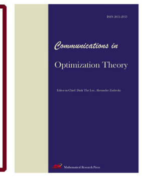
Commun. Optim. Theory 2026 (2026) 22 https://doi.org/10.23952/cot.2026.22

Communications in Optimization Theory
Available online at http://cot.mathres.org



SECOND ORDER CONDITIONS OF THE WEAK OPTIMALITY FOR PROBLEMS LINEAR IN THE CONTROL

ANDREI DMITRUK

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

Dedicated to Professor Vladimir Tikhomirov on the occasion of his 90th birthday

Abstract. An optimal control problem is considered in which the control system is linear in the control variables and nonlinear in the state variables, with no control constraints. For this problem, we propose no-gap necessary and sufficient conditions for the weak minimality of a quadratic order which is the integral of the squared variation of an additional state variable. Here, the second variation is automatically degenerate, and we transform it to the classical possibly non-degenerate form that essentially simplifies verification of its sign definiteness.

Keywords. Critical cone; Goh transformation; Lagrange function; Quadratic order; No-gap conditions; Second variation.

2020 Mathematics Subject Classification. 49K27, 49K15.

1. Statement of the Problem and the Main Result

Consider the following optimal control problem of the Mayer type:

$$J := \varphi_0(x(0), x(T)) \to \min, \tag{1.1}$$

$$\varphi_i(x(0), x(T)) \leqslant 0, \qquad i = 1, \dots, d(\varphi), \tag{1.2}$$

$$\eta_j(x(0), x(T)) = 0, \qquad j = 1, \dots, d(\eta), \tag{1.3}$$

$$\dot{x}(t) = f(t, x(t)) + F(t, x(t))\, u(t), \tag{1.4}$$

where the time $t \in [0,T]$ varies in a fixed interval, x is the $n-$dimensional state variable, u is the control of dimensions r. As usual, we assume that $x(t)$ is an absolute continuous function: $x \in AC^n[0,T]$, while $u(t)$ is an essentially bounded function: $u \in L^r_\infty[0,T]$.

Note that we do not allow the convenient constraint $u(t) \in U$, so actually, our problem is a general problem of the classical calculus of variations, linear in the derivatives. Obviously, any

E-mail address: optcon@mail.ru

Received January 5, 2025; Accepted March 26, 2025.

extremal in this problem is totally singular. Problems of this type often appear in various technical, economical and other applied studies, and they have been a subject of many mathematical studies (see e.g. [7]–[22]). However, the majority of those works consider only second order necessary optimality conditions, while second order sufficient conditions are proposed only in [21] with the proof in [22]. Since the last paper is written in Russian and now hardly available, here we give another, more simple proof of both necessary and close to them sufficient conditions of a special quadratic order.

The functions of finite-dimensional argument $(x(0),x(T))$ are defined on an open set $\mathscr{P} \subset \mathbb{R}^{2n}$, and the functions depending on (t,x) are defined on an open set $\mathscr{Q} \subset \mathbb{R}^{1+n}$. We assume that all these functions are twice smooth, i.e., belong to the class $C^2(\mathscr{P})$ or $C^2(\mathscr{Q})$, resp.

For brevity, the problem (1.1)–(1.4) will be called Problem A. Our main goal is to obtain second-order conditions for the weak minimality in this problem.

A pair of functions $w(t) = (x(t),u(t))$ on $[0,T]$ related by eq. (1.4), will be called *a process* of the problem. A process is called *admissible* if its endpoints $(x(0),x(T))$ belong to the set $\mathscr{P}$, there exists a compact set $D \subset \mathscr{Q}$ such that $(t,x(t)) \in D$ for almost all t, and all constraints of the problem are satisfied. As usual, we say that an admissible process $\widehat{w}(t) = (\widehat{x}(t),\widehat{u}(t)))$ delivers *a weak minimum* if there exists an $\varepsilon > 0$ such that $J(w) \geqslant J(\widehat{w})$ for any admissible process $w(t) = (x(t),u(t))$ satisfying the conditions $||x-\widehat{x}||_C < \varepsilon,\ ||u-\widehat{u}||_\infty < \varepsilon.$

Introduce the space $W = AC^n[0,T] \times L^r_\infty[0,T]$ with elements $w = (x,u)$, equipped with its natural norm

$$||w|| = ||x||_{AC} + ||u||_\infty = |x(0)| + \int_0^T |\dot{x}(t)|\,dt + \operatorname*{vraimax}_{t\in[0,T]} |u(t)|.$$

An easy fact is that the minimality in problem A w.r.t. this norm is equivalent to the weak minimality.

The first order necessary conditions for the weak minimality are well known. They can be obtained from a general Lagrange multipliers principle (see e.g. [1]–[6] and elsewhere).

Theorem 1.1. *If a process $\widehat{w}(t) = (\widehat{x}(t),\widehat{u}(t)))$ delivers a local minimum in problem A, then there exist a number $\alpha_0 \geqslant 0$, row vectors $\alpha \geqslant 0$, β of dimensions $d(\varphi)$, $d(\eta)$, respectively, and an essentially bounded function $\psi(t)$ of dimension n (adjoint or costate variable), not all equal to zero, satisfying the complementary slackness conditions*

$$\alpha_i\,\varphi_i(\widehat{x}_0,\widehat{x}_T) = 0 \ \text{ for all } \ i = 1,\ldots,d(\varphi),$$

and such that the Lagrange function

$$\mathscr{L}(x,u) = \big(\alpha_0\varphi_0 + \sum \alpha_i\varphi_i + \sum \beta_j\,\eta_j\big)(x_0,x_T) + \int_0^T \psi\big(\dot{x} - f(t,x) - F(t,x)\,u\big)\,dt,$$

is stationary at the point $(\widehat{w})$:

$$\mathscr{L}'(\widehat{x},\widehat{u}) = 0, \tag{1.5}$$

which means that

$$\mathscr{L}'(\widehat{x},\widehat{u})(\bar{x},\bar{u}) = 0 \ \text{ for all } \ (\bar{x},\bar{u}) \in W. \tag{1.6}$$

Introducing the *endpoint Lagrange function*

$$l(x_0,x_T) = \big(\alpha_0\varphi_0 + \sum \alpha_i\varphi_i + \sum \beta_j\,\eta_j\big)(x_0,x_T),$$

and the *Pontryagin function* $H(\psi,t,x,u) = \psi\big(f(t,x)+F(t,x)u\big)$, we can represent the Lagrange function in a shorter form

$$\mathscr{L}(x,u) = l(x_0,x_T) + \int_0^T \big(\psi\dot{x} - H(t,x,u)\big)\,dt,$$

whence condition (1.6) means that, for all $\bar{x}\in AC,\ \forall\,\bar{u}\in L_\infty$,

$$l_{x_0}\bar{x}(0) + l_{x_T}\bar{x}(T) + \int_0^T \big(\psi\dot{\bar{x}} - H_x\bar{x} - H_u\bar{u}\big)\,dt = 0. \tag{1.7}$$

An easy decryption of this condition shows that it is equivalent to the following set of conditions: $\psi(t)$ is a Lipschitz continuous function,

$$\begin{aligned}
&\text{(a)}\quad \dot{\psi}(t) = -H_x(\psi(t),t,\widehat{x}(t),\widehat{u}(t)),\\
&\text{(b)}\quad \psi(0) = l_{x_0}(\widehat{x}_0,\widehat{x}_T), \qquad \psi(T) = -l_{x_T}(\widehat{x}_0,\widehat{x}_T),\\
&\text{(c)}\quad H_u(\psi(t),t,\widehat{x}(t),\widehat{u}(t)) = 0 \quad \text{for a.a. } t\in[0,T].
\end{aligned}$$

Relation (a) is known as adjoint (costate) equation, (b) as transversality conditions, and (c) as condition of stationarity in the control.

Note that the function $\psi(t)$ is uniquely determined by the triple (α_0,α,β), whence the nontriviality condition can be reduced to the finite-dimensional multipliers only, i.e. to the form $\alpha_0+|\alpha|+|\beta|>0$. (Otherwise, if these multipliers vanish, then $l\equiv 0$, so $\psi(0)=0$, and since ψ satisfies a linear homogeneous ODE, we have $\psi(t)=0$ as well, which contradicts the "full" normalization.)

Any process of problem A satisfying the above conditions, is called *stationary*.

Now, let us pass to second order conditions.

Let $\widehat{w}(t) = (\widehat{x}(t)),\widehat{v}(t))$ be a stationary process. Fix it and consider the set Λ of all triples $\lambda=(\alpha_0,\alpha,\beta)$ normalized by $\alpha_0+|\alpha|+|\beta|=1$ and satisfying the above nonnegativity and slackness conditions, that, together with the corresponding function $\psi(t)$, realize the stationarity condition (1.7) for the process $\widehat{w}$, or, what is the same, the above conditions (a), (b), (c). Obviously, Λ is a nonempty finite-dimensional compact set.

For each $\lambda\in\Lambda$ consider the corresponding Lagrange function $\mathscr{L}[\lambda](x,u)$ (now we indicate its dependence on λ), and define its second variation

$$\Omega[\lambda](\bar{x},\bar{u}) = (l''_{x_0,x_0}\bar{x}(0),\bar{x}(0)) + 2(l''_{x_T,x_0}\bar{x}(0),\bar{x}(T)) + (l''_{x_T,x_T}\bar{x}(T),\bar{x}(T)) -$$

$$-\int_0^T \big((H''_{xx}\bar{x},\bar{x}) + 2(H''_{ux}\bar{x},\bar{u})\big)\,dt. \tag{1.8}$$

Following [20], along with these quadratic functionals, we introduce the functional

$$\Omega[\Lambda](\bar{x},\bar{u}) = \max_{\lambda\in\Lambda}\Omega[\lambda](\bar{x},\bar{u}).$$

Clearly, it is homogeneous of degree 2, but not quadratic if the set Λ is more than a singleton.

Without loss of generality we assume that $\varphi_0(\widehat{x}(0),\widehat{x}(T))=0$. Introduce the set I of active indices $i\in\{0,1,\dots,d(\varphi)\}$ such that $\varphi_i(\widehat{x}(0),\widehat{x}(T))=0$, and *the cone of critical variations* K_0

consisting of all $\bar w = (\bar x, \bar u) \in W$ that satisfy the relations:

$$\begin{aligned} \varphi'_{ix_0}\bar x(0) + \varphi'_{ix_T}\bar x(T) &\leqslant 0 \qquad i \in I, \\ \eta'_{jx_0}\bar x(0) + \eta'_{jx_T}\bar x(T) &= 0 \qquad \forall\, j, \\ \dot{\bar x} &= f'_x \bar x + (F'_x \bar x)\widehat u + F\,\bar u. \end{aligned} \tag{1.9}$$

Here, all the data functions are evaluated at the reference process $\widehat w$.

Finally, we introduce a quadratic functional, which will be called *the order of minimum*:

$$\gamma(\bar w) \,=\, |\bar x(0)|^2 + \int_0^T |\bar y(t)|^2\,dt \,+\, |\bar y(T)|^2, \tag{1.10}$$

where $\bar y(t)$ is an additional, artificial state variable subjected to equation

$$\dot{\bar y}(t) \,=\, \bar u(t), \qquad \bar y(0) = 0. \tag{1.11}$$

Now we are ready to formulate the conditions of order γ ("second order" conditions) for a weak minimum. Let $\widehat w$ be a stationary process of problem A.

Theorem 1.2. *a) If $\widehat w$ delivers a weak minimum, then*

$$\Omega[\Lambda](\bar w) \,\geqslant\, 0 \qquad \textit{for all } \; \bar w \in K_0\,. \tag{1.12}$$

b) If there exists a constant $c > 0$ such that

$$\Omega[\Lambda](\bar w) \,\geqslant\, c\,\gamma(\bar w) \qquad \textit{for all } \; \bar w \in K_0\,, \tag{1.13}$$

then $\widehat w$ delivers a strict weak minimum; moreover, there exists a constant $b > 0$ and a neighborhood $\mathscr{O}(\widehat w)$ in the space W, where the violation function

$$\sigma(w) := \sum_{i\in I} \varphi_i^+(w) + |\eta(w)| + \int_0^T \Big|\dot x - f(t,x) - F(t,x)\,u\Big|\,dt$$

satisfies the estimate $\sigma(w) \geqslant b\,\gamma(w - \widehat w)$.

Note that, for all admissible w, one has $\sigma(w) = J(w)$, so the last estimate yields the quadratical growth of the cost: $J(w) \geqslant b\,\gamma(w - \widehat w)$ in $\mathscr{O}(\widehat w)$.

Example 1. $n = 3, qr = 2$, $\dot x_1 = u_1, \dot x_2 = u_2 + x_2 u_1$, $\dot x_3 = 2x_1u_1 + 2x_2u_2 + x_1^2 + x_2^2$, $x(0) = 0$, $t \in [0,1]$, $J = x_3(1) \to \min$. The reference process is $\hat x = 0$, $\hat u = 0$. Here, $H = \psi_1 u_1 + \psi_2(u_2 + x_2u_1) + \psi_3(2x_1u_1 + 2x_2u_2 + x_1^2 + x_2^2)$, $l = \alpha_0 x_3 + \beta_1 x_1(0) + \beta_2 x_2(0) + \beta_3 x_3(0)$, so $\dot\psi_1 = \dot\psi_2 = \dot\psi_3 = 0$, $H_{u_1} = \psi_1 = 0$, $H_{u_2} = \psi_2 = 0$, $\psi_3(1) = \alpha_0 := 1$. The set Λ consists of a single element $\lambda = (\alpha_0 = 1,\ \beta_1 = \beta_2 = \beta_3 = 0)$, the second variation is

$$\Omega = \int_0^1 (2\bar x_1 \bar u_1 + 2\bar x_2 \bar u_2 + \bar x_1^2 + \bar x_2^2)\,dt,$$

and the critical cone K is given by the relations $\dot{\bar x}_1 = u_1$, $\dot{\bar x}_2 = u_2$, $\bar x_1(0) = \bar x_2(0) = 0$, whence $\bar x_1 = \bar y_1, \bar x_2 = \bar y_2$, and

$$\Omega = \bar x_1^2(1) + \bar x_2^2(1) + \int_0^1 (\bar x_1^2 + \bar x_2^2)\,dt \,=\, \gamma(\bar x),$$

so the sufficiency condition (1.13) is satisfied, and hence, the reference process provides a weak minimum.

Theorem 1.2 gives a pair of no-gap conditions of the above order γ. It was announced in [21] and proved in [22] on the base of a general theory of higher order conditions proposed by Levitin, Milyutin and Osmolovskii in [20]. That theory is very powerful but rather complicated. Moreover, the paper [22] is now hardly available even to the readers in Russia, not to speak of other countries. Here, in Sec 6 below, we give another proof, based on a more simple theory, namely, on the so-called *two-norm approach*, proposed in [23] and used also in [24] and [25]. Note that we use a slightly modified version of this approach.

Meanwhile, let us show how conditions (1.12) and (1.13) can be in a sense refined by narrowing the set of λ under the max sign. This will be done in the next sections.

2. The Case When Λ is a Singleton

Consider first the case when the set Λ consists of a single element $\lambda = (\alpha_0, \alpha, \beta, \psi)$, and represent the quadratic functional (1.8) in a more general form, without the above exact relation to Problem A:

$$\Omega(\bar{x},\bar{u}) = (S_{00}\bar{x}(0),\bar{x}(0)) + 2(S_{T0}\bar{x}(0),\bar{x}(T)) + (S_{TT}\bar{x}(T),\bar{x}(T)) + \\ + \int_0^T \Big((Q(t)\bar{x},\bar{x}) + 2(P(t)\bar{x},\bar{u}) \Big)\, dt, \tag{2.1}$$

where all coefficients are matrices of corresponding dimensions, among which S_{00}, S_{TT} and $Q(t)$ are symmetric.

Comparing with (1.8) and taking into account that the control $\widehat{u}(t)$ is measurable and bounded, we assume that the matrix Q has measurable and bounded entries, while P has Lipschitz continuous entries.

This quadratic functional is considered on a cone $K_0 \subset W = AC \times L_\infty$ of the form

$$\begin{aligned} a_{i0}\bar{x}(0) + a_{iT}\bar{x}(T) &\leqslant 0, \qquad i \in I, \\ b_{j0}\bar{x}(0) + b_{jT}\bar{x}(T) &= 0, \qquad j \in J, \\ \dot{x} &= A\bar{x} + B\bar{u}. \end{aligned} \tag{2.2}$$

Here I, J are some finite sets of indices, the matrix A has measurable and bounded entries, B is Lip-continuous, and a_{i0}, a_{iT}, b_{j0}, b_{jT} are some vectors of dimension $|I|$, $|J|$, respectively.

In view of Theorem 1.2, we have to analyze conditions (1.12) and (1.13) for the quadratic functional (2.1) on the cone K_0. We proceed as follows.

2.1. **Goh transformation.** First, we make the following transformation. Introduce the above state variable $\bar{y}$ of dimension r subjected to equation (1.11), and a new state variable $\bar{\xi} = \bar{x} - B\bar{y}$ of dimension n. Obviously, the last one satisfies the relations

$$\dot{\bar{\xi}}(t) = A\bar{\xi} + B_1\bar{y}, \qquad \bar{\xi}(0) = \bar{x}(0), \tag{2.3}$$

where the $r\times r-$matrix $B_1 = AB - \dot{B}$ is measurable and bounded. Note that this equation does not contain $\bar{u}$; the last one comes (in the simplest form!) only into equation (1.11) for $\bar{y}$.

Now, since $\bar{x} = \bar{\xi} + B\bar{y}$, we can replace the initial state variable $\bar{x}$ by the pair of new state variables $\bar{\xi}$, $\bar{y}$. This passage, proposed by B.S. Goh in [8], is called *Goh transformation*.

The endpoint conditions, determining the cone K_0, now take the form:

$$\begin{aligned} a_{i0}\,\bar\xi(0)+a_{iT}\big(\bar\xi(T)+B(T)\bar y(T)\big) &\leqslant 0, \qquad i\in I,\\ b_{j0}\,\bar\xi(0)+b_{jT}\big(\bar\xi(T)+B(T)\bar y(T)\big) &= 0, \qquad j\in J. \end{aligned} \tag{2.4}$$

Let us rewrite Ω in the new variables, aiming to exclude its explicit dependence on the control $\bar u$. Define by $q(\bar x(0),\bar x(T))$ the endpoint terms in (2.1); it is a finite-dimensional quadratic form in the space $\mathbb{R}^{2n}$. Setting $\bar x=\bar\xi+B\bar y$, we get

$$\Omega = q\Big(\bar\xi(0),\,\bar\xi(T)+B(T)\bar y(T)\Big) + \int_0^T \Big((Q\bar\xi,\bar\xi)+2(Q\bar\xi,B\bar y)+(QB\bar y,B\bar y)+ \\ + 2(P\bar\xi,\bar u) + 2(PB\,\bar y,\bar u)\Big)\,dt. \tag{2.5}$$

We see that here $\bar u$ comes only in the last two terms. Consider each of them. The term $(P\bar\xi,\bar u)$ can be integrated by parts in view of equation (2.3):

$$\int_0^T (P\bar\xi,\bar u)\,dt = (P\bar\xi,\bar y)\Big|_0^T - \int_0^T ((P\bar\xi)^{\bullet},\bar y)\,dt = \\ = \big(P(T)\bar\xi(T),\bar y(T)\big) - \int_0^T \big((\dot P\bar\xi,\bar y)+(P(A\bar\xi+B_1\bar y),\bar y)\big)\,dt.$$

All the terms in the resulting expression do not contain $\bar u$ and so, can be added to preceding terms in expression (2.5).

Consider the term $(PB\,\bar y,\bar u)$. The Lip-continuous $r\times r$-matrix $P(t)B(t)$ can be represented as the sum of two matrices: $PB(t)=S(t)+V(t)$, where $S=PB+B^*P^*$ is symmetric and $V=PB-B^*P^*$ skew-symmetric ones. Since $(S\bar y,\bar y)^{\bullet}=(\dot S\bar y,\bar y)+2(S\bar y,\bar u)$, we have

$$\int_0^T (S\bar y,\bar u)\,dt = \frac{1}{2}\big(S(T)\bar y(T),\bar y(T)\big) - \frac{1}{2}\int_0^T (\dot S\bar y,\bar y)\,dt.$$

The obtained expression does not contain $\bar u$ and, again, can be added to preceding terms in (2.5). However, the term $(V\bar y,\bar u)$ cannot be removed a priori.

Gathering similar terms, we come to the following quadratic functional:

$$\Omega = q_1\big(\bar\xi(0),\bar\xi(T),\bar y(T)\big) + \int_0^T \Big((Q\bar\xi,\bar\xi)+2(P_1\bar\xi,\bar y)+(R_1\bar y,\bar y)+(V\bar y,\bar u)\Big)\,dt, \tag{2.6}$$

where

$$q_1\big(\bar\xi(0),\bar\xi(T),\bar y(T)\big) = \\ = q\big(\bar\xi(0),\,\bar\xi(T)+B(T)\,\bar y(T)\big)+2(P(T)\,\bar\xi(T),\bar y(T))+\tfrac{1}{2}(S(T)\bar y(T),\bar y(T)),$$

$$S=PB+B^*P^*, \qquad V=PB-B^*P^*, \qquad B_1=AB-\dot B,$$

$$P_1=B^*Q-\dot P-PA, \qquad R_1=B^*QB-PB_1-B_1^*P^*-\tfrac{1}{2}\,\dot S.$$

These formulas will not be used in below. Note only, that due to the imposed assumptions, the skew-symmetric matrix $V(t)$ is Lip-continuous, and the symmetric matrix $R_1(t)$ is measurable and bounded.

The resulting expression contains only one term with $\bar{u}$, this is $(V\bar{y},\bar{u})$. In general, this term cannot be removed. Nevertheless, it can in some special case, and this is the main step in the promised refinement.

Theorem 2.1. *If the functional* $\Omega(\bar{w}) \geqslant 0$ *on the cone* K_0 *given by* (2.4) *and* (2.3)*, then the following two pointwise conditions hold true on* $[0,T]$*:*

$$a)\ V(t) = 0, \qquad b)\ R_1(t) \geqslant 0. \tag{2.7}$$

Here, condition a) holds for all $t \in [0,T]$, while b) holds for almost all $t \in [0,T]$.

These conditions are well-known as *Goh conditions* (of equality and inequality types, respectively); the proof was first given in [8] for the case without endpoint constraints and in [22] for the general case.

2.2. **Passage to a new control.** By the above theorem, we come to a quadratic functional of the form

$$\Omega = q_1\big(\bar{\xi}(0), \bar{\xi}(T), \bar{y}(T)\big) + \int_0^T \Big((Q\bar{\xi},\bar{\xi}) + 2(P_1\bar{\xi},\bar{y}) + (R_1\bar{y},\bar{y})\Big)\,dt, \tag{2.8}$$

that does not contain explicitly the control $\bar{u}$, which comes only into equation $\dot{\bar{y}} = \bar{u}$.

Then, one may remove this equation and, in view of (2.3), consider the variable $\bar{y}$ as a new control(!). Moreover, this variable, initially belonging to the space $\mathrm{Lip}_0^r[0,T]$ (consisting of all Lip-continuous functions $y(t)$ on $[0,T]$ with the value $y(0) = 0$), can be now taken from the wider space $L_2^r[0,T]$ (see the next two lemmas). The state variable $\bar{\xi}$ is uniquely determined from equation (2.3) by the initial condition $\bar{\xi}(0)$ and the control $\bar{y}$.

One may notice an obstacle concerning the endpoint conditions (2.4) that involves the value $\bar{y}(T)$, which is not admissible in the space $L_2^r[0,T]$. However, this obstacle is easily overcome by replacing $\bar{y}(T)$ with an arbitrary vector $\bar{h} \in \mathbb{R}^r$. Thus, the quadratic functional (2.8) now depends on the triple $(\bar{\xi}(0), \bar{y}, \bar{h})$ and have the following form (where we drop the bars):

$$\Omega = q_1\big(\xi(0), \xi(T), h)\big) + \int_0^T \Big((Q\xi,\xi) + 2(P_1\xi,y) + (R_1 y,y)\Big)\,dt, \tag{2.9}$$

subject to equation

$$\dot{\xi}(t) = A\,\xi + B_1 y, \tag{2.10}$$

and the endpoints relations

$$\begin{aligned} &a_{i0}\,\xi(0) + a_{iT}(\xi(T) + B(T)h) \leqslant 0, \qquad i \in I,\\ &b_{j0}\,\xi(0) + b_{jT}(\xi(T) + B(T)h) = 0, \qquad j \in J. \end{aligned} \tag{2.11}$$

This extension of the space

$$\mathfrak{U} := \{\, (\xi(0), y(\cdot), y(T)) \mid (\xi(0), y(\cdot)) \in \mathbb{R}^n \times \mathrm{Lip}_0^r[0,T] \,\}$$

to the space $H := \mathbb{R}^n \times L_2[0,T] \times \mathbb{R}^r$ of elements $(\xi(0), y(\cdot), h)$ is justified by the following two simple lemmas on density (see e.g. [31, 32]), and by the continuity of Ω in the extended space H.

Denote by K the subset of H determined by relations (2.11).

Lemma 2.2. *The subspace $\mathfrak{U}$ in $L_2[0,T] \times \mathbb{R}^r$ consisting of the pairs $(y(\cdot), y(T))$, where $y \in \mathrm{Lip}\,[0,T]$ with $y(0) = 0$, is dense in $L_2[0,T] \times \mathbb{R}^r$.*

Lemma 2.3. *Let H be a locally convex topological vector space, K be a finite-faced cone in H, and $\mathfrak{U}$ be a linear variety (algebraic subspace) dense in H. Then $K \cap \mathfrak{U}$ is dense in K.*

In our case, $K \cap \mathfrak{U} = K_0$, so Lemma 2.3 says that K_0 is dense in K. Therefore, if the functional Ω of the form (2.8) is nonnegative on K_0 in the space $\mathfrak{U}$, then it is nonnegative on K in the space H.

Thus, we have to analyze the sign-definiteness of quadratic functional (2.9) on the cone K given by relations (2.10)–(2.11) in the space H. This functional has almost classical type; the only difference from the classics is the presence of the parameter h that comes into the above relations. However, concerning this parameter we can give one more optimality condition, along with conditions $a)$ and $b)$ of Theorem 2.1:

Lemma 2.4. *If $\Omega(\xi, y, h) \geqslant 0$ on the cone K, then, for any pair (ξ, y) satisfying (2.10),*

$$c) \quad \inf_h \Omega(\xi, y, h) \;\geqslant\; 0,$$

where the infimum is taken over all h such that the corresponding triple (ξ, y, h) satisfies (2.3) and (2.4).

Despite the seeming "awkwardness" of this condition, it proves to be fairly effective in determining the non-optimal processes. Thus, the nonnegativity of Ω on K implies three necessary conditions: $a), b), c)$. In a typical case, when Ω is strictly non-degenerate in the control y, i.e. when $R_1(t) \geqslant \text{const} > 0$, its sign-definiteness can be analyzed by methods of the classical calculus of variations (see e.g. [31]–[33]).

Next, consider the general case when the set Λ is not a single point.

3. The Set Λ is More Than a Singleton

In this case, we must get back to the functional (2.1), involving the original control u, and then, after the above reductions, to the functional (2.6) on the space $\mathfrak{U}$, where all matrices linearly depend on λ from the set Λ. Consider its convex hull $M = \mathrm{co}\,\Lambda$. Obviously, it is compact in $\mathbb{R}^m$, where $m = |I| + |J|$. Since $\Omega[\lambda](\bar{w})$ is linear in λ, conditions (1.12)–(1.13) remain equivalent if one replace Λ by $\mathrm{co}\,\Lambda$.

Now, let G be the set of all $\lambda \in \mathbb{R}^m$ such that quadratic functional $\Omega[\lambda](\bar{w})$ satisfies the two conditions (2.7). Clearly, it is convex and closed, may be empty. Define the set $G(M) = G \cap M$ and the functional

$$\Omega[G(M)](\bar{w}) \;=\; \sup_{\lambda \in G(M)} \Omega[\lambda](\bar{w}).$$

(As usual, we assume that $\sup \varnothing = -\infty$.)

The following refined version of Theorem 1.2 holds true.

Theorem 3.1. *If $\Omega[M](\bar{w}) \geqslant 0$ on the cone K_0 in $\mathfrak{U}$, then also $\Omega[G(M)](\bar{w}) \geqslant 0$ on the cone K in H. In particular, it follows that $G(M)$ is nonempty.*

Since $G(M) \subset M$ and $K_0 \subset K$, the last inequality is obviously stronger than the first one.

The proof will be done in two steps, according to conditions a) and b) in (2.7). The first step is aimed to prove condition a). Since M is a set in $\mathbb{R}^m$, we can assume that we are given a vector-valued quadratic functional $\Omega(w) = \big(\Omega_1(w), \ldots, \Omega_m(w)\big)$, where each Ω_i is a functional of the form (2.6), and so, for any $\lambda \in \mathbb{R}^m$ we have a quadratic functional

$$\Omega[\lambda](w) = (\lambda, \Omega(w)) = \sum_i \lambda_i \Omega_i(w).$$

Along with this, we consider the vector function $(Vy,u) = \big((V_1 y,u), \ldots, (V_m y,u)\big)$ and similar vector functions for all other coefficients of Ω. (Notation V and Ω without indices or square brackets are used for ordered sets of matrices $(V_1, \ldots, V_m)$ and functionals $(\Omega_1, \ldots, \Omega_m)$. We hope this will not cause confusion.)

Now, fix any $t_* \in [0,T]$. For any Lip-continuous function $y(\tau)$ on the interval $[0,1]$ with the endpoint values $y(0) = y(1) = 0$, we define the vector

$$\sigma(y) := \int_0^1 \big(V(t_*)\, y(\tau), \dot{y}(\tau)\big)\, d\tau \in \mathbb{R}^m.$$

Denote by $L(t_*)$ the set of all vectors $\sigma(y)$ generated by all such functions $y(\tau)$. Note that $\sigma(y) = \oint \big(V(t_*)\, y, dy\big)$, where the integral is taken along the curve traced by the function $y(\tau)$. Therefore, if $\sigma \in L(t_*)$, then $k\sigma \in L(t_*)$ for any real k, and if σ', $\sigma'' \in L(t_*)$, then $\sigma' + \sigma'' \in L(t_*)$, because the corresponding curves y', y'' can be traced consecutively one after another. Thus, $L(t_*)$ is a linear subspace in $\mathbb{R}^m$.

Lemma 3.2. *For any $\sigma \in L(t_*)$, there exists a sequence of Lip-continuous functions $y_n(t)$ supported on an interval $\Delta_n \to t_*$ such that $||\dot{y}_n||_1 \leqslant O(1)$, and*

$$\int_0^T \big(V(t)\, y_n(t), \dot{y}_n(t)\big)\, dt \;\to\; \sigma. \tag{3.1}$$

Proof. Fix any $\hat{\sigma} \in L(t_*)$ and a corresponding function $\hat{y}(\tau) \in \mathrm{Lip}\,[0,1]$. Take any sequence of intervals $\Delta_n = [t_1, t_2]$ containing the point t_* such that $\delta_n = t_2 - t_1 \to 0$, and define the functions $y_n(t) = \hat{y}(\frac{t-t_1}{\delta_n})$ for $t \in \Delta_n$ and $y_n(t) = 0$ otherwise. Clearly, $\sigma(y_n) = \hat{\sigma}$ for all n; moreover, $||y_n||_C = ||\hat{y}||_C \leqslant O(1)$ and $||\dot{y}_n||_1 = ||\dot{\hat{y}}||_1 \leqslant O(1)$. Finally, since the matrix V is continuous, whence $\max|V(t) - V(t_*)| \to 0$ on Δ_n, we obviously have

$$\int_0^T \Big|\big((V(t) - V(t_*))\, y_n(t), \dot{y}_n(t)\big)\Big|\, dt \;\leqslant\; o(1) \cdot ||y_n||_C \cdot ||\dot{y}_n||_1 \;\to\; 0,$$

hence the sequence y_n satisfies (3.1), q.e.d. □

Now, take an arbitrary finite set $\theta = \{t_1, \ldots, t_N\}$ in $[0,T]$, define a subspace $L(\theta) = \sum L(t_i)$, and then a subspace

$$\mathscr{L} = \bigcup_\theta L(\theta) \;=\; \sum_\theta L(\theta),$$

where the union and the sum is taken over all possible finite θ-s.

The last equality holds because for any t_1, t_2 we obviously have $L(t_1) \cup L(t_2) \subset L(t_1) + L(t_2) = L(t_1, t_2)$.

The key fact in the proof of Theorem 3.1 is the following

Lemma 3.3. *Suppose that, for any $\bar w \in K_0$,*

$$\max_{\lambda\in M} \Omega[\lambda](\bar w)) \;\geqslant\; 0. \tag{3.2}$$

Then, for any $\bar w \in K_0$ and any $\sigma \in \mathscr{L}$,

$$\max_{\lambda\in M} \Big(\Omega[\lambda](\bar w) + (\lambda,\sigma)\Big) \;\geqslant\; 0. \tag{3.3}$$

Proof. Take any $\bar w = (\bar\xi, \bar y, \bar u) \in K_0$ and any $\sigma \in \mathscr{L}$. Let y_n be a sequence from Lemma 3.2 that generates σ by (3.1), and let $w_n = (\xi_n, y_n, u_n)$ be the corresponding sequence from equations (1.11) and (2.3) with $\xi_n(0) = 0$. Since $||y_n||_1 \to 0$, by the Gronwall lemma $||\xi_n||_C \to 0$, whence also $\Omega(w_n) \to \sigma$, i.e.. $\Omega[\lambda](w_n) \to (\lambda,\sigma)$ for any $\lambda \in \mathbb{R}^m$.

Now, we consider the sequence of functions $\Omega[\lambda](\bar w + w_n)$, where $\lambda \in M$ and $n = 1,2,\ldots$. The mixed terms in their decomposition are

$$\int_0^T \Big((Q[\lambda]\bar\xi, \xi_n) + (P_1[\lambda]\bar\xi, y_n) + (P_1[\lambda]\xi_n, \bar y) + (R_1[\lambda]\bar y, y_n) + \\ + (V[\lambda]y_n, \bar u) + (V[\lambda]\bar y, u_n)\Big)\, dt.$$

Since $||y_n||_1 \to 0$ and $||\xi_n||_C \to 0$, the first five terms here tend to zero. The last term can be reduced to the preceding ones by the integration by parts:

$$\int_0^T (V[\lambda]\bar y, u_n)\, dt \;=\; -\int_0^T (\dot V[\lambda]\bar y, y_n)\, dt - \int_0^T (V[\lambda]\bar u, y_n)\, dt \;\to\; 0.$$

Recall that matrices $V[\lambda](t)$ are Lip-continuous, so $|\dot V[\lambda](t)| \leqslant \mathrm{const}$. Therefore,

$$\Omega[\lambda](\bar w + w_n) = \Omega[\lambda](\bar w) + \Omega[\lambda](w_n) + o(1) \;\to\; \Omega[\lambda](\bar w) + (\lambda,\sigma).$$

Note that the sequence $\bar w + w_n$ may not belong to the cone K_0 given by relations (2.4). However, since $\bar w$ satisfies these relations, they can be violated only by $\mathscr{O}(|\xi_n(T)|) \to 0$, hence, by the Hoffman lemma (see Appendix) there exists a sequence $w_n' \in K_0$ such that $||w_n' - (\bar w + w_n)||_{\mathfrak{U}} \to 0$. Since $\Omega[\lambda]$ is Lip-continuous on any bounded set in $\mathfrak{U}$, we have

$$\Omega[\lambda](w_n') \;=\; \Omega[\lambda](\bar w + w_n) + o(1) \;\to\; \Omega[\lambda](\bar w) + (\lambda,\sigma).$$

Applying (3.2) for $w_n' \in K_0$, we get in the limit:

$$\max_{\lambda\in M} (\lambda,\, \Omega(\bar w) + \sigma) \;\geqslant\; 0, \qquad \text{q.e.d.} \qquad \Box$$

Since the last inequality holds for all $\sigma \in \mathscr{L}$, we have

$$\inf_{\sigma\in\mathscr{L}} \max_{\lambda\in M} (\lambda,\, \Omega(\bar w) + \sigma) \;\geqslant\; 0.$$

Here, the expression under the max sign is linear both in λ and σ, the sets M and $\mathscr{L}$ are convex, moreover, M is compact. Then, by the Neumann minimax theorem [28] we can interchange the operations of *inf* and *max*, thus obtaining $\max_{\lambda\in M} \inf_{\sigma\in\mathscr{L}} (\lambda,\, \Omega(\bar w) + \sigma) \;\geqslant\; 0$. Since $\mathscr{L}$ is a subspace, any $\lambda \notin \mathscr{L}^\perp$ gives $\inf\limits_{\sigma\in\mathscr{L}} (\lambda,\, \Omega(\bar w) + \sigma) \;=\; -\infty$, so it can be removed from the max-operation. On the other hand, any $\lambda \in \mathscr{L}^\perp$ gives

$$\inf_{\sigma\in\mathscr{L}} (\lambda,\, \Omega(\bar w) + \sigma) = (\lambda,\, \Omega(\bar w)),$$

so we get

$$\max_{\lambda\in M\cap\mathscr{L}^{\perp}} (\lambda,\,\Omega(\bar w)) \,\geqslant\, 0, \tag{3.4}$$

which holds for any $\bar w\in K_0$. In particular, it follows that $M\cap\mathscr{L}^{\perp}$ is nonempty.

Now we note that for any $\lambda\in\mathscr{L}^{\perp}$ the matrix (λ,V) is such that, for all $t_*\in[0,T]$ and all cycles $y(\tau)$ on $[0,1]$ with $y(0)=y(1)=0$ we have $\oint((\lambda,V(t_*))\,y,dy)=0$, which obviously implies $(\lambda,V(t_*))=0$. Then the matrix $(\lambda,V(t))$ is identically zero, so the functional $\Omega[\lambda]$ does not involve the control u.

Thus, like in the case of $M=\{\cdot\}$ (see Sec. 2.2), the control u is now removed from the functionals $\Omega[\lambda]$ for all $\lambda\in\mathscr{L}^{\perp}$, and we again can pass to the independent variables $(\xi_0,\,y(\cdot),\,h)$ from the extended space $H=\mathbb{R}^n\times L_2[0,T]\times\mathbb{R}^r$, where y is a new control, ξ is a new state variable, and h an additional parameter. The cone K_0 can now be replaced by the cone K given by relations (2.10)–(2.11), and so, we come to the condition

$$\max_{\lambda\in M\cap\mathscr{L}^{\perp}} (\lambda,\,\Omega(\bar w)) \,\geqslant\, 0, \qquad \forall\,\bar w\in K. \tag{3.5}$$

This was the first step in the proof of Theorem 3.1.

The second step concerns condition b) in (2.7) and can be more clearly presented in the following abstract setting.

3.1. **An abstract refinement theorem.** Let H be a Hilbert space, in which we are given a finite number of quadratic functionals

$$\Omega_i(x)=(Q_ix,x),\quad x\in H,\qquad i=1,\dots,m,$$

where each $Q_i: H\to H$ is a linear symmetric operator. Thus, we have a vector-valued quadratic mapping $\Omega(x)=\big(\Omega_1(x),\,\dots,\,\Omega_m(x)\big)$. For any vector $\lambda\in\mathbb{R}^m$, define the quadratic functional

$$\Omega[\lambda](x)=(\lambda,\Omega(x))\,=\sum_i\lambda_i\,\Omega_i(x)\,=\sum_i\lambda_i(Q_ix,x).$$

Let S be a convex compact set in $\mathbb{R}^m$. Define the functional

$$\Omega[M](x)=\max_{\lambda\in S}(\lambda,(Qx,x)),$$

and let S_0 be the set of all $\lambda\in S$ such that the functional $\Omega[\lambda](x)$ is lower semi-continuous in the weak topology (w.l.s.c.) in H. Obviously, S_0 is a convex closed subset of S (may be empty). The last is true because any weakly converging sequence $x_n\to x_0$ is bounded, the convergence $\Omega[\lambda_m](x)\to\Omega[\lambda_0](x)$ for $\lambda_m\to\lambda_0$ is uniform on any bounded set, and the uniform limit of semi-continuous functions is semi-continuous.

Define the set $\mathscr{R}$ of all vectors $\sigma\in\mathbb{R}^m$ such that $\Omega(x_n)\to\sigma$ for some sequence $x_n\overset{wk}{\rightharpoonup}0$. Obviously, it is a cone, and we claim that it is convex. Indeed, let $x_n\overset{wk}{\rightharpoonup}0$, $\Omega(x_n)\to\sigma$, and $y_n\overset{wk}{\rightharpoonup}0$, $\Omega(y_n)\to\rho$. Then there exists a subsequence $k_n\to\infty$ such that for any i we have $(Q_ix_n,y_{k_n})\to0$, whence

$$\Omega(x_n+y_{k_n})=\Omega(x_n)+\Omega(y_{k_n})+o(1)\;\to\;\sigma+\rho,\qquad\text{q.e.d.}$$

Denote by $\mathscr{R}^*$ the dual (conjugate) cone to $\mathscr{R}$.

Lemma 3.4. *The cone $\mathscr{R}^*$ consists of all $\alpha\in\mathbb{R}^m$ such that the functional $(\alpha,\Omega(x))$ is weakly lower semi-continuous.*

Proof. Indeed, if $\alpha \in \mathscr{R}^*$, then, for any $x_n \xrightarrow{wk} 0$, $\liminf\,(\alpha, \Omega(x_n)) \to 0$. Otherwise, choosing a subsequence, we obtain $\Omega(x_n) \to \sigma \in \mathscr{R}$ and $(\alpha, \sigma) < 0$, a contradiction.

Conversely, if $(\alpha, \Omega(x))$ is w.l.s.c., then for any $\sigma \in \mathscr{R}$ and a corresponding sequence $x_n \xrightarrow{wk} 0$ with $\Omega(x_n) \to \sigma$, we have $(\alpha, \sigma) = \lim\,(\alpha, \Omega(x_n)) \geqslant 0$, i.e., $\alpha \in \mathscr{R}^*$, q.e.d. $\square$

Finally, let K be a cone in H given by a finite number of linear inequalities $(a_i, x) \leqslant 0$, $i = 1, \ldots, s$, where all $a_i \in H$.

Lemma 3.5. *For any $h \in K$ and any $\sigma \in \mathscr{R}$, there exists a sequence $x'_n \in K$ such that $\Omega(x'_n) \to \Omega(h) + \sigma$.*

Proof. Take any sequence $x_n \xrightarrow{wk} 0$ with $\Omega(x_n) \to \sigma$. Clearly, $\Omega(h + x_n) \to \Omega(h) + \sigma$. Note that we cannot guarantee that $h + x_n \in K$. However, since $(a_i, h) \leqslant 0$ for all i, then $(a_i,\, h + x_n) \leqslant o(1)$, whence by the Hoffman lemma (see Appendix) there is a sequence $x'_n \in K$ such that $||x'_n - (h + x_n)|| \to 0$. Since Ω is Lip-continuous on any bounded set, we get

$$\Omega(x'_n) = \Omega(h + x_n) + o(1) \;\to\; \Omega(h) + \sigma, \qquad \text{q.e.d.} \qquad \square$$

The following assertion, similar to Theorem 3.1 and proved in [30], enjoys perhaps even wider generality. For the reader's convenience, we give the proof here.

Theorem 3.6. *If $\Omega[S](x) \geqslant 0$ on K, then $S_0 = S \cap \mathscr{R}^*$ is nonempty and $\Omega[S_0](x) \geqslant 0$ on K. The converse is trivial.*

Proof. Let $\Omega[S](x) \geqslant 0$ on K. Fix any $h \in K$. By Lemma 3.5, for any $\sigma \in \mathscr{R}$, there exists a sequence $x'_n \in K$ such that $\Omega(x'_n) \to \Omega(h) + \sigma$. Since

$$\Omega(S)(x'_n) \;=\; \max_{\lambda \in S}\,(\lambda, \Omega(x'_n) \geqslant 0,$$

we have in the limit:

$$\max_{\lambda \in S}\,(\lambda,\, \Omega(h) + \sigma) \;\geqslant\; 0,$$

and then

$$\inf_{\sigma \in \mathscr{R}}\, \max_{\lambda \in S}\,(\lambda,\, \Omega(h) + \sigma) \;\geqslant\; 0.$$

Here, the expression under the max-operation is linear both in λ and σ, the sets S and $\mathscr{R}$ are convex, moreover, S is compact. Then, applying again the minimax theorem [28], we interchange the operations of *inf* and *max*:

$$\max_{\lambda \in S}\, \inf_{\sigma \in \mathscr{R}}\,(\lambda,\, \Omega(h) + \sigma) \;\geqslant\; 0.$$

Now, since $\mathscr{R}$ is a cone, any $\lambda \notin \mathscr{R}^*$ gives $\inf\limits_{\sigma \in \mathscr{R}}\,(\lambda,\, \Omega(h) + \sigma) = -\infty$, so it can be removed from the max-operation. On the other hand, any $\lambda \in \mathscr{R}^*$ gives $\inf\limits_{\sigma \in \mathscr{R}}\,(\lambda,\, \sigma) = 0$, so we obtain

$$\max_{\lambda\, \in S \cap \mathscr{R}^*}\,(\lambda,\, \Omega(h)) \;\geqslant\; 0.$$

This holds for any $h \in K$. In particular, it follows that $S_0 = S \cap \mathscr{R}^*$ is nonempty. Theorem 3.6 is proved. $\square$

Proof of Theorem 3.1. Recall that we study a family of functionals $\Omega[\lambda](w)$ of the form (1.8) on the cone K_0 of the form (1.9), where λ belongs to a convex compact set $M = \operatorname{co}\Lambda \subset \mathbb{R}^m$.

By the Goh transformation, this situation is reduced to functionals of the form (2.6), considered on the cone (2.4).

Assume that $\Omega[M](\bar{w}) \geqslant 0$ on K_0. Lemma 3.3 and (3.4) imply that $\Omega[S](\bar{w}) \geqslant 0$ on K_0, where $S = M \cap \mathscr{L}^{\perp}$ consists of all $\lambda \in M$ such that $\Omega[\lambda]$ satisfies condition a) in (2.7), i.e. $V[\lambda](t) = 0$.

Further, since the original control $\bar{u}$ is eliminated, we can pass to the new control $\bar{y} \in L_2^r[0,T]$ and obtain $\Omega[S](\bar{w}) \geqslant 0$ on the cone K in the extended space H. Then by Theorem 3.6 we have $\Omega[E(S)](\bar{w}) \geqslant 0$ on K, where $E(S) = S \cap \mathscr{R}^*$, and $\mathscr{R}^*$ is the set of all $\lambda \in \mathbb{R}^m$ such that $\Omega[\lambda]$ is w.l.s.c. in the space H.

Finally, equality $E(S) = M \cap \mathscr{L}^{\perp} \cap \mathscr{R}^* = G(M)$ is guaranteed by the following simple fact.

Lemma 3.7. *The quadratic functional* (2.8) *is w.l.s.c. in* $y \in L_2^r[0,T] \iff$ *it satisfies the Legendre condition b) in* (2.7)*, i.e.* $R(t) \geqslant 0$ *a.e. on* $[0,T]$.

Thus, $\Omega[\,G(M)](\bar{w}) \geqslant 0$ on K. This accomplishes the proof of Theorem 3.1.

Combining this result with assertion $a)$ of Theorem 1.2, we obtain the following

Theorem 3.8. *If* $\widehat{w}$ *delivers a weak minimum, then*

$$\Omega[G(co\Lambda)](\bar{w}) \;\geqslant\; 0 \qquad \textit{for all } \bar{w} \in K. \tag{3.6}$$

In Section 3.3 below we consider the case when $G(co\Lambda)$ can be replaced by a more narrow set $G(\Lambda)$.

3.2. **Refinement of the sufficient condition.** Consider the sufficient condition $b)$ of Theorem 1.2 (condition (1.13)) and the similar conditions $b')$, obtained from that by replacing the set Λ with $M = \operatorname{co}\Lambda$, i.e., there exists a constant $c > 0$ such that

$$\Omega[\operatorname{co}\Lambda](\bar{w}) \;\geqslant\; c\,\gamma(\bar{w}) \qquad \text{for all } \bar{w} \in K_0\,. \tag{3.7}$$

Since the function $\Omega[\lambda](\bar{w})$ is linear in λ, its maximum over Λ coincides with that over $\operatorname{co}\Lambda$, so conditions $b)$ and $b')$ are equivalent. Now, we suppose (3.7) holds. Define a family of quadratic functionals $\widetilde{\Omega}[\lambda](\bar{w}) = \Omega[\lambda](\bar{w}) - c\,\gamma(\bar{w})$, where $(\lambda, c) \in \widetilde{M} = M \times \{c\}$ lies in $\mathbb{R}^m \times \mathbb{R}$ and $m = \dim\lambda$. Since γ has the form (1.10), this functional is of the same form, hence condition (3.7) implies $\widetilde{\Omega}[\widetilde{M}](\bar{w}) \geqslant 0$. Then, by Theorem 3.1 we can replace the set $\widetilde{M}$ with $G(\widetilde{M})$, so condition (3.7) is equivalent to the following one:

$$\widetilde{\Omega}[\,G(\widetilde{M})](\bar{w}) \;\geqslant\; 0 \qquad \text{for all } \bar{w} \in K. \tag{3.8}$$

The set $G(\widetilde{M})$ consists of all $(\lambda, c) \in \widetilde{M}$ such that $\widetilde{\Omega}[\lambda]$ satisfies conditions (2.7), i.e. $\Omega[\lambda]$ satisfies conditions $V[\lambda](t) = 0$ and $R[\lambda](t) \geqslant c$. Therefore, (3.8) means that

$$\Omega[\,G_c(M)](\bar{w}) \;\geqslant\; c\,\gamma(\bar{w}) \qquad \text{for all } \bar{w} \in K, \tag{3.9}$$

where the set $G_c(M)$ consists of all $\lambda \in M$ such that $V[\lambda](t) = 0$ and $R[\lambda](t) \geqslant c$ a.e. on $[0,T]$, or, in other words, of all $\lambda \in G(M)$ such that $R[\lambda](t) \geqslant c$ a.e. on $[0,T]$. Obviously, (3.9) implies that

$$\Omega[\,G(M)](\bar{w}) \;\geqslant\; c\,\gamma(\bar{w}) \qquad \text{for all } \bar{w} \in K. \tag{3.10}$$

An advantage of conditions (3.9) and (3.10), which are equivalent to (1.13), is that the corresponding functionals might possibly be simpler than that in (1.13), because the maximization in λ is performed over more narrow sets.

3.3. **On possible strengthening the necessary conditions.** Consider the case when the equality constraints (1.3)–(1.4) are non-degenerate, i.e. when the set Λ does not contain a triple $(\alpha_0,\alpha,\beta)\in\mathbb{R}^{1+d(\varphi)}\times\mathbb{R}^{d(\eta)}$ with $\alpha_0=0,\ \alpha=0$, so that only $\beta\neq 0$. In this case, the function $|\beta|/(\alpha_0+|\alpha|)$, being continuous, is bounded on the compact set Λ from above: $|\beta|/(\alpha_0+|\alpha|)\leqslant C$, i.e. $|\beta|\leqslant C(\alpha_0+|\alpha|)$ with some C. Hence, the normalization $\alpha_0+|\alpha|+|\beta|=1$ is equivalent to $\alpha_0+|\alpha|=1$, i.e. to $\sum_{i=0}^{d(\varphi)}\alpha_i=1$. In the last normalization, the set Λ is convex, so $G(\operatorname{co}\Lambda)=G(\Lambda)$. In the original (and any equivalent) normalization we can assert that

$$G(\operatorname{co}\Lambda) \;=\; \operatorname{co}(G(\Lambda)).$$

However, since $\Omega[\lambda]$ is linear in λ, we have $\Omega[\operatorname{co}(G(\Lambda))](\bar w)=\Omega[G(\Lambda)](\bar w)$, and therefore, Theorem 3.8 reads as follows: if $\widehat w$ delivers a weak minimum, then

$$\Omega[G(\Lambda)](\bar w)\;\geqslant\;0 \qquad \text{for all } \bar w\in K. \tag{3.11}$$

In the degenerate case, i.e., when the set Λ contains a triple (α_0,α,β) with $\alpha_0=0,\ \alpha=0$, only the following inclusion holds true:

$$G(\operatorname{co}\Lambda)\;\supset\;\operatorname{co}(G(\Lambda)).$$

Indeed, $\operatorname{co}\Lambda\supset\Lambda$, then $G(\operatorname{co}\Lambda)\supset G(\Lambda)$, and since the first set equals $G\cap\operatorname{co}\Lambda$, it is convex, whence $G(\operatorname{co}\Lambda)\supset\operatorname{co}(G(\Lambda))$.

Thus, if the equality constraints in Problem A are degenerate, our considerations do not allow to replace $G(\operatorname{co}\Lambda)$ with $G(\Lambda)$ in Theorem 3.8. Nevertheless, as was shown by Milyutin [26], this replacement can be justified, but by quite another, essentially more difficult technique. The last one is based on the replacement of some degenerate equality constraints by some specific (not straightforward!) inequality constraints, so that the remaining equality constraints become non-degenerate. Milyutin proved the following

Theorem 3.9. *If $\widehat w$ delivers a weak minimum in Problem A, then*

$$\Omega[\Lambda_+](\bar w)\;\geqslant\;0 \qquad \textit{for all } \bar w\in K, \tag{3.12}$$

where Λ_+ is the set of all $\lambda\in\Lambda$, such that $\Omega[\lambda](\bar w)\geqslant 0$ on a subspace $L_\lambda\subset W$ of a finite codimension c_λ that depend on λ.

It can be easily shown that $\Lambda_+\subset G(\Lambda)$, so this theorem implies (3.11). Moreover, it also implies the result of [13]. Later, some upper bounds on the codimension c_λ were obtained in [29]. Detailed review of all those results is far beyond the scope of this paper.

4. The Abstract Two-norm Approach

Now we pass to the proof of Theorem 1.2. To this aim, consider the following general abstract problem:

$$f_0(x)\to\min,\qquad f_i(x)\leqslant 0,\quad i=1,\dots,s;\qquad g(x)=0. \tag{4.1}$$

Here X,Y are Banach spaces, the functionals $f_i:X\to\mathbb{R}$ and the mapping $g:X\to Y$ are defined in an open set $\mathscr{D}\subset X$, and we are interested in the local minimality at a given point $x_0\in\mathscr{D}$.

For now, we impose the following first order assumptions:

A1) the functionals $f_i\,,\ i=0,1,\dots,s$, are Frechet differentiable at x_0, the mapping g is strictly differentiable at x_0 (smoothness of the data functions);

A2) the image of the derivative $g'(x_0)$ is closed in Y (weak regularity of equality constraint). Under these assumptions, the following well known Lagrange multipliers rule holds.

Theorem 4.1. *Let x_0 be a local minimum in problem (4.1). Then there exist Lagrange multipliers $\alpha_i \geqslant 0,\ i = 0,1,\ldots,s,$ and $y^* \in Y^*$, satisfying the normalization $\sum_i \alpha_i + ||y^*|| = 1$, the complementary slackness conditions*

$$\alpha_i f_i(x_0) = 0, \qquad i = 1,\ldots,s, \tag{4.2}$$

and such that the Lagrange function

$$\mathscr{L}(x) = \alpha_0 f_0(x) + \sum_{i=1}^{s} \alpha_i f_i(x) + y^* g(x)$$

is stationary at x_0: $\mathscr{L}'(x_0) = 0$, i.e.,

$$\alpha_0 f_0'(x_0) + \sum_{i=1}^{s} \alpha_i f_i'(x_0) + y^* g'(x_0) = 0. \tag{4.3}$$

The set of all tuples $\lambda = (\alpha_0, \alpha, y^*)$ satisfying the above properties, where $\alpha = (\alpha_1, \ldots, \alpha_s)$, will be denoted by $\Lambda(x_0)$, or simply Λ. Theorem 4.1 claims that, at any local minimum this set is nonempty. These conditions are of first order, because they involve only first derivative of all data functions in the problem at a given point.

To obtain second order conditions, we will use the well-known *two-norm approach* (see, e.g. [23, 24, 25]), though in a bit modified form. To describe it, we assume that, along with the norm $||x||$, the space X is also equipped by another norm $||x||_a$ which is weaker than the basic one: $||x||_a \leqslant \text{const}\,||x||$. Naturally, we also have to impose some assumptions of the second order.

Let be given a mapping $F: X \to Z$ between two Banach spaces, defined in a neighborhood of a point $x_0 \in X$. Introduce the following notion. (Which will be used for mappings to the spaces $\mathbb{R}$ and Y.)

Definition 4.2. We say that the mapping $F: X \to Z$ admits an expansion up to second order terms (or simply, a second order expansion) at the point x_0 if

$$F(x_0 + \bar{x}) = F(x_0) + F'(x_0)\,\bar{x} + \frac{1}{2}\,Q_F(\bar{x}) + r_F(\bar{x}) \qquad \forall\, \bar{x} \in X, \tag{4.4}$$

where $F'(x_0): X \to Y$ is a linear bounded operator (the Frechet derivative), and $Q_F(\bar{x}) = D_F(\bar{x},\bar{x})$, where $D_F: X \times X \to Y$ is a symmetric bilinear mapping satisfying the estimate

$$||D_F(\bar{x}_1, \bar{x}_2)|| \leqslant c_F\, ||\bar{x}_1||_a\, ||\bar{x}_2||, \qquad c_F = \text{const}\,, \tag{4.5}$$

and $||r_F(\bar{x}) = o(||\bar{x}||_a^2)$ as $||\bar{x}|| \to 0$.

Note that two different norms of $\bar{x}$ are used in the right-hand side of inequality (4.5). This is the point that differ our version of the two-norm approach from the standard one, where both $\bar{x}_1$ and $\bar{x}_2$ are taken with the weak norm $||\cdot||_a$. Therefore, our assumption (4.5) is more week than the standard one. The term $Q_F(\bar{x})$ will be called *the second variation* of the mapping F at x_0. Note that estimate (4.5) implies

$$||Q_F(\bar{x})|| \leqslant c_F\, ||\bar{x}||_a\, ||\bar{x}||. \tag{4.6}$$

Let x_0 be a stationary point of problem (4.1), i.e. have a nonempty $\Lambda(x_0)$. In addition to assumptions A1 and A2, we impose the following assumption:

A3) the data functions f_i and g have expansions up to second order terms at the point x_0.

Now, we introduce the quadratic order

$$\gamma(\bar{x}) \,=\, ||\bar{x}||_a^2.$$

Our aim is to obtain necessary and sufficient conditions of this order for a local minimum at x_0 in problem (4.1).

Let $I = I(x_0)$ be the set of all active indices at x_0. Without loss of generality, we may assume that $I = \{0, 1, \ldots, s\}$ and $f_0(x_0) = 0$. It follows from expansion (4.4) that, for any $\lambda \in \Lambda(x_0)$, the corresponding Lagrange function $L[\lambda](x)$ has the expansion

$$L[\lambda](x_0 + \bar{x}) = L[\lambda](x_0) + L'[\lambda](x_0)\bar{x} + \frac{1}{2}\, Q_L[\lambda](\bar{x}) + r_L[\lambda](\bar{x}). \tag{4.7}$$

Since $L[\lambda](x_0) = 0$, and by (4.3) also $L'[\lambda](x_0) = 0$, we have

$$L[\lambda](x_0 + \bar{x}) \,=\, \frac{1}{2}\,\Omega[\lambda](\bar{x}) + r_L[\lambda](\bar{x}), \tag{4.8}$$

where the quadratic form

$$\Omega[\lambda](\bar{x}) \,:=\, Q_L[\lambda](\bar{x}) \,=\, \sum_{i\in I} \alpha_i\, Q_i(\bar{x}) + y^* Q_g(\bar{x})$$

(which will be called *the second variation of the Lagrange function*) is generated by the bilinear form

$$D_L[\lambda](\bar{x}_1, \bar{x}_2) \,=\, \sum_{i\in I} \alpha_i D_i(\bar{x}_1, \bar{x}_2) + y^* D_g(\bar{x}_1, \bar{x}_2) \tag{4.9}$$

that satisfies (4.5), and the remaining term is

$$r_L[\lambda](\bar{x}) = \sum_{i\in I} \alpha_i\, r_i(\bar{x}) + y^* r_g(\bar{x}) \,=\, o(||\bar{x}||_a^2) \quad \text{as} \quad ||\bar{x}|| \to 0.$$

Also, we introduce the cone of critical variations

$$K = \{\, \bar{x} \mid f_i'(x_0)\,\bar{x} \leqslant 0,\ \forall\, i \in I,\ g'(x_0)\,\bar{x} = 0\,\}. \tag{4.10}$$

Note that it is always nonempty, since contains $\bar{x} = 0$.

Following Levitin–Milyutin–Osmolovskii [20], we introduce the function

$$\Omega[\Lambda](\bar{x}) \,=\, \sup_{\lambda\in\Lambda} \Omega[\lambda](\bar{x})$$

and the violation function $\sigma(x) := \sum_{i\in I} f_i^+(x) + ||g(x)||$, where $a^+ \,=\, \max\{a, 0\}$.

Theorem 4.3. *a) If x_0 is a local minimum, then*

$$\Omega[\Lambda](\bar{x}) \,\geqslant\, 0 \qquad \forall\, \bar{x} \in K. \tag{4.11}$$

b) If there exists $c > 0$ such that

$$\Omega[\Lambda](\bar{x}) \,\geqslant\, c\,\gamma(\bar{x}) \qquad \forall\, \bar{x} \in K, \tag{4.12}$$

then x_0 is a strict local minimum of the order γ. The last means that there exists $b > 0$ such that, in a neighborhood of x_0, the following estimate holds: $\sigma(x) \geqslant b\,\gamma(x - x_0)$.

Note that, if the two norms coincide: $||x||_a = ||x||$, this theorem reduces to a well known theorem from [20].

5. Proof of Theorem 4.3

Let us pass to the proof. We begin with assertion a), the necessity. First of all, note that if $g'(x_0)X \neq Y$, then Λ contains an element $\hat{\lambda}$ with all $\alpha_i = 0$ and $y^* \neq 0$. Then also $-\hat{\lambda} \in \Lambda$, and hence, for any $\bar{x} \in X$ we have $\Omega[\Lambda](\bar{x}) \geqslant |\Omega[\hat{\lambda}](\bar{x})| \geqslant 0$, so the inequality (4.11) holds independently of the local minimality at x_0. Therefore, in below we assume that $g'(x_0)X = Y$.

Recall that, in this regular case, the Lyusternik theorem says (see e.g. [2]–[6]) that there is a neighborhood $\mathscr{O}(x_0)$ and a constant b such that for any $x \in \mathscr{O}(x_0)$ there exists a correction $\tilde{x}$ such that $g(x+\tilde{x}) = 0$ and $||\tilde{x}|| \leqslant b\,||g(x)||$.

Lemma 5.1. *If x_0 is a local minimum, then, for any $\bar{x} \in K$, the following system is incompatible with respect to h:*

$$\begin{aligned} f_i'(x_0)\,h + Q_i(\bar{x}) &< 0, \\ g'(x_0)\,h + Q_g(\bar{x}) &= 0. \end{aligned} \tag{5.1}$$

Proof. Suppose the contrary: such $\bar{x}$ and h exist. Then $\exists C > 0$ such that

$$\begin{aligned} f_i'(x_0)\,h + Q_i(\bar{x}) &< -2C, \\ g'(x_0)\,h + Q_g(\bar{x}) &= 0. \end{aligned} \tag{5.2}$$

Multiplying the inequality in (4.10) by $\varepsilon > 0$, the system (5.2) by $\frac{1}{2}\,\varepsilon^2$, and summing the results up, we obtain the relations:

$$\begin{aligned} f_i'(x_0)(\varepsilon\bar{x} + \tfrac{1}{2}\,\varepsilon^2 h) + \tfrac{1}{2}\,\varepsilon^2 Q_i(\bar{x}) &< -C\varepsilon^2, \\ g'(x_0)(\varepsilon\bar{x} + \tfrac{1}{2}\,\varepsilon^2 h) + \tfrac{1}{2}\,\varepsilon^2 Q_g(\bar{x}) \; &= 0. \end{aligned} \tag{5.3}$$

Note that

$$Q_g\big(\varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h\big) = D_g\big(\varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h,\, \varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h\big) = \varepsilon^2 D_g(\bar{x},\bar{x}) + O(\varepsilon^3)$$

and $r_g(\varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h) = o(\varepsilon^2)$. Setting $x_\varepsilon = x_0 + \varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h$, we obtain

$$g(x_\varepsilon) = g(x_0) + g'(x_0)\big(\varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h\big) + \frac{1}{2}\,\varepsilon^2 Q_g(\bar{x}) + o(\varepsilon^2) = o(\varepsilon^2) \tag{5.4}$$

in view of (5.3). By the Lyusternik theorem (see, e.g., [3, 6, 27]), there is a correction $\tilde{x}_\varepsilon$ such that $g(x_\varepsilon + \tilde{x}_\varepsilon) = 0$ and $||\tilde{x}_\varepsilon|| = o(\varepsilon^2)$. Then, applying (4.4) to all f_i and taking into account (5.3), we obtain for $x_\varepsilon' = x_0 + \varepsilon\bar{x} + \frac{1}{2}\,\varepsilon^2 h + \tilde{x}_\varepsilon$ with all sufficiently small $\varepsilon > 0$:

$$f_i(x_\varepsilon') = f_i(x_0 + \varepsilon\bar{x} + \tfrac{1}{2}\,\varepsilon^2 h + \tilde{x}_\varepsilon) =$$
$$= f_i(x_0) + f_i'(x_0)(\varepsilon\bar{x} + \tfrac{1}{2}\,\varepsilon^2 h + \tilde{x}_\varepsilon) + \tfrac{1}{2}\,\varepsilon^2 D_i(\bar{x},\bar{x}) + o(\varepsilon^2) \leqslant -C\varepsilon^2 + o(\varepsilon^2) < 0.$$

So, the point x_ε' satisfies all the constraints and gives strictly smaller value to the cost than x_0 does, which contradicts the local minimality at x_0.

Thus, Lemma 5.1 is proved. □

Now, we get back to the proof of assertion a) in Theorem 4.3. Fix any $\bar{x} \in K$ and consider system (5.1) with respect to h. It is a particular case of a general linear system of the form

$$\begin{aligned} (p_i, x) + \xi_i &< 0, \qquad i \in I, \\ Ax + \eta \; &= 0, \end{aligned} \tag{5.5}$$

where x is an element of a Banach space X, all $p_i \in X^*$, all ξ_i are real numbers, I is a finite set of indices, $A : X \to Y$ is a linear bounded surjective operator, Y is a Banach space, and $\eta \in Y$.

Lemma 5.2. *System* (5.5) *is incompatible* $\Longleftrightarrow$ *there exist multipliers* $\alpha_i \geqslant 0$, $i \in I$, *and* $y^* \in Y^*$, *such that* $\sum \alpha_i > 0$,

$$\sum \alpha_i\, p_i + y^* A \;=\; 0, \tag{5.6}$$

and

$$\sum \alpha_i\, \xi_i \;+\; \langle y^*, \eta\rangle \;\geqslant\; 0. \tag{5.7}$$

The proof is relegated to Appendix.

Applying this lemma to system (5.1), we get multipliers $\alpha_i \geqslant 0$, $i \in I$, and $y^* \in Y^*$, normalized by $\sum \alpha_i + ||y^*|| = 1$, such that

$$\sum \alpha_i f_i'(x_0) \,+\, y^* g'(x_0) \;=\; 0, \tag{5.8}$$

and

$$\sum \alpha_i\, Q_i(\bar{x}) \;+\; \langle y^*, Q_g(\bar{x})\rangle \;\geqslant\; 0. \tag{5.9}$$

The first relation means that the collection $\lambda = (\alpha_i,\; i \in I,\; y^*)$ belong to Λ, and the second one that $\Omega[\lambda](\bar{x}) \geqslant 0$. Then, the more so, $\Omega[\Lambda](\bar{x}) \geqslant 0$, and this is true for any $\bar{x} \in K$. Thus, assertion a) of Theorem 4.3 (the necessity part) is proved.

Let us prove assertion b), i.e., the sufficiency part. Suppose there is no $\gamma-$minimum at x_0. This means that there exists a sequence $\delta x_n \neq 0$, $\delta x_n \to 0$, such that

$$f_i(x_0 + \delta x_n) \;\leqslant\; o(\gamma_n), \qquad i \in I,$$
$$g(x_0 + \delta x_n) \;=\; o(\gamma_n),$$

where we denote $\gamma_n := \gamma(\delta x_n) = ||\delta x_n||_a^2$.

Taking expansion (4.4) for each f_i and g in the above relations, we get

$$\begin{aligned} f_i'(x_0)\,\delta x_n + \tfrac{1}{2}\,Q_i(\delta x_n) + r_i(\delta x_n) \;&\leqslant\; o(\delta x_n),\\ g'(x_0)\,\delta x_n + \tfrac{1}{2}\,Q_g(\delta x_n) + r_g(\delta x_n) \;&=\; o(\gamma_n). \end{aligned} \tag{5.10}$$

According to (4.4), all $r_i(\delta x_n) = o(\gamma_n)$ and $r_g(\delta x_n) = o(\gamma_n)$. Since

$$||Q_i(\delta x_n)|| = ||D_i(\delta x_n, \delta x_n)|| \;\leqslant\; \text{const}\; ||\delta x_n||_a \cdot ||\delta x_n|| \;=\; o(\sqrt{\gamma_n}),$$

and similarly $|Q_g(\delta x_n)| = o(\sqrt{\gamma_n})$, relations (5.10) imply

$$\begin{aligned} f_i'(x_0)\,\delta x_n \;&\leqslant\; o(\sqrt{\gamma_n}),\\ g'(x_0)\,\delta x_n \;&=\; o(\sqrt{\gamma_n}). \end{aligned} \tag{5.11}$$

Thus, δx_n violates the cone K by the value $o(\sqrt{\gamma_n})$. By the Hoffman lemma (see Appendix), there exists a correction $\tilde{x}_n$ such that $||\tilde{x}_n|| \leqslant o(\sqrt{\gamma_n})$ and $\bar{x}_n := \delta x_n + \tilde{x}_n \in K$. Then

$$\gamma(\bar{x}_n) = ||\delta x_n + \tilde{x}_n||_a^2 \;\leqslant\; \gamma_n + 2\sqrt{\gamma_n} \cdot ||\tilde{x}_n||_a + ||\tilde{x}_n||_a^2 \;=\; \gamma_n + o(\gamma_n).$$

Consider system (5.10). Take any $\lambda \in \Lambda$. Multiplying (5.10) by α_i and y^*, respectively, and summing up in view of (4.3), we obtain

$$\frac{1}{2}\,\Omega[\lambda](\delta x_n) \;\leqslant\; -r_\lambda(\delta x_n) + o(\gamma_n) \;\leqslant\; o(\gamma_n),$$

where $\Omega[\lambda](\delta x_n) = \sum_i Q_i(\delta x_n) + Q_g(\delta x_n)$ and $r_\lambda(\delta x_n) = \sum r_i(\delta x_n) + r_g(\delta x_n) = o(\gamma_n)$. Then, for the sequence $\bar{x}_n = \delta x_n + \tilde{x}_n \in K$, we have

$$\begin{aligned}\Omega[\lambda](\bar{x}_n) = \; & \Omega[\lambda](\delta x_n) + 2D[\lambda](\delta x_n, \tilde{x}_n) + \Omega[\lambda](\tilde{x}_n) \leqslant \\ & \leqslant o(\gamma_n) + \text{const}\, ||\delta x_n||_a \cdot ||\tilde{x}_n|| + o(\gamma_n) = \\ & = o(\gamma_n) + \text{const}\, \sqrt{\gamma_n} \cdot o(\sqrt{\gamma_n}) = o(\gamma_n),\end{aligned}$$

where $D[\lambda](\delta x, \bar{x})$ is defined in (4.9). This estimate is true for any $\lambda \in \Lambda$. Since the set Λ is bounded, all the constants here are mutually bounded, whence

$$\Omega[\Lambda](\bar{x}_n) = \max_{\lambda \in \Lambda} \Omega[\lambda](\bar{x}_n) \leqslant o(\gamma_n),$$

which contradicts the inequality (4.12). Theorem 4.3 is completely proved. □

In the next section, we apply this theorem to the above class of optimal control problems (1.1)–(1.4).

REMARK. The concept of local minimality can be also formulated in terms of sequences. Namely, let, as before, $f_0(x_0) = 0$. Obviously, the point x_0 is a local minimum iff there is no sequence $x_n \to x_0$ such that $f_0(x_n) < 0$, $f_i(x_n) \leqslant 0$ for all $i \neq 0$, and $g(x_n) = 0$. This concept is not symmetric w.r.t. the cost f_0 and the inequality constraints f_i. The strongest among symmetrical necessary conditions (i.e., the closest to the local minimality) is the following one, proposed by Levitin–Milyutin–Osmolovskii in [20] and called the $s-$necessity (strongest necessity): there is no sequence $x_n \to x_0$ such that $f_i(x_n) < 0$ for all i, and $g(x_n) = 0$. (This is close to the *weak Pareto efficiency* in multi-criterial optimization). In fact, practically all known necessary conditions of local minimality follow not from the local minimality itself, but from the $s-$necessity. In particular, Theorem 4.1 and part a) of Theorem 4.3 follow from the $s-$necessity.

As to sufficient condition, they always guarantee not just local minimality, but *strict* local minimality, which means that there is no sequence $x_n \to x_0$, $x_n \neq x_0$, such that $f_i(x_n) \leqslant 0$ for all i, and $g(x_n) = 0$. This is the weakest condition, symmetric w.r.t. the cost and inequality constraints, that guarantees the local minimality.

Similar concepts are proposed in [20] at the $\gamma-$level, where γ is some positive functional with $\gamma'(0) = 0$. (Typically, γ is a quadratic functional.) A point x_0 delivers $\gamma-$necessity if for any $\varepsilon > 0$ there is no sequence $x_n \to x_0$ such that $f_i(x_n) < -\varepsilon\gamma(x_n - x_0)$ for all i, and $g(x_n) = 0$. A point x_0 delivers $\gamma-$sufficiency if there exists $c > 0$ and a neighborhood $\mathscr{O}(x_0)$ in which $\sigma(x) \geqslant c\,\gamma(x_n - x_0)$. A majority of necessary conditions of local minimality in different classes of optimization problems are in fact necessary conditions of some order γ, specific for each class, and follow not from the local minimality itself, but from the $\gamma-$necessity, and a majority of sufficient conditions are sufficient conditions of some order γ and imply not only local minimality, but also the $\gamma-$sufficiency. The gap between $\gamma-$necessity and $\gamma-$sufficiency cannot be reduced at the level γ. Further development of these concepts see in [20] and [27].

6. A General Problem Linear in the Control

Now we get back to Problem A and try to apply the abstract Theorem 4.3 to it. Here we take γ of the form (1.10) and (in view of second order conditions) the following alternative norm:

$$||w||_a^2 = \gamma(w) = |x(0)|^2 + \int_0^T |y(t)|^2\, dt + |y(T)|^2,$$

where $\dot y = u$ and $y(0) = 0$.

Clearly, this norm is weaker than the original one: $||w||_a \leqslant \text{const}\,||w||$. However, since Problem A is stated in the space $W = AC \times L_\infty$ of elements $w = (x,u)$, we rather quickly face a difficulty, caused by the fact that relation (1.4), being an equality constraint in the space L_1, does not satisfy assumption A3, because its second order expansion contains the term $(C(t)\,\bar x(t))\bar u(t)$ with some matrix $C(t)$, the symmetrized form of which is

$$(C(t)\,\bar x_1(t))\bar u_2(t) + (C(t)\,\bar x_2(t))\bar u_1(t),$$

that cannot be estimated in the L_1-norm by the product $||\bar w_1||_a\,||\bar w_2||$, where $\bar w_1 = (\bar x_1, \bar u_1)$ and $\bar w_2 = (\bar x_2, \bar u_2)$, and so, the estimate (4.5) does not hold here.

Because of this, we take as independent variables the initial state value $x(0)$ and the control $u(t)$, i.e. the pair $(x(0),u) \in \mathbb{R}^n \times L_\infty^r[0,T]$, while the endpoint state value $x(T)$ will be regarded as a function of them: $x(T) = \rho(x_0,u)$, according to the dynamic equation

$$\dot x(t) = f(t,x(t)) + F(t,x(t))\,u(t), \qquad x(0) = x_0\,. \tag{6.1}$$

Then, Problem A now takes the form:

$$J := \varphi_0(x_0,\, \rho(x_0,u)) \to \min, \tag{6.2}$$

$$\varphi_i(x_0,\, \rho(x_0,u)) \leqslant 0, \qquad i = 1,\dots,d(\varphi), \tag{6.3}$$

$$\eta_j(x_0,\, \rho(x_0,u)) = 0, \qquad j = 1,\dots,d(\eta), \tag{6.4}$$

which will be called Problem B. (Note that the mapping $\rho : \mathbb{R}^n \times L_\infty^r[0,T] \to \mathbb{R}^n$ has a finite-dimensional image.)

The weak minimality in Problem A is the local minimality with respect to the norm $||x||_{AC} + ||u||_\infty$. Obviously, on the admissible set, this norm is equivalent to the norm $|x(0)| + ||u||_\infty$ of the pair $(x(0),u)$ in the space $W = \mathbb{R}^n \times L_\infty$, so we may consider the local minimality in Problem B w.r.t. the last norm.

We have to check all assumptions of the above two-norm approach for this problem, provided that all the smoothness assumptions on the data functions are satisfied.

The fulfilment of the first order assumptions A1–A2 is obvious. The fulfilment of second order assumption A3 is analyzed in the next section.

6.1. **Second order expansion of the mapping ρ.** Let us find a second order expansion of the mapping $\rho : (x_0,u) \mapsto x(T)$ near a given point $(x_0,u) \in W = \mathbb{R}^n \times L_\infty^r[0,T]$. Let $x(t)$ be the corresponding solution of Cauchy problem (6.1). Taking a small increment $(\bar x_0, \bar u)$, we get an increment δx satisfying the relations (dropping the argument t) :

$$(x+\delta x)^\bullet = f(x+\delta x) + F(x+\delta x)(u+\bar u), \qquad (x+\delta x)\Big|_0 = x_0 + \bar x_0\,,$$

where $|\bar x_0| + ||\bar u||_\infty \to 0$. (Here we use the nominal norm in W, since we study the local minimality in this norm.) Obviously, $||\delta x||_C \to 0$.

Expanding the right hand side with account of (6.1), we obtain

$$(\delta x)^\bullet = f'(x)\delta x + \frac{1}{2}(f''(x)\delta x, \delta x) + F(x)\bar{u} + r_f(\delta x) +$$
$$+ \big[F'(x)\delta x + \frac{1}{2}(F''(x)\delta x, \delta x) + r_F(\delta x)\big](u+\bar{u}), \qquad \delta x(0) = \bar{x}_0,$$

where $|r_f(t,\delta x)| + |r_F(t,\delta x)| \leqslant \alpha(\delta x)|\delta x|^2$ uniformly for $t \in [0,T]$ with some $\alpha(\delta x) \to 0$. Since F is a matrix, F' and F'' are some tensors of corresponding dimensions of the third and fourth range, respectively, by $(F'(x)\bar{x})\,u$ etc. we denote their natural action on the vectors δx, $\bar{x}$, $\bar{u}$.

Define a function $\bar{x}(t)$ satisfying the linear equation

$$\dot{\bar{x}} = f'(x)\bar{x} + (F'(x)\bar{x})\,u + F(x)\bar{u}, \qquad \bar{x}(0) = \bar{x}_0, \tag{6.5}$$

and let $\delta x = \bar{x} + \eta$. Then $||\bar{x}||_C \to 0$, while η satisfies the equation

$$\dot{\eta} = f'(x)\eta + \frac{1}{2}(f''(x)\bar{x},\bar{x}) + (f''(x)\bar{x},\eta) + \frac{1}{2}(f''(x)\eta,\eta) +$$
$$+ (F'(x)\bar{x})\,\bar{u} + (F'(x)\eta)(u+\bar{u}) + \frac{1}{2}(F''(x)\bar{x},\bar{x})(u+\bar{u}) + \tag{6.6}$$
$$+ \big[(F''(x)\bar{x},\eta) + \frac{1}{2}(F''(x)\eta,\eta)\big](u+\bar{u}) + R(\delta x, u, \bar{u}), \quad \eta(0) = 0,$$

where $R(\delta x,u,\bar{u}) = r_f(\delta x) + r_F(\delta x)(u+\bar{u})$ is bounded by

$$|R(\delta x,u,\bar{u})| \leqslant \alpha(\delta x)\cdot(1+|u+\bar{u}|)\cdot|\delta x|^2.$$

Let us make two important estimates. First, using the Goh transformation, we have $\bar{x} = \bar{\xi} + \bar{y}$, where $||\bar{y}||_2 \leqslant \mathscr{O}(\sqrt{\gamma})$ and $||\bar{\xi}||_C \leqslant \text{const}\,||\bar{y}||_1 \leqslant \mathscr{O}(\sqrt{\gamma})$ whence $||\bar{x}||_2 \leqslant \mathscr{O}(\sqrt{\gamma})$. This yields that $||R||_1 = \int_0^T |R|\,dt \leqslant o(\gamma)$.

Second, by the Gronwall lemma, $||\eta||_C$ is estimated by the integral of the right hand side of (6.6), except the terms $f'(x)\eta$ and $(F'(x)\eta)(u+\bar{u})$:

$$||\eta||_C \leqslant b\int_0^T \Big(|\bar{x}|^2 + |\bar{x}|\cdot|\eta| + |\eta|^2 + |\bar{x}|\cdot|\bar{u}| + |\bar{x}|^2\,|u+\bar{u}| + |R|\Big)\,dt \leqslant$$
$$\leqslant b\Big(||x||_2^2 + ||\eta||_C\cdot||\bar{x}||_1 + ||\eta||_C^2 + ||\bar{x}||_2\cdot||\bar{u}||_2 + ||\bar{x}||_2^2\cdot||u+\bar{u}||_\infty\Big) + o(\gamma),$$

with some constant b. Therefore, moving the terms with $||\eta||_C$ from the right to the left side, we have

$$||\eta||_C\big(1 - b||\bar{x}||_1 - b||\eta||_C\big) \leqslant b\Big(||\bar{x}||_2^2 + ||\bar{x}||_2\cdot||\bar{u}||_2 + ||\bar{x}||_2^2\cdot||u+\bar{u}||_\infty) + o(\gamma) \leqslant$$
$$\leqslant b\big(\gamma + o(\sqrt{\gamma}) + o(\gamma)\Big) = o(\sqrt{\gamma}).$$

Since $||\bar{x}||_1 + ||\eta||_C \to 0$, we may assume that $\big(1 - b||\bar{x}||_1 - b||\eta||_C\big) \geqslant 1/2$, whence obtaining the final estimate: $||\eta||_C \leqslant o(\sqrt{\gamma})$.

Therefore, $\delta x = \bar{x} + \eta$ is estimated by $||\delta x||_2 \leqslant ||\bar{x}||_2 + ||\eta||_2 \leqslant o(\sqrt{\gamma})$, and so, equation (6.6) for η becomes:

$$\dot{\eta} = f'(x)\eta + (F'(x)\eta)(u+\bar{u}) + (F'(x)\bar{x})\,\bar{u} +$$
$$+ \frac{1}{2}(f''(x)\bar{x},\bar{x}) + \frac{1}{2}(F''(x)\bar{x},\bar{x})u + R_\eta, \quad \eta(0) = 0,$$

where $\int_0^T |R_\eta(\delta x, u, \bar{u})|\, dt \leqslant o(\gamma)$.

Now, introduce the function $\bar{z}$ satisfying the equation (obtained by changing $\eta \mapsto \bar{z}$ in the above equation and dropping the terms $(F'(x)\bar{z})\bar{u}$ and R_η):

$$\dot{\bar{z}} = f'(x)\bar{z} + (F'(x)\bar{z})\,u + (F'(x)\bar{x})\bar{u} + \\ + \frac{1}{2}\,(f''(x)\bar{x},\bar{x}) + \frac{1}{2}\,(F''(x)\bar{x},\bar{x})u, \qquad \bar{z}(0) = 0. \tag{6.7}$$

Obviously, $\bar{z}$ quadratically depends on $(\bar{x}_0, \bar{u})$ and is bounded by

$$||\bar{z}||_C \leqslant \text{const}\left(||\bar{x}||_2^2 + ||\bar{x}||_1 \cdot ||\bar{u}||_\infty\right) \leqslant o(\sqrt{\gamma}).$$

Finally, setting $\eta = \bar{z} + \sigma$, we get

$$\dot{\sigma} = f'(x)\sigma + (F'(x)\sigma)(u+\bar{u}) + (F'(x)\bar{z},\bar{u}) + R_\eta\,, \qquad \sigma(0) = 0, \tag{6.8}$$

with the same remainder R_η as above. It follows that

$$||\sigma||_C \leqslant \text{const}\,||(F'(x)\bar{z},\bar{u})||_1 = o(\sqrt{\gamma}),$$

which is not completely good. However, we are interested mainly in the estimation of the terminal value of σ, so we will proceed as follows. Since equation (6.8) is linear in σ, there exists a Lip-continuous matrix $\mathscr{A}(t) = \mathscr{A}(t,x,u,\bar{u})$ such that

$$\sigma(T) = \int_0^T \mathscr{A}(t)(F'(x)\bar{z},\bar{u})\,dt + o(\gamma).$$

Moreover, there exists a Lip-continuous tensor $\Phi(t)$ such that $\mathscr{A}(t)(F'(x)\bar{z},\bar{u}) = (\Phi(t)\bar{z})\bar{u}$. Then, recalling that $\dot{\bar{y}} = \bar{u}$, $\bar{y}(0) = 0$, and ignoring $o(\gamma)$, we have

$$\sigma(T) = \int_0^T (\Phi(t)\,\bar{z},\bar{u})\,dt = (\Phi\bar{z},\bar{y})\Big|_T - \int_0^T (\dot{\Phi}(t)\,\bar{z},\bar{y})\,dt - \int_0^T (\Phi(t)\,\dot{\bar{z}},\bar{y})\,dt.$$

Taking into account (6.7) and the estimates for $\bar{z}$, $\bar{y}$, we obtain $|\sigma(T)| = o(\gamma)$.

Summing up, we proved the following.

Theorem 6.1. *If $|\bar{x}_0| + ||\bar{u}||_\infty \to 0$, then the corresponding increment of the solution to* (6.1) *is $\delta x = \bar{x} + \bar{z} + \sigma$, where $\bar{x}$ is its first variation (the Frechet derivative of the mapping ρ), satisfying equation* (6.5) *with $||\bar{x}||_C \to 0$ and $||\bar{x}||_2 + |\bar{x}(T)| \leqslant \mathscr{O}(\sqrt{\gamma})$, the function $\bar{z}$ is its second variation, satisfying equation* (6.7) *with $||\bar{z}||_C \leqslant o(\sqrt{\gamma})$, and the remainder term with $||\sigma||_C = o(\sqrt{\gamma})$ and $|\sigma(T)| = o(\gamma)$. Consequently, $|\delta x(T)| \leqslant \mathscr{O}(\sqrt{\gamma})$.*

6.2. **Checking the estimate of the mixed quadratic term.** Now, let a scalar twice smooth function $l(x_0,x_T)$ of arguments $(x_0,x_T) \in \mathbb{R}^{2n}$ be given. Consider the functional $\varphi(x_0,u) = l(x_0,x(T))$ with $x(T) = \rho(x_0,u)$, where ρ is the above mapping. We must check the required estimate (4.5) of the mixed quadratic term in the expansion of φ.

By Theorem 6.1, taking an increment $(\bar{x}_0,\bar{u})$, we obtain an increment $\delta x(T)$ and the following expansion:

$$\Delta\varphi = l'_{x_0}\bar{x}_0 + l'_{x_T}\,\delta x(T) + \\ + \frac{1}{2}\,(l''_{x_0,x_0}\bar{x}_0,\bar{x}_0) + (l''_{x_T,x_0}\bar{x}_0,\delta x(T)) + \frac{1}{2}\,(l''_{x_T,x_T}\,\delta x(T),\delta x(T)) + o(\gamma).$$

Since $\delta x(T) = \bar{x}(T) + \bar{z}(T) + \sigma(T)$, where

$$|\bar{x}(0)| + |\bar{x}(T)| \leqslant O(\sqrt{\gamma}), \qquad |\bar{z}(T)| \leqslant o(\sqrt{\gamma}), \qquad |\sigma(T)| = o(\gamma),$$

we can ignore all the terms with $\sigma(T)$ and the second order terms with $\bar{z}(T)$, whence obtaining

$$\Delta\varphi = l'_{x_0}\bar{x}_0 + l'_{x_T}(\bar{x}(T)+\bar{z}(T)) +$$
$$+\frac{1}{2}(l''_{x_0,x_0}\bar{x}_0,\bar{x}_0) + (l''_{x_T,x_0}\bar{x}_0,\bar{x}(T)) + \frac{1}{2}(l''_{x_T,x_T}\bar{x}(T),\bar{x}(T)) + o(\gamma). \tag{6.9}$$

It follows that the first variation of φ is $d\varphi = l'_{x_0}\bar{x}_0 + l'_{x_T}\bar{x}(T)$, and its second variation is

$$d^2\varphi = l'_{x_T}\bar{z}(T) + \frac{1}{2}(l''_{x_0,x_0}\bar{x}_0,\bar{x}_0) + (l''_{x_T,x_0}\bar{x}_0,\bar{x}(T)) + \frac{1}{2}(l''_{x_T,x_T}\bar{x}(T),\bar{x}(T)), \tag{6.10}$$

where $\bar{x}$ and $\bar{z}$ satisfy (6.5) and (6.7), respectively.

The only second order term that needs verification of the estimate (4.5) is $l'_{x_T}\bar{z}(T)$ which has the form

$$l'_{x_T}\bar{z}(T) = \int_0^T \mathscr{P}(t)(F'(x)\bar{x})\bar{u} = \int_0^T (C(t)\bar{x},\bar{u})\,dt$$

with some Lip-continuous matrices $\mathscr{P}$ and C. For any two elements $w' = (x'_0,u')$ and $\widetilde{w} = (\widetilde{x}_0,\widetilde{u})$ in the space W with the corresponding functions $x'(t)$ and $\widetilde{x}(t)$ obtained by the linear equation (6.5), the symmetric bilinear operator of (4.5) has the form

$$D(w',\widetilde{w}) = \int_0^T (C(t)x',\widetilde{u})\,dt + \int_0^T (C(t)\widetilde{x},u')\,dt.$$

Recall that $||x'||_2 + ||y'||_2 + |y'(T)| \leqslant ||w'||_a$ and $||\widetilde{x}||_C + ||\widetilde{u}||_\infty \leqslant ||\widetilde{w}||$. Then, the first integral is estimated by

$$\int_0^T |(C(t)x',\widetilde{u})|\,dt \leqslant ||x'||_1 \cdot ||\widetilde{u}||_\infty \leqslant \text{const}\,||w'||_a \cdot ||\widetilde{w}||,$$

which is OK. Since $(y')^\bullet = u'$ and $y'(0) = 0$, the second integral can be taken by parts:

$$\int_0^T (C(t)\widetilde{x},u')\,dt = (C\widetilde{x},y')\Big|_T - \int_0^T (\dot{C}(t)\widetilde{x},y')\,dt - \int_0^T (C(t)\dot{\widetilde{x}},y')\,dt.$$

The first two terms here are estimated by

$$\leqslant |\widetilde{x}(T)|\cdot|y'(T)| + ||\widetilde{x}||_C \cdot ||y'||_1 \leqslant \text{const}\,||\widetilde{w}||\cdot||w'||_a\,.$$

Since $\dot{\widetilde{x}} = A(t)\widetilde{x} + B(t)\widetilde{u}$, where A and B are some bounded matrices, $||\dot{\widetilde{x}}||_\infty \leqslant ||\widetilde{w}||$, so the last integral has the same estimate:

$$\int_0^T |(C(t)\dot{\widetilde{x}},y')|\,dt \leqslant \text{const}\,||\widetilde{w}||\cdot||w'||_a\,.$$

Summing up we finally have $|D(w',\widetilde{w})| \leqslant \text{const}\,||w'||_a \cdot ||\widetilde{w}||$, thus Problem B satisfies all assumptions (A1)–(A3) of the abstract problem (4.1), and so, falls into the scope of Theorem 4.1 and Theorem 4.3, the latter of which yields the following

Theorem 6.2. *a) If $(\hat{x},\hat{u})$ is a local minimum in Problem B, then*

$$\Omega_B[\Lambda](\bar{x}) \geqslant 0 \qquad \forall\,\bar{x}\in K_0\,. \tag{6.11}$$

b) If there exists $c>0$ such that

$$\Omega_B[\Lambda](\bar{x}) \geqslant c\,\gamma(\bar{x}) \qquad \forall\,\bar{x}\in K_0\,, \tag{6.12}$$

then $(\hat{x},\hat{u})$ is a local minimum of the order γ. The last means that there exists $b>0$ such that, in a neighborhood of $(\hat{x},\hat{u})$, the following estimate holds: $\sigma(x,u) \geqslant b\,\gamma(x-\hat{x},u-\hat{u})$.

6.3. **Two representations of the second variation.** Applying Theorem 4.3 to Problem B, we have to consider, for any $\lambda \in \Lambda$, the corresponding second variation of the functional $\varphi[\lambda](x_0,u) = l[\lambda](x_0,\rho(x_0,u))$, where $\rho(x_0,u) = x(T)$ is the above mapping $\mathbb{R}^n \times L^r_\infty[0,T] \to \mathbb{R}^n$. However, the main Theorem 1.2 uses another form (1.8) of the second variation. Here we show that both these forms coincide. (Below, we drop the argument λ.)

Consider the functional $\varphi(x_0,u) = l(x_0,\rho(x_0,u))$. Its second variation $d^2\varphi$ is given by (6.10), where $\bar{x}$ satisfies (6.5) and $\bar{z}$ satisfies (6.7). For brevity, introduce the matrix $A(t) = f'(x) + (F'(x)u$ (where the derivative of the matrix F is taken separately for each entry.) Then equations (6.5) and (6.7) have the form:

$$\dot{\bar{x}} = A\bar{x} + F(x)\bar{u}, \qquad \bar{x}(0) = \bar{x}_0\,, \tag{6.13}$$

$$\dot{\bar{z}} = A\bar{z} + \omega(t), \qquad \bar{z}(0) = 0, \tag{6.14}$$

where $\omega(t) = \frac{1}{2}\,(f''(x)\bar{x},\bar{x}) + \frac{1}{2}\,(F''(x)\bar{x},\bar{x})u + (F'(x)\bar{x})\bar{u}$. Since $\lambda \in \Lambda$, we can take the corresponding function $\psi(t)$ satisfying the relations

$$\dot{\psi} = -\psi A, \qquad \psi F(x) = 0, \qquad \psi(0) = l'_{x_0}\,, \qquad \psi(T) = -l'_{x_T}\,,$$

whence $\psi\,\omega = \frac{1}{2}\,(H''_{xx}\bar{x},\bar{x}) + (H''_{ux}\bar{x},\bar{u})$, and so

$$l'_{x_T}\,\bar{z}(T) \,=\, 0 - \psi(T)\bar{z}(T) \,=\, \psi(0)\bar{z}(0) - \psi(T)\bar{z}(T) \,=\, -\int_0^T (\psi\bar{z})^{\bullet}\,dt =$$

$$= \int_0^T \psi\,\omega\,dt \,=\, -\int_0^T \Big(\frac{1}{2}\,(H''_{xx}\bar{x},\bar{x}) + (H''_{ux}\bar{x},\bar{u})\Big)dt.$$

Summing this up with the second line of (6.9), we obtain exactly the second variation $\Omega_A(\bar{x},\bar{u})$ of Problem A in the form (1.8). Therefore, $d^2\varphi(\bar{x}_0,\bar{u}) = \Omega_A(\bar{x},\bar{u})$ for any pair $(\bar{x}_0,\bar{u})$ satisfying (6.13), whence Theorem 6.2 is equivalent to Theorem 1.2, and so, the last one is proved.

7. On More Fine Quadratic Orders

One may ask is it possible to use, instead of γ given by (1.10), the following seemingly "more natural" order

$$\gamma'(\bar{w}) = |\bar{x}(0)|^2 + \int_0^T |\bar{x}(t)|^2\,dt + |\bar{x}(T)|^2 \;?$$

The answer is: generally no. Consider the following

Example 2. $n = r = 1$, $\dot{x} = b(t)u - txu$, $x(0) = 0$,

$$J = \int_0^1 2xu\,dt + N\Big(x^2(1) + \int_0^1 x^2(t)\,dt\Big).$$

(The integral part of J can be reduced to a terminal term, but we do not do this.) Here $N > 0$ is an arbitrary number, and $b(t)$ is an arbitrary Lip-continuous function that equals 2^{-n} on each interval $\Delta_n = [2^{-n},\, 3\cdot 2^{-(n+1)}]$, $n = 1,2,\ldots$. Obviously, such a function does exist, because the distance between the consecutive intervals Δ_n and Δ_{n+1} is $2^{-n} - 3\cdot 2^{-(n+2)} = 2^{-(n+2)}$, which is not smaller than the twice increment of the function, $2^{-n} - 2^{-(n+1)} = 2\cdot 2^{-(n+2)}$.

The reference process is $\hat{x} = \hat{u} = 0$. The critical cone K is given by the relations $\dot{\bar{x}} = b(t)\bar{u}$, $\bar{x}(0) = 0$, and $\Omega = J$ on K. It can be shown that $\int_0^1 2\bar{x}\bar{u}\,dt \geqslant 0$ on K, whence $\Omega \geqslant N\gamma'$. But if one takes the sequence x_n of the form of isosceles triangle on Δ_n with the height $2^{-3n/2}$ one can find that the integral $\int_0^1 2\bar{x}\bar{u}\,dt$ is negative, and its absolute value is $> N\gamma'(x_n)$, so $J(x_n) < 0$. Therefore, the reference process does not provide a weak minimum.

However, let us get back to Problem A and select a special case when a more fine quadratic order does work. Recall that the $n\times r$-matrices $B(t)$ and $P^*[\lambda](t)$ are assumed to be Lip-continuous. Suppose now that there exists a measurable bounded $k\times r$-matrix $D(t)$, called *a divisor*, such that $B(t) = \widetilde{B}(t)D(t)$ and $P^*[\lambda](t) = \widetilde{P}^*[\lambda]D(t)$ for all $\lambda \in \Lambda$, where the $n\times k-$matrices $\widetilde{B}(t)$ and $\widetilde{P}^*[\lambda](t)$ are Lip-continuous. (Since the set Λ is finite-dimensional, it suffices to represent in such a way only a finite number of matrices $P^*[\lambda](t)$.)

Then Theorem 1.2 remains valid if one replace the order $\gamma(\bar{w})$ by the order $\gamma_D(\bar{w})$ defined by the same formula (1.10), where the additional state $\bar{y}$ is now defined by

$$\dot{\bar{y}} = D(t)\,\bar{u}, \qquad \bar{y}(0) = 0.$$

The proof remains the same modulo obvious alterations [22]. One can easily show that $\gamma_D(\bar{w}) \leqslant \text{const}\,\gamma(\bar{w})$ on K, but the reverse estimate does not hold in general.

8. Appendix

1. The Hoffman lemma (a simplified version). Let X, Y be Banach spaces, and let K be a cone in X given by linear relations:

$$(p_i, x) \leqslant 0, \quad i = 1,\ldots,m, \qquad Ax = 0,$$

where all $p_i \in X^*$ and $A : X \to Y$ is a linear surjective operator. Then there exists a constant C such that, for any $x \in X$, there exists $x' \in K$ such that $||x'-x|| \leqslant C\big(\sum (p_i,x)^+ + ||Ax||\big)$.

The proof can be found in, e.g., [4] and elsewhere.

2. Proof of Lemma 5.2. ($\Longrightarrow$) Let C_i, $i \in I$, be the sets defined by the corresponding inequalities, and E by the set defined by the equality in (5.5). If $\exists\, i_0$ such that $p_{i_0} = 0$ and $\xi_{i_0} \geqslant 0$, then C_{i_0} is empty. In this case, taking $\alpha_{i_0} > 0$, all other $\alpha_i = 0$ and $y^* = 0$, we get (5.8)–(5.9). If $p_{i_0} = 0$, but $\xi_{i_0} < 0$, then $C_{i_0} = X$, and setting $\alpha_{i_0} = 0$, the index i_0 can be completely excluded. Thus, we may assume that all $p_i \neq 0$, whence all C_i are nonempty. Clearly, the set E is nonempty.

Since $\cap_i C_i \cap E = \emptyset$, the Dubovitskii–Milyutin theorem says [1, 2] that there exist multipliers $\lambda_i \in X^*$, $i \in I$, and $\mu \in X^*$, not all equal zero, such that $\sum \lambda_i + \mu = 0$, and

$$\sum \inf(\lambda_i, C_i) \;+\; \inf(\mu, E) \;\geqslant\; 0. \tag{8.1}$$

Take any $i \in I$. We claim that $(p_i, \bar{x}) \leqslant 0$ implies $(\lambda_i, \bar{x}) \geqslant 0$. Indeed, otherwise $\exists\,\bar{x}$ such that $(p_i, \bar{x}) \leqslant 0$ and $(\lambda_i, \bar{x}) < 0$. Then, taking arbitrary $\hat{x} \in C_i$, we get $(p_i, \hat{x} + r\bar{x}) \leqslant 0$, so $\hat{x} + r\bar{x} \in C_i$ for all $r > 0$, but $(\lambda_i, \hat{x} + r\bar{x}) = (\lambda_i, \hat{x}) + r(\lambda_i, \bar{x}) \to -\infty$ as $r \to +\infty$, which contradicts (8.1). Therefore, our claim is proved, and hence, all $\lambda_i = -\alpha_i p_i$ for some $\alpha_i \geqslant 0$.

Similar (and even more simple) reasons show that $A\bar{x} = 0$ implies $\mu\,\bar{x} = 0$, which yields $\mu = y^*A$ for some $y^* \in Y^*$, q.e.d.

($\Longleftarrow$) Suppose (5.8)–(5.9) hold, but $\exists\,\hat{x}$ satisfying (5.5). Multiplying the last relations by α_i and y^*, and summing them up, we get

$$\sum_i \alpha_i\big((p_i, \hat{x}) + \xi_i\big) \;+\; y^*(A\hat{x} + \eta) \;< 0,$$

because all $\alpha_i \geqslant 0$ and $\sum a_i > 0$. Taking into account (5.8), we obtain $\sum \alpha_i \xi_i + (y^*, \eta) < 0$, a contradiction with (5.9), q.e.d. □

REFERENCES

[1] A.Ya. Dubovitskii, A.A. Milyutin. Extremum problems in the presence of restrictions, USSR Comput. Math. and Math. Phys. 5 (1965) 1–80.

[2] I.V. Girsanov, Lectures on Mathematical Theory of Extremum Problems, Lecture Notes in Econ. and Math. Systems, 1972.

[3] A.D. Ioffe, V.M. Tikhomirov, Theory of Extremal Problems, North-Holland, Amsterdam, 1979.

[4] V.M. Alekseev, V.M. Tikhomirov, and S.V. Fomin, Optimal Control, (Transl. from the Russian), Consultants Bureau, New York, 1987.

[5] A. Dmitruk, N. Osmolovskii, A general Lagrange multipliers theorem, In: Constructive Nonsmooth Analysis and Related Topics (CNSA-2017), 2017, doi: 10.1109/CNSA.2017.7973951.

[6] A.V. Dmitruk, N.P. Osmolovskii, A general Lagrange multipliers theorem and related questions, control systems and math. methods in economics, G. Feichtinger et al. (eds.), Lecture Notes in Econ. and Math. Systems, vol. 687, pp. 165–194, Springer 2018.

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

[8] B.S. Goh, Necessary conditions for singular extremals involving multiple control variables, SIAM J. Control, 4 (1966) 716–731.

[9] R. Gabasov, F.M. Kirillova, Singular Optimal Controls, Nauka, Moscow, 1973.

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

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

[12] A.J. Krener, The high order maximal principle and its application to singular extremals, SIAM J. Control, 15 (1977), 256–293.

[13] A.A. Agrachev, R.V. Gamkrelidze, A second order optimality principle for a time-optimal problem, Mathematics of the USSR–Sbornik, 29 (1976), 547–576.

[14] A.V. Sarychev, High-order necessary conditions of optimality for nonlinear control systems, Systems & Control Lett. 16 (1991) 369–378.

[15] M.I. Zelikin, Conditions for optimality of singular trajectories in the problem of minimizing a curvilinear integral Soviet Math. Doklady, 26 (1982) 631–634.

[16] F. Lamnabhi-Lagarrigue, G. Stefani, Singular optimal control problems: on the necessary conditions of optimality, SIAM J. Control & Optim. 28 (1990) 823–840.

[17] J.L. Speyer, D.H. Jacobson, Necessary and sufficient conditions of optimality for singular control problems. A transformation approach, J. Math. Anal. Appl. 33 (1971) 163–186.

[18] V.I. Gurman, Singular Optimal Control Problems, Moscow, Nauka, 1977.

[19] V.A. Dykhta, Conditions of a local minimum for singular regimes in systems with linear control, Automatics and Remote Control, 42 (1982) 1583-1587.

[20] 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, 33 (1978) 97–168.

[21] A.V. Dmitruk, Quadratic order conditions of a weak minimum for singular regimes in optimal control problems, Sov. Math., Dokl. 18 (1977) 418–422.

[22] A.V. Dmitruk, Quadratic conditions for a weak minimum for singular regimes in optimal control problems, Proc. of VIII Winter School on Math. Programming and Related Topics, pp. 102–119, CEMI RAN, 1976.

[23] A.D. Ioffe, Necessary and sufficient conditions for a local minimum: 3, SIAM J. Control Optim. 17 (1979) 266–288.

[24] H. Maurer, First and second order sufficient optimality conditions in mathematical programming and optimal control, Math. Program. Stud. 14 (1981) 163–177.

[25] K. Malanowski, Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems, Adv. Math. Sci. Appl. 2 (1993) 397–443.

[26] A.A. Milyutin, Quadratic conditions of an extremum in smooth problems with a finite-dimensional image, In: ”Metody teorii ekstremal’nyh zadach v ekonomike (Methods of the theory of extremal problems in economics)”, V.L. Levin (ed.), pp. 138–177, Nauka, Moscow, 1981. (in Russian).

[27] A.A. Milyutin, N.P. Osmolovskii, Calculus of Variations and Optimal Control, American Math. Soc., Providence, Rhode Island, vol. 180, 1998.

[28] J.-P. Aubin, I. Ekeland, Applied Nonlinear Analysis, Wiley and Sons, 1984.

[29] A.V. Arutyunov, Second order necessary conditions in optimal control problems, Doklady Russian Acad. Sci., Mathematics, 2000, v. 371, no. 1.

[30] A.V. Dmitruk, Jacobi type conditions for the problem of Bolza with inequalities, Math. Notes of the Acad. Sci. USSR, 35 (1984), 427–435.

[31] A.V. Dmitruk, Jacobi type conditions for singular extremals, Control and Cybernetics, 37 (2008) 285–306.

[32] A.V. Dmitruk, K.K. Shishov, Analysis of a quadratic functional with a partly singular Legendre condition, Moscow University Comput. Math. and Cybernetics, 34 (2010) 16–25.

[33] A.V. Dmitruk, N.A. Manuilovich, Minimization of degenerate integral quadratic functionals. Proc. Steklov Institute Math. 315 (2021) 98-117.