$, such that $\alpha_{i_0} > 0$. This contradicts the singularity of i_0. Lemma A1 is proved.

Let us return to our Problem A in the space W. First we note that for some $\varepsilon > 0$, $\forall\, t$, in the ε-neighborhood of $\hat u(t)$ the set U coincides with the polyhedron M, given by inequalities (27), and hence, for some (may be smaller) $\varepsilon > 0$, $\forall\, t$, in the ball $B_\varepsilon(0)$ the set $U - \hat u(t)$ simply coincides with its tangent cone $N = \{\bar u \mid (a_s, \bar u) \le 0,\; s \in S_0\}$. Since U has nonempty interior, so does N. Put, for convenience in notation, $\hat w = 0$; then in a neighborhood of zero $U = N$.
Recall that we denote $\mathcal{N} = \{\bar w(t) = (\bar x(t), \bar u(t)) \in W \mid \bar u(t) \in N \text{ a.e. on } [0,T]\}$.

Consider the following system of linear equalities and strict inequalities in W:

$$\varphi_i'(0)\bar{p} < 0, \quad i \in I, \quad K'(0)\bar{p} = 0, \quad \bar{x} = A\bar{x} + B\bar{u}, \quad \bar{u} \in \operatorname{int}\mathcal{N}.$$

The last inequality can be written as

$$\Phi(\bar{u}) = \min_{s \in S_0} \ \operatorname{ess\,sup}\,(a_s, \bar{u}(t)) < 0,$$

where obviously Φ is a sublinear functional. The stationarity of $\hat{w}$ means that this system has no solution, which by the Dubovitskii–Milyutin theorem is equivalent to the existence of Lagrange multipliers $\lambda = (\alpha_i, \ \beta \in R^{dimK}, \ \psi \in L_\infty, \ \mu_s \in L_1)$, such that $\alpha_i \ge 0 \ \ \forall\, i \in I, \ \ \mu_s(t) \ge 0$ a.e. on $[0,T] \ \ \forall\, s \in S_0$, and the Euler–Lagrange equation holds: for all $\bar{w} \in W$

$$\sum \alpha_i \varphi_i'\,\bar{p} + K'\,\bar{p} + \int \psi\,(\bar{x} - A\bar{x} - B\bar{u})\,dt + \sum \int \mu_s\,\bar{u}\,dt = 0$$

For Problem A this condition is equivalent to MP.

The total singularity of $\hat{w}$ (i.e. its singularity w.r.t. all U) is equivalent to that all $\mu_s(t) = 0$ for all such λ-s, and by Lemma A1 the last property is equivalent to the existence of $\bar{w}_* \in W$, such that

$$\varphi_i'(0)\,\bar{p}_* < 0, \qquad K'(0)\,\bar{p}_* = 0, \qquad \dot{\bar{x}}_* = A\bar{x}_* + B\bar{u}_*, \tag{66}$$

$$\bar{u}_* \in \operatorname{int}\mathcal{N}, \quad \text{i.e. } \exists\,\varepsilon_* > 0 \ \text{ such that } \ B_{\varepsilon_*}(\bar{u}_*(t)) \subset N \ \text{ a.e. on } [0,T].$$

(Recall that system (66) defines the critical cone $\mathcal{K}$.) We can take $||\bar{w}_*|| \le 1$. These $\bar{w}_*$ and the corresponding $\varepsilon_* > 0$ will be used in what follows.

Let us establish a simple property of the cone N in R^k. Define the function $\sigma_N(u) = \sum_{s \in S_0}(a_s, u)^+$, and pick any $\bar{u} \in \operatorname{int} N$ and $\varepsilon > 0$, such that $B_\varepsilon(\bar{u}) \subset N$.

Lemma A2. *For any ρ there exists $C = C(N, \varepsilon, \rho)$, such that $\forall\, \delta > 0$, $\forall\, u \in R^k$ with $\sigma_N(u) \le \delta$, the point $u + \delta C\bar{u}$ lies in the $\delta\rho$-interior of N, i.e.*

$$u + \delta C \bar{u} + B_{\delta\rho}(0) \subset N.$$

Proof. In view of homogeneity of this property, we can consider only the case when $\delta = 1$, i.e. to prove that for some C, if $\sigma_N(u) \le 1$, then $(u + C\bar{u}) + B_\rho(0) \subset N$.

Obviously, $\exists\, C_0 = C_0(N, \varepsilon)$ such that for all these u we have $u + C_0\bar{u} \in N$. Then, $\exists\, C_1 = C_1(N, \varepsilon, \rho)$ such that for all $u' \in N$ the point $u' + C_1\bar{u}$ lies in the ρ-interior of N. Taking $C = C_0 + C_1$, we get the required property.

Now we recall the known estimate of the distance to the level set of an operator g, mapping a Banach space X to another Banach space Y (see e.g. [11]). Suppose, g

is strictly differentiable at a point $x_0 \in X$ (e.g. continuously Frechet differentiable at x_0), $g(x_0) = 0$, and denote $\mathcal{M} = \{x \mid g(x) = 0\}$. If $g'(x_0)$ is onto (i.e. g satisfies Lyusternik condition at x_0), then there exist a number L and a neighborhood $\mathcal{B}(x_0)$, such that

$$\forall\, x \in \mathcal{B}(x_0) \qquad \text{dist}\,(x, \mathcal{M}) \le L\,||g(x)||. \tag{67}$$

Finally, we prove an estimate for the increment of the function

$$\sigma_0(w) = \sum \varphi_i^+(p) + |\,K(p)\,| + \int |\,\dot{x} - f_0(x) - F(x)u\,|\,dt, \tag{68}$$

which is crucial for our purposes.

Lemma A3. *Let* $w_n \to 0$, $\bar{w} \in \mathcal{K}$, *and* $\varepsilon_n \to 0$.
Then $\sigma_0(w_n + \varepsilon_n \bar{w}) \le \sigma_0(w_n) + o\,(\varepsilon_n)$.

Proof. Consider the increment of the first term in (68) (we omit i):

$$\varphi^+(p_n + \varepsilon_n \bar{p}) = [\,\varphi(p_n) + \varphi'(p_n) \cdot \varepsilon_n \bar{p} + o\,(\varepsilon_n)\,]^+ \le$$

(because $(\cdot)^+$ is a sublinear functional)

$$\le \varphi^+(p_n) + \varepsilon_n\,[\,\varphi'(p_n)\,\bar{p}\,]^+ + o(\varepsilon_n). \tag{69}$$

Since $\varphi'(p_n) \to \varphi'(0)$, and $\varphi'(0)\,\bar{p} \le 0$ (because $\bar{w} \in \mathcal{K}$), the second term in (69) is $o(\varepsilon_n)$, whence for the increment of the first term in (68) we obtain the desired estimate. Similar arguments work for other two terms in (68). Lemma A3 is proved.

Proof of Theorem 5.2 (ii). In view of assertion *(i),* we have to show, that if $\exists\, a > 0$ and a neighborhood $\mathcal{B}(\hat{w})$, such that $\forall\, w \in \mathcal{B}(\hat{w})$ satisfying (1), (4) with $u \in N$, the inequality $\sigma(w) \ge a\gamma(w)$ holds, then $\exists\, a' > 0$ and a neighborhood $\mathcal{B}'(\hat{w})$, in which the inequality $\sigma(w) \ge a'\gamma(w)$ holds. (The reverse implication is trivial.) We prove the contrary-reverse implication, i.e. if $\exists\, w_n \to 0$, such that $\sigma(w_n) = o(\gamma(w_n))$, then $\exists\, w_n' \to 0$ satisfying (1), (4) with $u_n' \in N$, such that $\sigma(w_n') = o(\gamma(w_n'))$.

Thus, suppose that $\exists\, w_n \to 0$ with $\sigma(w_n) \le \alpha_n\,\gamma(w_n)$, where $\alpha_n \to 0$. In particular, we have $\sigma_N(u_n) \le \alpha_n\,\gamma(w_n)$. Denote by L the constant from the Lyusternik estimate (67) for the operator g, corresponding to equality constraints (1) and (4) (see Def. 5.1).

By Lemma A2, taking $\bar{u} = \bar{u}_*(t)$, $\varepsilon = \varepsilon_*$ (uniform for a.a. t), and $\rho = 2L$, we get $\delta_n = \alpha_n\,\gamma(w_n)$, and $w_n' = w_n + C\,\delta_n\,\bar{w}_*$, such that

$$B_{\rho\delta_n}(u_n') \subset N. \tag{70}$$

In particular, $u_n' \in N$, and hence $\sigma(w_n') = \sigma_0(w_n')$. By Lemma A3 we have

$$\sigma_0(w_n') \le \sigma_0(w_n) + o(\alpha_n \gamma(w_n)) \le 2\,\alpha_n\,\gamma(w_n)$$

for large enough n. The point w_n' violates equalities (1) and (4) not more than by $\sigma_0(w_n')$. Using estimate (67), we get a point w_n'' satisfying these equalities, such that $||w_n'' - w_n'|| \le L \cdot 2\,\alpha_n\,\gamma(w_n) = \rho\,\alpha_n\,\gamma(w_n)$. In particular, $|\,u_n''(t) - u_n'(t)\,| \le \rho\,\alpha_n\,\gamma(w_n)$, and due to (70) we have $u_n''(t) \in N$. Since the function σ_0 is Lipschitzian (at least in a neighborhood of zero), then

$$\sigma(w_n'') = \sigma_0(w_n'') \le \sigma_0(w_n') + \text{const}\,||w_n'' - w_n'|| \le$$

$$\le 2\,\alpha_n\,\gamma(w_n) + \text{const} \cdot 2L\,\alpha_n\,\gamma(w_n) = o(\gamma(w_n)).$$

Finally, since $||w_n'' - w_n|| \le o(\gamma(w_n))$, we obviously have $\gamma(w_n'') \sim \gamma(w_n)$, so we may write $\sigma(w_n'') \le o(\gamma(w_n''))$. Thus, we get the desired sequence $w_n'' \to 0$, satisfying (1), (4) and the last estimate, with $u_n''(t) \in N$, q.e.d.

Proof of Theorem 7.2 (ii) is similar to the previous one. We have to show, that if there exists a sequence $\{w_n\} \in \Pi(U, \hat{w})$, such that $\sigma(w_n) = o(\gamma(w_n))$, then there exists a sequence $\{w_n'\} \in \Pi(U, \hat{w})$ satisfying equalities (1), (4) with $u_n' \in N$, such that $\sigma(w_n') = o(\gamma(w_n'))$. The only two points in the previous proof, where we used that $||w_n|| \to 0$, were those, when we used Lyusternik estimate (67) and the Lipschitz continuity of σ_0 in a neighborhood of zero. The validity of these properties in a neighborhood of zero is not sufficient now, because Pontryagin sequences do not, generally, converge to zero in the norm of W. However, these properties are valid not only in a neighborhood of zero, but on a more broad set $||w||_1 = |x(0)| + ||\dot{x}||_1 + ||u||_1 < \varepsilon$, for some $\varepsilon > 0$. For estimate (67) this was proved in [33, Part II], [34], while the Lipschitz continuity of σ_0 on this set (w.r.t. $||w||$, or even w.r.t. $||w||_1$) is rather obvious. Thus, in both these points the previous proof remains valid, and so Theorem 7.2 *(ii)* is proved.

References

[1] G.A.Bliss, *Lectures on the calculus of variations,* Chicago, 1946.

[2] A.Ja.Dubovitskii, A.A.Milyutin, *Extremal problems in the presence of constraints,* USSR Comput. Math. and Math. Physics, 1965, v. 5, No. 3.

[3] A.Ja.Dubovitskii, A.A.Milyutin, *Theory of the Maximum Principle,* "Metody teorii ekstremal'nyh zadach v ekonomike" ("Application of the theory of extremal problems to economics") (V.L.Levin ed.), Moscow, "Nauka", CEMI, 1981, p. 6–47 (in Russian).

[4] A.A.Milyutin, *Quadratic conditions of an extremum in smooth problems with a finite-dimensional image,* – ibid., p. 138–177 (in Russian).

[5] E.S.Levitin, A.A.Milyutin, N.P.Osmolovskii. *Conditions of higher–order for a local minimum in extremal problems with constraints,* Russian Math. Surveys, 1978, v. 33, No. 6, p. 97–168.

[6] N.P.Osmolovskii. *Necessary and sufficient conditions of a high order for a Pontryagin and a bounded-strong minima in an optimal control problem,* Soviet Phys. Doklady, 1988, v. 33, No. 12, p. 883–885.

[7] N.P.Osmolovskii, *Quadratic conditions for nonsingular extremals in optimal control (a theory),* Russian Journal of Mathematical Physics, Wiley, 1994, v. 2, No. 4, 487–516.

[8] V.A.Dubovitskii, *Necessary and sufficient conditions of a Pontryagin minimum in an optimal control problem with singular regimes,* Russian Math. Surveys, 1982, v. 37, No. 6.

[9] E.S.Levitin, *Perturbation theory in the mathematical programming and its applications*, Moscow, Nauka, 1992.

[10] A.A.Milyutin, *Maximum Principle in the general optimal control problem*, Moscow, "Phasis", 1998, to appear (in Russian).

[11] A.V.Dmitruk, A.A.Milyutin, N.P.Osmolovskii, *Lyusternik's theorem and the theory of extremum,* Russian Math. Surveys, 1980, v. 35, No. 6, p. 11–51.

[12] A.V.Dmitruk, A.A.Milyutin, *New Legendre type conditions for optimal control problems, linear in the control,* Proc. of the European Control Conference, Brussels, 1997.

[13] A.A.Milyutin, *On a duality formula for multidimensional vector fields,* Russian Journal of Mathematical Physics, Wiley, 1995, v. 3, No. 1, p. 81–112.

[14] A.V.Dmitruk, *The Euler-Jacobi equation in the calculus of variations,* Math. Notes of the Acad. Sci. USSR, 1976, v. 20, p. 1032–1038.

[15] A.V.Dmitruk, *Jacobi type conditions for the problem of Bolza with inequalities,* ibid., 1984, v. 35, p. 427–435.

[16] A.V.Dmitruk, *A condition of Jacobi type for a quadratic form to be nonnegative on a finite-faced cone,* Mathematics of the USSR, Izvestija, 1982, v. 18, No. 3, p. 525–535.

[17] H.J.Kelley, R.E.Kopp, H.G.Moyer, *Singular extremals,* Topics in Optimization (ed. G.Leitman), Acad. Press, New York–London, 1967, p. 63–101.

[18] B.S.Goh, *Necessary conditions for singular extremals involving multiple control variables,* SIAM J. on Control, 1966, v. 4, No. 4, p. 716–731.

[19] R.Gabasov, F.M.Kirillova, *Singular optimal controls,* Nauka, Moscow, 1973.

[20] D.J.Bell, D.H.Jacobson, *Singular Optimal Control Problems,* Academic Press, NY, 1975.

[21] H.W.Knobloch, *Higher order necessary conditions in optimal control theory,* Lecture Notes in Control and Inf. Sciences, v. 34, 1981.

[22] A.J.Krener, *The high order maximal principle and its application to singular extremals,* SIAM J. on Control, 1977, v. 15, No. 2, p. 256–293.

[23] A.A.Agrachiov, R.V.Gamkrelidze, *Second order optimality principle for a time-optimal problem,* Math. USSR, Sbornik, 1976, v. 100, No. 4.

[24] A.V.Sarychev, *High-order necessary conditions of optimality for nonlinear control systems,* Systems & Control Letters, 1991, v. 16, p. 369–378.

[25] M.I.Zelikin, *An optimality condition for singular trajectories in the problem of minimization of a curvilinear integral,* Soviet Math. Doklady, 1982, v. 267, No. 3.

[26] F.Lamnabhi-Lagarrigue, G.Stefani, *Singular optimal control problems: on the necessary conditions of optimality,* SIAM J. on Control & Optim., 1990, v. 28, No. 4, p. 823–840.

[27] J.L.Speyer, D.H.Jacobson. *Necessary and sufficient conditions of optimality for singular control problems. A transformation approach,* J. Math. Analysis and Appl., 1971, v. 33, No. 1, p. 163–186.

[28] I.B.Vapnyarskii, *Existence theorem in the problem of Bolza, and necessary conditions of sliding and singular regimes,* USSR Comput. Math. and Math. Physics, 1967, v. 7, No. 2.

[29] V.I.Gurman, *Singular optimal control problems,* Moscow, Nauka, 1977.

[30] V.A.Dykhta, *Conditions of a local minimum for singular regimes in systems with linear control,* Automatics and Remote Control, 1981, No. 12.

[31] A.V.Dmitruk, *Quadratic order conditions of a weak minimum for singular regimes in optimal control problems,* Soviet Math. Doklady, 1977, v. 18, No. 2.

[32] A.V.Dmitruk, *Quadratic conditions for a Pontryagin minimum in an optimal control problem, linear in the control, with a constraint on the control,* Soviet Math. Doklady, 1983, v. 28, No. 2, p. 364–368.

[33] A.V.Dmitruk, *Quadratic conditions for a Pontryagin minimum in an optimal control problem, linear in the control,* Mathematics of the USSR, Izvestija, 1987, v. 28, No. 2, p. 275–303, and 1988, v. 31, No. 1, p. 121–141.

[34] A.V.Dmitruk, *Fine theorems on weakening equality constraints in optimal control problems, linear in control,* Siberian Math. Journal, 1990, v. 31, No. 2, p. 37–51.

[35] A.V.Dmitruk, *Second order necessary and sufficient conditions of a Pontryagin minimum for singular regimes,* Lecture Notes in Control and Inf. Sciences, 1992, v. 180, p. 334–343.

[36] A.V.Dmitruk, *Second order optimality conditions for singular extremals,* in Computational Optimal Control (R.Bulirsch and D.Kraft eds.), Intern. Ser. on Numer. Mathematics, Birkhaueser, Basel, v. 115 (1994), p. 71–81.

[37] A.V.Dmitruk, *Quadratic order conditions of a Pontryagin minimum for singular extremals in optimal control problems,* Dr.Sci. Dissertation, St. Petersburg University, 1994.

[38] A.V.Dmitruk, *Second Order Necessary and Sufficient Conditions of a Pontryagin Minimum for Singular Boundary Extremals,* Proc. of the Third Intern. Congress on Industrial and Appl. Math., Hamburg, 1995: ZAMM, Issue 3, p. 411–412.

[39] A.V.Dmitruk, *Quadratic order conditions of a local minimum for abnormal extremals,* Nonlinear Analysis, Theory, Methods & Appl., 1997, v. 30, No. 4, p. 2439–2448.

[40] D.J.W.Ruxton, D.Bell, *Junction times in singular optimal control,* Appl. Math. Comput., 1995, v. 70, No. 2-3, p. 143–154.

[41] A.Visintin, *Strong convergence results related to strict convexity,* Commun. Partial Diff. Equations, 1984, v. 9, No. 5, p. 439–466.

[42] A.F.Filippov, *On some topics in the theory of optimal regulation,* Vestnik Moskovskogo Universiteta, Ser. mat. mech. phys., 1959, No. 2 (in Russian).

[43] Ch.Castaing, *Sur les multi-applications measurables,* Rev. Francaise Inf. Rech. Oper., 1967, v. 1, p. 91–126.

Central Economics and Mathematics Institute of the Russian Academy of Sciences, Russia 117418, Moscow, Nakhimovskii prospekt, 47

e-mail: dmitruk@member.ams.org

Published in *Proceedings of Symposia on Pure Mathematics, v. 64 "Differential Geometry and Control"* (Eds. G.Ferreyra et al.), Amer. Math. Soc., 1999, p. 163-198.