17/2008

Raport Badawczy

Research Report

RB/3/2008

Second-order sufficient optimality conditions for control problems with linearly independent gradients of control constraints

Nikolai P. Osmolovskii

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Pracowni zgłaszający pracę:
Doc. dr hab. inż. Antoni Żochowski

Warszawa 2008

Second-order sufficient optimality conditions for control problems with linearly independent gradients of control constraints

Nikolai P. Osmolovskii*

September 3, 2008

Abstract

We consider optimal control problems with initial-final state equality and inequality constraints and control inequality constraints given by smooth functions satisfying the hypothesis of linear independence of gradients of active control constraints. For such problems, we derive second-order sufficient conditions of a bounded strong minimum with quadratic growth of 'violation function' [3].

1 Pontryagin and bounded strong minima. First order necessary conditions

Consider the following optimal control problem on a fixed interval $[0, T]$:

$$\dot{y}(t) = f(u(t), y(t)) \quad \text{for a.a. } t \in [0, T], \tag{1}$$

$$u(t) \in U, \quad \text{for a.a. } t \in [0, T], \tag{2}$$

$$\phi_i(y(0), y(T)) \leq 0, \ i = 1, \dots, r_1, \tag{3}$$

$$\phi_i(y(0), y(T)) = 0, \ i = r_1 + 1, \dots, r, \tag{4}$$

$$J(w) := \phi_0(y(0), y(T)) \to \min, \tag{5}$$

where $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ and $\phi_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $i = 0, \dots, r$ are twice continuously differentiable (C^2) mappings, U is a closed subset of $\mathbb{R}^m$. Denote by $\mathcal{U} := L^\infty(0, T; \mathbb{R}^m)$ and $\mathcal{Y} := W^{1,1}(0, T; \mathbb{R}^n)$ the control and state space. We consider problem (1)-(5) in the space $\mathcal{W} := \mathcal{U} \times \mathcal{Y}$, and we refer to this problem as problem (P). Define the norm of element $w = (u, y) \in \mathcal{W}$ by $\|w\|_{\mathcal{W}} := \|u\|_\infty + \|y\|_{1,1} = \text{ess sup}_{[0,T]}|u(t)| + |y(0)| + \int_0^T |\dot{y}(t)| \, \mathrm{d}t$. Elements of $\mathcal{W}$ satisfying (1)-(4) are said to be feasible. The set of feasible points

†Systems Research Institute, ul. Newelska 6, 01-447 Warszawa, Poland; Politechnika Radomska and Akademia Podlaska, Poland
Email: nikolai@osmolovskii.msk.ru

is denoted by $F(P)$. We shall use abbreviations $y(0) = y_0$, $y(T) = y_T$, $(y_0, y_T) = \eta$.

It is well known that any control problem with a cost functional in the integral form $J = \int_0^T F(u,y)\,dt$ can be represented in the endpoint form by introducing a new state variable z defined by the state equation $\dot{z} = F(u,y)$, $z(0) = 0$. This yields the cost functional $J = z(T)$. The new variable z is called unessential component in the augmented problem. The general definition of an unessential component[5] is as follows. The state variable y_i, i.e., the i-th component of the state vector y is called *unessential* if the function f does not depend on y_i and if the functions ϕ_j, $j = 0, 1, \dots, r$ are affine in $y_{i0} := y_i(0)$ and $y_{iT} := y_i(T)$. Let $\underline{y}$ denote the vector of all essential components of the state vector y.

Let us define two concepts of minimum. We say that $w^0 = (u^0, y^0) \in F(P)$ is a *bounded strong minimum* if $J(w^0) \leq J(w^k)$ for large enough k for any sequence $w^k \in F(P)$, bounded in $\mathcal{W}$, such that $\underline{y}^k \to \underline{y}^0$ uniformly and $y^k(0) \to y^0(0)$. We say that $w^0 \in F(P)$ is a *Pontryagin minimum* if $J(w^0) \leq J(w^k)$ for large enough k for any sequence $w^k \in F(P)$, bounded in $\mathcal{W}$, such that $y^k \to y^0$ uniformly and $\|u^k - u^0\|_1 \to 0$, where $\|u\|_1 = \int_0^T |u(t)\,dt$.

Equivalently, w^0 is a bounded strong minimum iff for any $M > 0$, there exist $\varepsilon > 0$ such that if $w \in F(P)$ is such that $\|u\|_\infty \leq M$, $\|\underline{y} - \underline{y}^0\|_\infty \leq \varepsilon$, and $|y(0) - y^0(0)| < \varepsilon$, we have that $J(w^0) \leq J(w)$. A point w^0 is a Pontryagin minimum iff for any $M > 0$, there exist $\varepsilon > 0$ such that if $w \in F(P)$ is such that $\|u\|_\infty \leq M$, $\|y - y^0\|_\infty \leq \varepsilon$, and $\|u - u^0\|_1 < \varepsilon$, we have that $J(w^0) \leq J(w)$. Obviously, a bounded strong minimum implies a Pontryagin minimum.

Let us recall the formulation of Pontryagin's principle at the point $w \in F(P)$. Denote by $\mathbb{R}^{q*}$ the dual to $\mathbb{R}^q$ identified with the set of q dimensional raw vectors. Set

$$\varphi^\mu(y_0, y_T) = \varphi(y_0, y_T, \mu) := \sum_{i=0}^{r} \mu_i \phi_i(y_0, y_T), \tag{6}$$

where $y_0 = y(0)$, $y_T = y(T)$, $\mu = (\mu_0, \dots \mu_r) \in \mathbb{R}^{(r+1)*}$. Consider the Hamitonian function $H : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n*}$ defined by

$$H(u, y, p) = pf(u, y). \tag{7}$$

We call *costate associated with* $\mu \in \mathbb{R}^{(r+1)*}$ the solution $p = p^\mu$ (whenever it exists) of

$$\begin{aligned} -\dot{p}(t) &= H_y(u(t), y(t), p(t)), \quad \text{a.a. } t \in [0,T]; \\ p(0) &= -\varphi^\mu_{y_0}(y(0), y(T)); \quad p(T) = \varphi^\mu_{y_T}(y(0), y(T)). \end{aligned} \tag{8}$$

Definition 1.1. We say that $w = (u, y) \in F(P)$ satisfies Pontryaguin's principle if there exist a nonzero $\mu \in \mathbb{R}^{(r+1)*}$ and $p \in W^{1,\infty}(0, T, \mathbb{R}^{n*})$ such that

(8) holds and

$$\mu_i \geq 0, \quad i = 0, \ldots, r_1, \quad \mu_i \phi_i(y(0), y(T)) = 0, \quad i = 1, \ldots, r_1, \tag{9}$$

$$H(u(t), y(t), p(t)) \leq H(v, y(t), p(t)), \quad \text{for all } v \in U, \text{ a.a. } t \in (0,T). \tag{10}$$

The following theorem holds [3],[4],[5]:

Theorem 1.2. *A Pontryagin minimum satisfies Pontryagin's principle.*

In the sequel, we assume that the set U is given in the form $U = \{u \in \mathbb{R}^m \mid g(u) \leq 0\}$, where $g : \mathbb{R}^m \to \mathbb{R}^q$ is C^2 mapping. In other words, the control constraints are defined by

$$g_j(u(t)) \leq 0, \quad \text{for a.a.} \quad t \in [0,T], \quad j = 1, \ldots, q. \tag{11}$$

We assume that the following *qualification hypothesis of linear independence* holds: the gradients $g_i'(u)$, $i \in I_g(u)$ are linearly independent at each point $u \in \mathbb{R}^m$ such that $g(u) \leq 0$, where $I_g(u) = \{i \in \{1, \ldots, q\} \mid g_i(u) = 0\}$ is the set of active indices.

Let us recall a first order necessary condition of a weak minimum, which is a local minimum in $\mathcal{W}$. To this end, define the augmented Pontryagin function $\overline{H} : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n*} \times \mathbb{R}^{q*} \to \mathbb{R}$ by

$$\overline{H}(u, y, p, a) = H(u, y, p) + ag(u). \tag{12}$$

For $w = (u, y) \in F(P)$, denote by Λ_0 the set of all tuples $\lambda = (\mu, p, a) \in \mathbb{R}^{(r+1)*} \times W^{1,\infty}(0,T; \mathbb{R}^{n*}) \times L^\infty(0,T; \mathbb{R}^{q*})$ of Lagrange multipliers such that the following relations hold

$$\begin{array}{l} \mu_i \geq 0,\ i = 0, \ldots, r_1,\ \mu_i \phi_i(y(0), y(T)) = 0,\ i = 1, \ldots, r_1, \\ a(t) \geq 0,\ a(t) g(u(t)) = 0,\ \text{a.a. } t \in (0,T), \\ -\dot{p}(t) = H_y(w(t), p(t)),\ \text{a.a. } t \in (0,T), \\ p(0) = -\varphi^\mu_{y_0}(\eta),\ p(T) = \varphi^\mu_{y_T}(\eta), \\ \overline{H}_u(w(t), p(t), a(t)) = 0,\ \text{a.a. } t \in (0,T);\ |\mu| = 1\}. \end{array} \tag{13}$$

The following result is well-known [1].

Theorem 1.3. *Let w be a weak minimum. Then the set Λ_0 is nonempty and bounded. Moreover, the projector $(p, \lambda, \mu) \to \mu$ is injective on Λ_0.*

Denote by M_0 the set of all $\lambda = (\mu, p, a) \in \Lambda_0$ such that inequality (10) of Pontryagin's principle is satisfied. Obviously, $M_0 \subset \Lambda_0$, and the condition $M_0 \neq \emptyset$ is equivalent to Pontryagin's principle.

2 Growth condition of order γ

Let us fix a pair $w = (u, y) \in F(P)$. By $\delta w = (\delta u, \delta y)$ we denote a *variation*, i.e., an arbitrary element of the space $\mathcal{W}$, and the notation $\{\delta w_k\}$ stands for an arbitrary sequence of variations in $\mathcal{W}$. For any $\delta w \in \mathcal{W}$ we set

$$\delta f = f(u(t) + \delta u(t), y(t) + \delta y(t)) - f(u(t), y(t)) = f(w(t) + \delta w(t)) - f(w(t)),$$

i.e., δf is the increment of the function f (at the point $w(t)$) which corresponds to the variation $\delta w(t)$. Similarly, we set

$$\delta\phi_0 = \phi_0(y(0) + \delta y(0), y(T) + \delta y(T)) - \phi_0(y(0), y(T)) = \phi_0(\eta + \delta\eta) - \phi_0(\eta),$$

etc. In order to define a growth condition of the order γ, we must define the so-called 'order function'. A function $\Gamma : \mathbb{R}^m \to \mathbb{R}$ is said to be an *order function* if there exists a number $\varepsilon_\Gamma > 0$ such that (a) $\Gamma(u) = |u|^2$ if $|u| < \varepsilon_\Gamma$; (b) $\Gamma(u) > 0$ if $|u| \geq \varepsilon_\Gamma$; (c) $\Gamma(u)$ is Lipschitz continuous on each compact set $\mathcal{C} \subset \mathbb{R}^m$. Obviously, the function $\Gamma(u) = |u|^2$ is an order function. For an arbitrary order function $\Gamma(u)$, we set

$$\gamma(\delta w) = \|\delta y\|_\infty^2 + \int_0^T \Gamma(\delta u(t))\,\mathrm{d}t. \tag{14}$$

We call the functional $\gamma : \mathcal{W} \to \mathbb{R}$ the *higher order* [3].

Next, let us define the *violation function* [3]

$$\sigma(\delta w) = (\delta J)_+ + \sum_{i=1}^{r_1} \phi_i(\eta + \delta\eta)_+ + \sum_{i=r_1+1}^{r} |\phi_i(\eta + \delta\eta)| + \|\delta\dot{y} - \delta f\|_1, \tag{15}$$

where $\eta = (y_0, y_T) = (y(0), y(T))$, $\delta\eta = (\delta y_0, \delta y_T) = (\delta y(0), \delta y(T))$, $(\delta J)_+ = (\phi_0(\eta + \delta\eta) - \phi_0(\eta))_+$, $\alpha_+ = \max\{\alpha, 0\}$.

We say that $\{\delta w_k\}$ is a *bounded strong sequence* if $\limsup_k \|\delta u_k\|_\infty < \infty$, $|\delta y_k(0)| + \|\underline{\delta y_k}\|_\infty \to 0$ $(k \to \infty)$. Denote by S the set of all bounded strong sequences satisfying (a) $u(t) + \delta u_k(t) \in U$ for a.a. $t \in [0, T]$ for all k, and (b) $\sigma(\delta w_k) \to 0$ $(k \to \infty)$.

We say that a *bounded strong γ-sufficiency* holds (at the point w) [3] if there exists a constant $C > 0$ such that for any sequence $\{\delta w_k\} \in S$ we have $\sigma(\delta w_k) \geq C\gamma(\delta w_k)$ for all sufficiently large k.

Equivalently, a *bounded strong γ-sufficiency* holds iff there exists $C > 0$ such that for any $M > 0$ there exists $\varepsilon > 0$ such that the conditions

$$\begin{aligned} &u(t) + \delta u(t) \in U \text{ for a.a. } t \in [0, T], \quad \|\delta u\|_\infty < M, \\ &\sigma(\delta w) < \varepsilon, \quad |\delta y(0)| < \varepsilon, \quad \|\underline{\delta y}\|_\infty < \varepsilon \end{aligned}$$

imply the inequality $\sigma(\delta w) \geq C\gamma(\delta w)$.

We say that a *bounded strong γ-growth condition holds for the cost function J* if there exists $C > 0$ such that for any $M > 0$ there exists $\varepsilon > 0$ such that the conditions $w + \delta w \in F(P)$, $\|\delta u\|_\infty < M$, $\|\underline{\delta y}\|_\infty < \varepsilon$, $|\delta y(0)| < \varepsilon$, $\delta J < \varepsilon$ imply the inequality $\delta J \geq C\gamma(\delta w)$. Obviously, a bounded strong γ-sufficiency implies a bounded strong γ-growth condition for the cost function, and the latter implies a strict bounded strong minimum.

Set

$$C_\gamma(\sigma, S) := \inf_{\{\delta w_k\} \in S} \left(\liminf_k \frac{\sigma(\delta w_k)}{\gamma(\delta w_k)} \right),$$

where the lower bound is taken over the set of all sequences from S that do not vanish. The following proposition easily follows from definitions.

Proposition 2.1. *The inequality $C_\gamma(\sigma, S) > 0$ is equivalent to the bounded strong γ-sufficiency.*

Our goal is to obtain conditions which guarantee this inequality. To this end, we will estimate $C_\gamma(\sigma, S)$ from below.

Let $\lambda = (\mu, p, a) \in M_0$. We say that the function $H(v, y(t), p(t))$ satisfies a *growth condition of the order* Γ if there exists $C > 0$ such that for a.a. $t \in [0, T]$ we have

$$H(v, y(t), p(t)) - H(u(t), y(t), p(t)) \geq C\Gamma(v - u(t)) \quad \text{for all} \quad v \in U. \tag{16}$$

For any $C > 0$, denote by $M(C\Gamma)$ the set of all $\lambda \in M_0$ such that the condition (16) is satisfied for a.a. $t \in [0, T]$. One may show that if a bounded strong γ-sufficiency holds, then there exists $C > 0$ such that the set $M(C\Gamma)$ is nonempty.

3 Second order sufficient conditions

A direction (variation) $\delta w = (\delta u, \delta y) \in \mathcal{W}$ is said to be *critical* [1] at the point w if the following relations hold

$$\begin{aligned} &\phi_i'(\eta)\delta\eta \leq 0, \ i \in I_\phi(\eta) \cup \{0\}; \ \phi_j'(\eta)\delta\eta = 0, \ j = r_1 + 1, \ldots, r, \\ &\delta\dot{y} = f'(w)\delta w, \ (g_j'(u)\delta u)\chi_{\{g_j(u)=0\}} \leq 0, \ j = 1, \ldots, q, \end{aligned} \tag{17}$$

where $I_\phi(\eta) = \{i \in \{1, \ldots, r_1\} \mid \phi_i(y(0), y(T)) = 0\}$ is the set of active indices, $\chi_{\{g_j(u)=0\}}$ is the characteristic function of the set $\{t \in [0, T] \mid g_j(u(t)) = 0\}$, $j = 1, \ldots, q$, $\eta = (y(0), y(T))$, and $\delta\eta = (\delta y(0), \delta y(T))$. Denote by $\mathcal{K}$ the set of all critical directions $\delta w \in \mathcal{W}$ at the point w. Obviously, $\mathcal{K}$ is a convex cone in $\mathcal{W}$. We call it the *critical cone*.

For any $\lambda = (\mu, p, a) \in \Lambda_0$, let us define a quadratic form at the point w by relation

$$\Omega(\delta w, \lambda) := \frac{1}{2}\langle \varphi_{\eta\eta}(\eta, \mu)\delta\eta, \delta\eta\rangle + \frac{1}{2}\int_0^T \langle \overline{H}_{ww}(w, p, a)\delta w, \delta w\rangle \, dt. \tag{18}$$

For any $C \geq 0$, set

$$\Omega_{M(C\Gamma)}(\delta w) = \max_{\lambda \in M(C\Gamma)} \Omega(\delta w, \lambda). \tag{19}$$

Theorem 3.1. *For a feasible point $w = (u, y)$, assume that there exist an order function Γ and a number $C > 0$ such that the set $M(C\Gamma)$ is nonempty and there exists $C_{\mathcal{K}} > 0$ such that*

$$\Omega_{M(C\Gamma)}(\delta w) \geq C_{\mathcal{K}}\Big(|\delta y(0)|^2 + \int_0^T |\delta u|^2\, dt\Big) \quad \textit{for all} \quad \delta w \in \mathcal{K}. \tag{20}$$

Then, for the corresponding higher order γ (14), a bounded strong γ-sufficiency holds at the same point.

Remark 3.2. One may prove that the sufficient optimality condition given by Theorem 3.1 is a natural strengthening of the following necessary condition of Pontryaguin minimum: the set M_0 is nonempty (i.e., Pontryaguin's principle holds) and $\max_{\lambda \in M_0} \Omega(\delta w, \lambda) \geq 0$ for all $\delta w \in \mathcal{K}$.

Below, we will give a proof of Theorem 3.1.

4 Passage to Pontryagin's sequences, the basic constant

In what follows, we assume that, for given order function Γ and a number $C > 0$, the set $M(C\Gamma)$ is nonempty. For any $\lambda = (\mu, p, a) \in \Lambda_0$, we set

$$\Psi(\delta w, \lambda) := \delta\varphi^{\mu} - \int_0^T p(\delta\dot{y} - \delta f)\,\mathrm{d}t = \delta\varphi^{\mu} - \int_0^T (p\delta\dot{y} - \delta H)\,\mathrm{d}t, \tag{21}$$

where $\delta\varphi^{\mu} = \varphi^{\mu}(\eta + \delta\eta) - \varphi^{\mu}(\eta)$, $\delta H = p\delta f$. Since Λ_0 is a bounded set, it is easy to show that there exists $k_0 > 0$ such that

$$\max_{\lambda \in \Lambda_0} \Psi(\delta w, \lambda) \leq k_0 \sigma(\delta w). \tag{22}$$

Moreover, in virtue of (8), we have

$$\begin{aligned} &\textstyle\int_0^T p\delta\dot{y}\,\mathrm{d}t = p\delta y\,|_0^T - \int_0^T \dot{p}\delta y\,\mathrm{d}t \\ = \;& \varphi^{\mu}_{y_0}(\eta)\delta y(0) + \varphi^{\mu}_{y_T}(\eta)\delta y(T) + \textstyle\int_0^T H_y(w,p)\delta y\,\mathrm{d}t \end{aligned} \tag{23}$$

for any $\lambda \in \Lambda_0$. Consequently,

$$\Psi(\delta w, \lambda) = \delta\varphi^{\mu} - \varphi^{\mu}_{\eta}(\eta)\delta\eta + \int_0^T (\delta H - H_y(w,p)\delta y)\,\mathrm{d}t, \forall\, \lambda \in \Lambda_0. \tag{24}$$

In the sequel, we shall omit k in notation of sequences. Under the assumption that $M(C\Gamma) \neq \emptyset$, $C > 0$, the following lemma holds.

Lemma 4.1. *If $\{\delta w\} \in S$ then $\|\delta u\|_1 \to 0$, $\|\delta y\|_\infty \to 0$, and hence $\gamma(\delta w) \to 0$.*

The proof of this lemma consists of two propositions.

Proposition 4.2. *Let $\{\delta w\} \in S$ and $(\mu, p, a) \in \Lambda_0$. Then $\left(\int_0^T \delta H \, \mathrm{d}t\right)_+ \to 0$.*

Proof. According to (22) and (24), $\delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T (\delta H - H_y(w,p)\delta y)\,\mathrm{d}t \leq k_0\sigma(\delta w)$. Since $\|\underline{\delta y}\|_\infty \to 0$, the condition $\sigma(\delta w) \to 0$ implies $\left(\int_0^T \delta H \, \mathrm{d}t\right)_+ \to 0$. □

Proposition 4.3. *If $\{\delta w\} \in S$, then $\int_0^T |\delta u|^2 \,\mathrm{d}t \to 0$.*

Proof. Let $\{\delta w\} \in S$ and $(\mu, p, a) \in M(C\Gamma)$ $(C > 0)$. We obviously have

$$\delta H \quad := \quad H(w + \delta w, p) - H(w, p) = \bar{\delta}_y H + \delta_u H,$$

where $\bar{\delta}_y H = H(u + \delta u, y + \delta y, p) - H(u + \delta u, y, p)$, $\delta_u H = H(u + \delta u, y, p) - H(w, p)$. The conditions $\|\underline{\delta y}\|_\infty \to 0$ and $\limsup \|\delta u\|_\infty < \infty$ imply that $\|\bar{\delta}_y H\|_\infty \to 0$. Therefore, the condition $\left(\int_0^T \delta H \,\mathrm{d}t\right)_+ \to 0$ (which holds by Proposition 4.2) implies $\left(\int_0^T \delta_u H \,\mathrm{d}t\right)_+ \to 0$. But $\delta_u H \geq C\Gamma(\delta u)$, since $(\mu, p, a) \in M(C\Gamma)$ and $u + \delta u \in U$. Consequently, $\int_0^T \Gamma(\delta u)\,\mathrm{d}t \to 0$ which easily implies that $\int_0^T |\delta u|^2 \,\mathrm{d}t \to 0$. □

Proof of Lemma 4.1 Let $\{\delta w\} \in S$. Then by Cauchy-Schwartz inequality and Proposition 4.3 we have

$$\|\delta u\|_1 = \int_0^T |\delta u| \,\mathrm{d}t \leq \sqrt{T} \left(\int_0^T |\delta u|^2 \,\mathrm{d}t \right)^{\frac{1}{2}} \to 0.$$

Moreover, from conditions $\|\delta u\|_1 \to 0$, $|\delta y(0)| + \|\underline{\delta y}\|_\infty \to 0$, and $\|\delta\dot{y} - \delta f\|_1 \to 0$ we easily deduce that $\|\delta y\|_{1,1} \to 0$ and hence $\|\delta y\|_\infty \to 0$. Consequently, $\gamma(\delta w) \to 0$. □

We shall say that $\{\delta w\}$ is a *Pontryagin's sequence* if the following conditions are satisfied: $\limsup \|\delta u\|_\infty < \infty$, $\|\delta u\|_1 \to 0$, and $\|\delta y\|_\infty \to 0$. Denote by Π the set of all Pontryagin's sequences. For any $\{\delta w\} \in \Pi$ we obviously have: $\sigma(\delta w) \to 0$. Thus, under the assumption that $M(C\Gamma) \neq \emptyset$, $C > 0$, we have proved that

$$S = \{\{\delta w\} \in \Pi \mid u(t) + \delta u(t) \in U\} \tag{25}$$

(in this formula, the condition $u(t) + \delta u(t) \in U$ is assumed to be satisfied for a.a. $t \in [0, T]$ and for all members of the sequence$\{\delta w\}$). Set

$$\Pi_{\sigma\gamma} = \{\{\delta w\} \in \Pi \mid u(t) + \delta u(t) \in U,\ \sigma(\delta w) \leq O(\gamma(\delta w))\}. \tag{26}$$

From equality (25) and definition (26) we easily deduce that

$$C_\gamma(\sigma, S) = C_\gamma(\sigma, \Pi_{\sigma\gamma}). \tag{27}$$

Set

$$\Psi_{\Lambda_0}(\delta w) = \max_{\lambda\in\Lambda_0} \Psi(\delta w, \lambda). \tag{28}$$

Then according to (22) we have $\Psi_{\Lambda_0}(\delta w) \le k_0\sigma(\delta w)$. Consequently,

$$C_\gamma(\sigma, \Pi_{\sigma\gamma}) \ge k_0^{-1} C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}), \tag{29}$$

where

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) := \inf_{\{\delta w\}\in\Pi_{\sigma\gamma}} \left(\liminf \frac{\Psi_{\Lambda_0}(\delta w)}{\gamma(\delta w)}\right) \tag{30}$$

(the lower bound in this formula is taken over the set of sequences from $\Pi_{\sigma\gamma}$ that do not vanish). We call $C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma})$ the *basic constant on the set of Pontryagin's sequences.* From (27) and (29) we obtain the following inequality

$$C_\gamma(\sigma, S) \ge k_0^{-1} C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}). \tag{31}$$

In the sequel, we will estimate $C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma})$ from below.

5 Extension of the set $\Pi_{\sigma\gamma}$

Let $\lambda = (\mu, p, a) \in \Lambda_0$ and $\{\delta w\} \in \Pi$ satisfies the condition $u(t) + \delta u(t) \in U$ a.e. on $[0, T]$ for all members of the sequence. Then relation (24) combined with the equalities $a\delta g + a\delta g_- = 0$ (where $\alpha_- = \max\{-\alpha, 0\} \ge 0$ and $g_- = (g_{1-}, \dots, g_{q-})$ and $\delta\overline{H} = \delta H + a\delta g$ imply

$$\Psi(\delta w, \lambda) = \delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T (\delta\overline{H} - H_y(w, p)\delta y)\, dt + \int_0^T a\delta g_-\, dt. \tag{32}$$

Let $\{\delta w\} \in \Pi_{\sigma\gamma}$ and hence $\sigma(\delta w) \le O(\gamma(\delta w))$. For given sequence, we deduce from (22) and (32) that

$$\max_{\lambda\in\Lambda_0} \left\{\delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T (\delta\overline{H} - H_y(w, p)\delta y)\, dt + \int_0^T a\delta g_-\, dt\right\} \le O(\gamma), \tag{33}$$

where $\gamma = \gamma(\delta w)$. Since $\overline{H}_u(w, p, a) = 0$, $H_y = \overline{H}_y$ and $\limsup \|\delta w\|_\infty < \infty$, the following estimate holds uniformly on Λ_0: $\int_0^T |\delta\overline{H} - H_y(w, p)\delta y|\, dt \le O(\gamma)$. Moreover, $|\delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta| \le O(\gamma)$ uniformly on Λ_0. Therefore, condition (33) implies

$$\max_{\lambda\in\Lambda_0} \int_0^T a\delta g_-\, dt \le O(\gamma). \tag{34}$$

This estimate is satisfied for any $\{\delta w\} \in \Pi_{\sigma\gamma}$. Thus, we get

$$\Pi_{\sigma\gamma} = \Big\{\{\delta w\} \in \Pi \mid g(u+\delta u) \le 0,\ \sigma \le O(\gamma),\ \max_{\lambda\in\Lambda_0} \int_0^T a\delta g_-\,\mathrm{d}t \le O(\gamma)\Big\}. \tag{35}$$

Since we estimate the basic constant from bellow, we may extend the set of sequences $\Pi_{\sigma\gamma}$. Namely, let us define a set of sequences

$$\Pi_{o(\sqrt{\gamma})} = \Big\{\{\delta w\} \in \Pi \mid g(u+\delta u) \le 0,\ \sigma = o(\sqrt{\gamma}),\ \max_{\lambda\in\Lambda_0} \int_0^T a\delta g_-\,\mathrm{d}t \le O(\gamma)\Big\}. \tag{36}$$

Obviously, $\Pi_{\sigma\gamma} \subset \Pi_{o(\sqrt{\gamma})}$, and hence

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) \ge C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt{\gamma})}). \tag{37}$$

In what follows, we will estimate $C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt{\gamma})})$ from below.

6 Passage to the sequences of local variations.

Set $S^{loc} = \{\{\delta w\} \mid \|\delta w\|_\infty \to 0\}$. Sequences from S^{loc} will be called *local.* Note that $S^{loc} \subset \Pi$. Bellow, we will pass to the set of sequences

$$S^{loc}_{o(\sqrt{\gamma})} := \Pi_{o(\sqrt{\gamma})} \cap S^{loc}.$$

Lemma 6.1. *For any $\lambda = (\mu, p, a) \in \Lambda_0$ and for any sequence $\{\delta w\} \in S^{loc}$ satisfying the relation $g(u+\delta u) \le 0$ (for all members of the sequence) the following formula holds:*

$$\Psi(\delta w, \lambda) = \Omega(\delta w, \lambda) + \int_0^T a\delta g_-\,\mathrm{d}t + o(\gamma(\delta w)) \tag{38}$$

uniformly on Λ_0.

Proof. Formula (38) follows from (32) and relations: $H_y = \overline{H}_y$, $\overline{H}_u(w,p,a) = 0$ for all $\lambda \in \Lambda_0$. □

Let $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$, and let $\{\varepsilon\}$ be a sequence of positive numbers converging to zero, i.e., $\varepsilon \to +0$. For members $\delta w = (\delta u, \delta y)$ and ε of the sequences $\{\delta w\}$ and $\{\varepsilon\}$, respectively, which have the same numbers, we set

$$\delta u_\varepsilon(t) = \begin{cases} \delta u(t) & \text{if } \ |\delta u(t)| < \varepsilon, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta u^\varepsilon = \delta u - \delta u_\varepsilon, \quad \delta w_\varepsilon = (\delta u_\varepsilon, \delta y), \quad \delta w^\varepsilon = (\delta u^\varepsilon, 0).$$

Then $\|\delta u_\varepsilon\|_\infty \to 0$ and hence $\{\delta w_\varepsilon\} \in S^{loc}$. Moreover, $\{\delta w\} = \{\delta w_\varepsilon\} + \{\delta w^\varepsilon\}$.

Proposition 6.2. *For the sequences $\{\delta w\}$, $\{\delta w_\varepsilon\}$, and $\{\delta w^\varepsilon\}$, the following formula holds*

$$\delta f = \delta_\varepsilon f + \delta^\varepsilon f + r_f, \tag{39}$$

where

$$\begin{aligned}
&\delta f = f(w+\delta w) - f(w),\ \delta_\varepsilon f = f(w+\delta w_\varepsilon) - f(w),\\
&\delta^\varepsilon f = f(w+\delta w^\varepsilon) - f(w),\ r_f = (\bar\delta_y f - \delta_y f)\chi^\varepsilon,\\
&\delta_y f = f(u, y+\delta y) - f(u,y),\ \bar\delta_y f = f(u+\delta u, y+\delta y) - f(u+\delta u, y),
\end{aligned}$$

and χ^ε is a characteristic function of the set $\mathcal{M}^\varepsilon = \{t \mid \delta u^\varepsilon(t) \neq 0\}$.

Proof. We have

$$\begin{aligned}
\delta f &= \delta f \chi^\varepsilon + \delta f(1-\chi^\varepsilon)\\
&= \Big(f(u+\delta u^\varepsilon, y+\delta y) - f(u+\delta u^\varepsilon, y) + \delta^\varepsilon f\Big)\chi^\varepsilon + \delta_\varepsilon f(1-\chi^\varepsilon)\\
&= \bar\delta_y f\chi^\varepsilon + \delta^\varepsilon f + \delta_\varepsilon f - \delta_\varepsilon f\chi^\varepsilon = \delta^\varepsilon f + \delta_\varepsilon f + (\bar\delta_y f - \delta_y f)\chi^\varepsilon\\
&= \delta_\varepsilon f + \delta^\varepsilon f + r_f.
\end{aligned}$$

□

We continue to work with the sequences $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$, and $\{\varepsilon\}$. For the sequence $\{\delta w\}$, we obviously have: $\|r_f\|_\infty \to 0$. Let us now assume that the sequence $\{\varepsilon\}$ satisfies the following conditions

$$(a)\ \varepsilon \to +0, \quad (b)\ \frac{\|r_f\|_\infty}{\varepsilon^2} \to 0, \quad (c)\ \frac{\sqrt{\gamma(\delta w)}}{\varepsilon^2} \to 0.$$

(Note that $(c) \Rightarrow (b)$.) Then

$$\|r_f\|_1 \le \frac{\|r_f\|_\infty}{\varepsilon^2}\int_{\mathcal{M}^\varepsilon} \varepsilon^2\,\mathrm{d}t = o(\gamma^\varepsilon), \tag{40}$$

where $\gamma^\varepsilon = \gamma(\delta w^\varepsilon)$. Let $\lambda = (\mu, p, a) \in \Lambda_0$. Since $\delta H = p\delta f$, it follows from Proposition 6.2 and estimate (40) that $\int_0^T \delta H\,\mathrm{d}t = \int_0^T \delta_\varepsilon H\,\mathrm{d}t + \int_0^T \delta^\varepsilon H\,\mathrm{d}t + o(\gamma^\varepsilon)$ uniformly on Λ_0, where $\delta_\varepsilon H = H(u+\delta u_\varepsilon, y+\delta y, p) - H(u,y,p)$, $\delta^\varepsilon H = H(u+\delta u^\varepsilon, y, p) - H(u,y,p)$. Consequently, $\Psi(\delta w, \lambda) = \Psi(\delta w_\varepsilon, \lambda) + \int_0^T \delta^\varepsilon H\,\mathrm{d}t + o(\gamma^\varepsilon)$ uniformly on Λ_0, and hence

$$\Psi_{\Lambda_0}(\delta w) = \max_{\lambda\in\Lambda_0}\left\{\Psi(\delta w_\varepsilon, \lambda) + \int_0^T \delta^\varepsilon H\,\mathrm{d}t\right\} + o(\gamma^\varepsilon). \tag{41}$$

Furthermore, the condition $g(u+\delta u) \le 0$ implies

$$g(u+\delta u_\varepsilon) \le 0, \quad g(u+\delta u^\varepsilon) \le 0. \tag{42}$$

Consequently, $\delta^\varepsilon H \geq C\Gamma(\delta u^\varepsilon)$ for any $\lambda \in M(C\Gamma)$, and then $\int_0^T \delta^\varepsilon H\,\mathrm{d}t \geq C\gamma^\varepsilon$ for any $\lambda \in M(C\Gamma)$. Since $M(C\Gamma) \subset \Lambda_0$, we have

$$\max_{\lambda\in\Lambda_0}\left\{\Psi(\delta w_\varepsilon,\lambda) + \int_0^T \delta^\varepsilon H\,\mathrm{d}t\right\} \geq \max_{\lambda\in M(C\Gamma)}\left\{\Psi(\delta w_\varepsilon,\lambda) + \int_0^T \delta^\varepsilon H\,\mathrm{d}t\right\}$$
$$\geq \max_{\lambda\in M(C\Gamma)}\left\{\Psi(\delta w_\varepsilon,\lambda) + C\gamma^\varepsilon\right\} = \Psi_{M(C\Gamma)}(\delta w_\varepsilon) + C\gamma^\varepsilon,$$

where, by definition,

$$\Psi_{M(C\Gamma)}(\delta w) = \max_{\lambda\in M(C\Gamma)} \Psi(\delta w,\lambda). \tag{43}$$

This and formula (41) imply that

$$\Psi_{\Lambda_0}(\delta w) \geq \Psi_{M(C\Gamma)}(\delta w_\varepsilon) + C\gamma^\varepsilon + o(\gamma^\varepsilon). \tag{44}$$

Moreover, we obviously have $\gamma = \gamma_\varepsilon + \gamma^\varepsilon$, where $\gamma = \gamma(\delta w)$, $\gamma_\varepsilon = \gamma(\delta w_\varepsilon)$, $\gamma^\varepsilon = \gamma(\delta w^\varepsilon)$. Consider two possible cases.

(A) $\liminf \gamma_\varepsilon/\gamma = 0$. In this case, we take a subsequence such that $\gamma_\varepsilon = o(\gamma)$, and hence $\gamma^\varepsilon/\gamma \to 1$ on the subsequence. Assume that this condition is satisfied for the sequence $\{\delta w\}$ itself. Then, according to Lemma 6.1, and since $a\delta_\varepsilon g_- \geq 0$, we have

$$\begin{aligned}\Psi_{M(C\Gamma)}(\delta w_\varepsilon) &= \max_{\lambda\in M(C)}\left\{\Omega(\delta w_\varepsilon,\lambda) + \int_0^T a\delta_\varepsilon g_-\,\mathrm{d}t\right\} + o(\gamma_\varepsilon)\\ &\geq \Omega_{M(C\Gamma)}(\delta w_\varepsilon) + o(\gamma_\varepsilon).\end{aligned}$$

Obviously, $|\Omega_{M(C\Gamma)}(\delta w_\varepsilon)| \leq O(\gamma_\varepsilon) = o(\gamma)$. Therefore, inequality (44) implies in this case

$$\liminf \frac{\Psi_{\Lambda_0}(\delta w)}{\gamma(\delta w)} \geq C. \tag{45}$$

(B) $\liminf \gamma_\varepsilon/\gamma > 0$, and hence $\gamma \leq O(\gamma_\varepsilon)$ and $\gamma^\varepsilon \leq O(\gamma_\varepsilon)$. Let us show, in this case, that

$$\{\delta w_\varepsilon\} \in S^{loc}_{o(\sqrt{\gamma})}. \tag{46}$$

Indeed, the sequence $\{\delta w_\varepsilon\} = \{(\delta u_\varepsilon, \delta y)\}$ satisfies the conditions $\|\delta w_\varepsilon\| \to 0$, $\varphi(u+\delta u_\varepsilon) \leq 0$. Furthermore, the following estimate hold

$$\operatorname{meas}\mathcal{M}^\varepsilon = \frac{1}{\varepsilon^2}\int_{\mathcal{M}^\varepsilon} \varepsilon^2\,\mathrm{d}t \leq \frac{1}{\varepsilon^2}\gamma^\varepsilon \cdot O(1) \leq \frac{\sqrt{\gamma}}{\varepsilon^2}\sqrt{\gamma^\varepsilon}\cdot O(1) = o(\sqrt{\gamma^\varepsilon}), \tag{47}$$

since $\sqrt{\gamma}/\varepsilon^2 \to 0$. In virtue of Proposition 6.2 $\delta f = \delta_\varepsilon f + \delta^\varepsilon f + r_f$, where according to (40), $\|r_f\|_1 = o(\gamma^\varepsilon)$, $\|\delta^\varepsilon f\|_1 = O(\operatorname{meas}\mathcal{M}^\varepsilon) = o(\sqrt{\gamma^\varepsilon})$. This and the relations $\|\delta\dot{y} - \delta f\|_1 = o(\sqrt{\gamma})$, $\gamma = O(\gamma_\varepsilon)$ imply $\|\delta\dot{y} - \delta_\varepsilon f\|_1 = o(\sqrt{\gamma_\varepsilon})$. The sequence $\{\delta y\}$ was not changed. Finally, $\max_{\Lambda_0}\int_0^T a\delta_\varepsilon g_-\,\mathrm{d}t \leq \max_{\Lambda_0}\int_0^T a\delta g_-\,\mathrm{d}t \leq O(\gamma) = O(\gamma_\varepsilon)$. Thus, (46) is proved.

It follows from (46) and inequality (44) that

$$\liminf \frac{\Psi_{\Lambda_0}(\delta w)}{\gamma(\delta w)} \geq \liminf \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon) + C\gamma^\varepsilon}{\gamma}$$

$$= \liminf \left(\frac{\gamma_\varepsilon}{\gamma} \cdot \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon)}{\gamma_\varepsilon} + C \cdot \frac{\gamma^\varepsilon}{\gamma} \right)$$

$$\geq \liminf \left(\min \left\{ \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon)}{\gamma_\varepsilon}, \, C \right\} \right) = \min \left\{ \liminf \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon)}{\gamma_\varepsilon}, \, C \right\}$$

$$\geq \min \left\{ \inf_{S^{loc}_{o(\sqrt{\gamma})}} \left(\liminf \frac{\Psi_{M(C\Gamma)}}{\gamma} \right), \, C \right\} = \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})} \right), \, C \right\}.$$

Thus, we have proved that for each sequence $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$ there exists a subsequence such that for this subsequence we have $\liminf \Psi_{\Lambda_0}/\gamma \geq \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})} \right), \, C \right\}$. This implies that

$$\inf_{\Pi_{o(\sqrt{\gamma})}} \liminf \frac{\Psi_{\Lambda_0}}{\gamma} \geq \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})} \right), \, C \right\},$$

that is

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt{\gamma})}) \geq \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})} \right), \, C \right\}. \tag{48}$$

Along with (37) this implies that

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) \geq \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})} \right), \, C \right\}. \tag{49}$$

In what follows, we will estimate $C_\gamma \left(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})} \right)$ from below.

7 Passage to the set of sequences S_1

Our goal consists in passing to the set of sequences of critical variations defined by relations (17). In this section, we will do one more step in this direction.

Proposition 7.1. *Let $\lambda = (\mu, p, a) \in \Lambda_0$. Then the critical cone has the following equivalent representation:*

$$\begin{aligned} &\phi_i'(\eta)\delta\eta \leq 0, \;\; \mu_i \phi_i'(\eta)\delta\eta = 0, \;\; i \in I_\phi(\eta) \cup \{0\}, \\ &\phi_j'(\eta)\delta\eta = 0, \;\; j = r_1 + 1, \ldots, r, \;\; \delta\dot{y} = f'(w)\delta w, \\ &(g_j'(u)\delta u)\chi_{\{g_j(u)=0\}} \leq 0, \;\; a_j g_{ju}(u)\delta u = 0, \;\; j = 1, \ldots, q. \end{aligned} \tag{50}$$

Proof. Indeed, from the definition of the set Λ_0 it easily follows that for any $\lambda = (\mu, p, a) \in \Lambda_0$ we have

$$\sum_{i=0}^{r} \mu_i \phi_i'(\eta)\delta\eta - \int_0^T p(\delta\dot{y} - f'(w)\delta w)\,\mathrm{d}t + \int_0^T a g'(u)\delta u\,\mathrm{d}t = 0 \quad \forall \delta w \in \mathcal{W}.$$

Therefore, relations (17) imply that $\mu_i\phi_i'(\eta)\delta\eta = 0$, $i \in I_\phi(\eta)\cup\{0\}$, $ag'(u)\delta u = 0$. Thus, relations (50) are satisfied. Vice versa, if relations (50) are satisfied, then, obviously, relations (17) are satisfied too. $\square$

Since the convex hull $\operatorname{co}\Lambda_0$ of the set Λ_0 is a finite dimensional convex set, its relative interior $\operatorname{int}\operatorname{co}\Lambda_0$ is nonempty.

Proposition 7.2. *Let $\hat{\lambda} = (\hat{\mu}, \hat{p}, \hat{a}) \in \operatorname{int}\operatorname{co}\Lambda_0$. Then there exists $\hat{C} > 0$ such that, for any $\lambda = (\mu, p, a) \in \operatorname{co}\Lambda_0$, the following inequalities hold*

$$\mu_i \le \hat{C}\hat{\mu}_i,\ i = 0, \ldots, r_1; \quad a_j(t) \le \hat{C}\hat{a}_j(t)\ \text{a.e. on}\ [0,T],\ j = 1, \ldots, q. \tag{51}$$

Proof. Since $\hat{\lambda} = (\hat{\mu}, \hat{p}, \hat{a})$ is an interior point of the set $\operatorname{co}\Lambda_0$, there exists $\varepsilon > 0$ such that for any $\lambda = (\mu, p, a) \in \operatorname{co}\Lambda_0$ we have $\hat{\lambda} \pm \varepsilon(\lambda - \hat{\lambda}) \in \operatorname{co}\Lambda_0$. Condition $\hat{\lambda} - \varepsilon(\lambda - \hat{\lambda}) \in \operatorname{co}\Lambda_0$ implies $\hat{\mu}_i - \varepsilon(\mu_i - \hat{\mu}_i) \ge 0$, $i = 0, \ldots, r_1$, and $\hat{a}_j(t) - \varepsilon(a_j(t) - \hat{a}_j(t)) \ge 0$, $j = 1, \ldots, q$. Consequently,

$$\frac{1+\varepsilon}{\varepsilon}\hat{\mu}_i \ge \mu_i, \quad i = 0, \ldots, r_1; \quad \frac{1+\varepsilon}{\varepsilon}\hat{a}_j(t) \ge a_j(t), \quad j = 1, \ldots, q.$$

Thus, it suffices to set $\hat{C} = (1+\varepsilon)/\varepsilon$. $\square$

Let us fix an element $\hat{\lambda} = (\hat{\mu}, \hat{p}, \hat{a}) \in \operatorname{int}\operatorname{co}\Lambda_0$. It follows from Proposition 7.2, that, in the definition of the set of sequences $S^{loc}_{o(\sqrt{\gamma})}$, the condition $\max_{\Lambda_0}\int_0^T a(\delta g)_-\,\mathrm{d}t \le O(\gamma)$ is equivalent to the condition $\int_0^T \hat{a}(\delta g)_-\,\mathrm{d}t \le O(\gamma)$.

Define a new set of sequences S_1 by the relations

$$\|\delta w\|_\infty \to 0,\ \sigma(\delta w) = o(\sqrt{\gamma(\delta w)}), \tag{52}$$

$$g(u) + g'(u)\delta u \le 0,\ g_j'(u)\delta u\chi_{\{\hat{a}_j \ge \varepsilon\}} = 0,\ j = 1, \ldots, q,\ \varepsilon \to +0, \tag{53}$$

where $\chi_{\{\hat{a}_i \ge \varepsilon\}}(t)$ is the characteristic function of the set $\{t \mid \hat{a}_i(t) \ge \varepsilon\}$. Relations (53) mean that for a given sequence $\{\delta w_k\}$ there exists a sequence $\{\varepsilon_k\}$ such that $\varepsilon_k > 0$, $\varepsilon_k \to 0$, and $g_j'(u(t))\delta u_k(t) = 0$ a.e. on the set $\{t \mid \hat{a}_j(t) \ge \varepsilon_k\}$ for all $j = 1, \ldots, q$ and for all $k = 1, 2, \ldots$. Set

$$\Phi^1_C(\delta w) = \max_{\lambda \in M(C\Gamma)} \left\{ \Omega(\delta w, \lambda) + \int_0^T a(g'(u)\delta u)_-\,\mathrm{d}t \right\}. \tag{54}$$

Let us show that

$$C_\gamma(\Psi_{M(C\Gamma)}, S^{loc}_{o(\sqrt{\gamma})}) \ge C_\gamma(\Phi^1_C, S_1). \tag{55}$$

Take any $\{\delta w\} \in S^{loc}_{o(\sqrt{\gamma})}$. Then

$$\|\delta w\|_\infty \to 0, \quad g(u+\delta u) \le 0, \quad \sigma = o(\sqrt{\gamma}), \quad \int_0^T \hat{a}(\delta g)_-\, \mathrm{d}t \le O(\gamma). \tag{56}$$

Here the relation $\sigma = o(\sqrt{\gamma})$ is equivalent to the set of relations

$$\|\delta\dot{y} - f'(w)\delta w\|_1 = o(\sqrt{\gamma}), \tag{57}$$

$$\phi_i'(\eta)\delta\eta \le o(\sqrt{\gamma}),\ i \in I_\phi(\eta) \cup \{0\},\ |\phi_i'(\eta)\delta\eta| = o(\sqrt{\gamma}),\ i = r_1+1\ldots, r. \tag{58}$$

The last relation in (56) is equivalent to

$$\int_0^T \hat{a}_j(g_j'(u)\delta u)_-\, \mathrm{d}t \le O(\gamma),\ j = 1, \ldots, q. \tag{59}$$

Finally, it follows from the relation $g(u+\delta u) \le 0$ that

$$g_j(u(t)) + g_j'(u(t))\delta u(t) \le k_1|\delta u(t)|^2,\ j = 1, \ldots, q, \tag{60}$$

where $k_1 > 0$ does not depend on the member of the sequence. In virtue of the hypothesis of linear independence of gradients $g_j'(u)$, for any sequence $\varepsilon = \varepsilon(\delta u) \to +0$, there exist $k_2 > 0$ and a sequence $\{\bar{u}\}$ such that

$$|\bar{u}(t)| \le k_2\Big(|\delta u(t)|^2 + \sum_{j=1}^{q}(g_j'(u(t))\delta u(t))_-\chi_{\{\hat{a}_j(t)\ge\varepsilon\}}(t)\Big), \tag{61}$$

$$g_j(u(t)) + g_j'(u(t))(\delta u(t) + \bar{u}(t)) \le 0 \text{ a.e.}, \quad j = 1, \ldots, q, \tag{62}$$

$$g_j'(u(t))(\delta u(t) + \bar{u}(t)) = 0 \ \text{ if } \ \hat{a}_j(t) \ge \varepsilon, \ \ j = 1, \ldots, q, \tag{63}$$

$$g_j'(u(t))\bar{u}(t) = 0 \ \text{ if } \ 0 < \hat{a}_j(t) < \varepsilon,\ j = 1, \ldots, q. \tag{64}$$

Relations (61) imply $\|\bar{u}\|_\infty \le O(\|\delta u\|_\infty) = o(1)$. Since $\int_{\{\hat{a}_j \ge \varepsilon\}}(g_j'(u)\delta u)_-\, \mathrm{d}t \le \frac{1}{\varepsilon}\int_0^T \hat{a}_j(g_j'(u)\delta u)_-\, \mathrm{d}t$, relations (59) and (61) imply $\|\bar{u}\|_1 \le O(\gamma(\delta w))/\varepsilon$. Choose $\varepsilon = \varepsilon(\delta w) \to +0$ such that $\|\delta w\|_\infty/\varepsilon \to 0$. Then $\sqrt{\gamma(\delta w)}/\varepsilon \to 0$, and hence $\|\bar{u}\|_1 = o(\sqrt{\gamma(\delta w)})$. Moreover, since $\|\delta u\|_\infty/\varepsilon \to 0$, we have that

$$\int_0^T |\bar{u}|^2\, \mathrm{d}t \le \|\bar{u}\|_\infty\|\bar{u}\|_1 \le O(\|\delta u\|_\infty)\cdot O(\gamma(\delta w))/\varepsilon = o(\gamma(\delta w)), \tag{65}$$

$$\int_0^T |\delta u|\cdot|\bar{u}|\, \mathrm{d}t \le \|\delta u\|_\infty\|\bar{u}\|_1 \le \|\delta u\|_\infty O(\gamma(\delta w))/\varepsilon = o(\gamma(\delta w)), \tag{66}$$

$$\int_0^T |\delta y|\cdot|\bar{u}|\, \mathrm{d}t \le \|\delta y\|_\infty\|\bar{u}\|_1 \le \|\delta y\|_\infty O(\gamma(\delta w))/\varepsilon = o(\gamma(\delta w)). \tag{67}$$

Set $\{\delta w_1\} = \{(\delta u + \bar{u}, \delta y)\}$. Relations (65)-(67) imply that, for the sequences $\{\delta w_1\}$ and $\{\delta w\}$, we have uniformly on Λ_0

$$\Omega(\delta w_1, \lambda) = \Omega(\delta w, \lambda) + o(\gamma(\delta w)), \quad \gamma(\delta w_1) = \gamma(\delta w) + o(\gamma(\delta w)). \tag{68}$$

Since $\|\bar{u}\|_\infty \to 0$ and $\|\bar{u}\|_1 = o(\sqrt{\gamma(\delta w)})$, we have

$$\|\delta w_1\|_\infty \to 0, \quad \|\delta \dot{y} - f'(w)\delta w_1\|_1 = o(\sqrt{\gamma}). \tag{69}$$

Moreover, the sequence $\{\delta w_1\}$ satisfies relations (58) with $\delta\eta = \delta\eta_1$ (since $\delta y = \delta y_1$) and relations

$$g_j(u(t)) + g_j'(u(t))\delta u_1(t) \le 0 \text{ a.e.}, \ j = 1, \dots, q, \tag{70}$$

$$g_j'(u(t))\delta u_1(t) = 0 \ \text{ if } \ \hat{a}_j(t) \ge \varepsilon, \ \ j = 1, \dots, q, \tag{71}$$

$$g_j'(u(t))\delta u(t) = g_j'(u(t))\delta u_1(t) \ \text{ if } \ 0 < \hat{a}_j(t) < \varepsilon, \ j = 1, \dots, q. \tag{72}$$

Consequently $\{\delta w_1\} \in S_1$. Let us show that, for any $j = 1, \dots, q$, we have

$$\int_0^T a_j(g_j'(u)\delta u_1)_- \,\mathrm{d}t \le \int_0^T a_j(g_j(u+\delta u))_- \,\mathrm{d}t + o(\gamma) \tag{73}$$

uniformly on Λ_0. Indeed, in view of relations (51),(71), (72), and since $\varepsilon \to 0$, we have

$$\begin{aligned} & \int_0^T a_j(g_j'(u)\delta u_1)_- \,\mathrm{d}t = \int_{\hat{a}_j \le \varepsilon} a_j(g_j'(u)\delta u_1)_- \,\mathrm{d}t \\ = \ & \int_{\hat{a}_j \le \varepsilon} a_j(g_j'(u)\delta u)_- \,\mathrm{d}t \le \int_{a_j \le \varepsilon} a_j(g_j(u+\delta u))_- \,\mathrm{d}t + \varepsilon O(\gamma(\delta w)) \\ \le \ & \int_0^T a_j(g_j(u+\delta u))_- \,\mathrm{d}t + o(\gamma(\delta w)). \end{aligned}$$

Relations (38), (43), (54), (68), and (73) imply that $\Phi_C^1(\delta w_1) \le \Psi_{M(C\Gamma)}(\delta w) + o(\gamma(\delta w))$, where $\{\delta w_1\} \in S_1$. Along with the second relation in (68) this implies that

$$\liminf \frac{\Psi_{M(C\Gamma)}(\delta w)}{\gamma(\delta w)} \ge \liminf \frac{\Phi_C^1(\delta w_1)}{\gamma(\delta w_1)} \ge C_\gamma(\Phi_C^1, S_1).$$

Since $\{\delta w\}$ is an arbitrary sequence in $S^{loc}_{o(\sqrt{\gamma})}$, inequality (55) follows.

8 Support of the critical cone

Define $\mathcal{R}_0$ as a set of all variations $\delta w = (\delta u, \delta y) \in \mathcal{W}$ satisfying the relations

$$\delta\dot{y} = f'(w)\delta w, \quad g_j'(u)\delta u \chi_{\{g_j(u)=0\}} \le 0, \quad \hat{a}_j g_j'(u)\delta u = 0, \ j = 1, \dots, q. \tag{74}$$

Obviously, $\mathcal{R}_0$ is a closed convex cone, and by Proposition 7.1, $\mathcal{K} \subset \mathcal{R}_0$. Consider two sets of linear functionals

$$l_i : \delta w \in \mathcal{W} \to \phi_i'(\eta)\delta\eta, \quad i \in I_\phi \cup \{0\}, \tag{75}$$

$$l_i : \delta w \in \mathcal{W} \to \phi_i'(\eta)\delta\eta, \quad i = r_1 + 1, \dots, r, \tag{76}$$

where $\delta w = (\delta u, \delta y)$, $\delta\eta = (\delta y(0), \delta y(T))$. Let Q_0 be the cone generated by functionals (75), and Q_1 be the subspace generated by functionals (76). Set $Q = Q_0 + Q_1$. Then Q is a convex and finitely generated cone. Assume that there exists a linear functional $w_1^* \in Q$ which has an integral representation on the cone $\mathcal{R}_0$:

$$\langle w_1^*, \delta w\rangle = -\int_0^T a^1 g'(u)\delta u \, \mathrm{d}t \quad \forall\, \delta w \in \mathcal{R}_0, \tag{77}$$

where

$$a^1 \in L^\infty(0,T;\mathbb{R}^{q*}), \quad a^1 \geq 0, \quad a^1 g(u) = 0. \tag{78}$$

Obviously, $\langle w_1^*, \delta w\rangle \geq 0$ for all $\delta w \in \mathcal{R}_0$, i.e. $w_1^* \in \mathcal{R}_0^*$. Assume that there exists $\delta w_1 \in \mathcal{R}_0$ such that $\langle w_1^*, \delta w_1\rangle > 0$. Set $\mathcal{R}_1 = \{\delta w \in \mathcal{R}_0 \mid \langle w_1^*, \delta w\rangle = 0\}$. From (74), (77), and (78) it follows that $\mathcal{R}_1 = \{\delta w \in \mathcal{R}_0 \mid a^1 g'(u)\delta u = 0\}$. On the other hand, since $w_1^* \in Q$, we have $\langle w_1^*, \delta w\rangle \leq 0$ for all $\delta w \in \mathcal{K}$. But $\mathcal{K} \subset \mathcal{R}_0$ and $w_1^* \in \mathcal{R}_0^*$, consequently $\langle w_1^*, \delta w\rangle \geq 0$ for all $\delta w \in \mathcal{K}$. Thus, $\langle w_1^*, \delta w\rangle = 0$ for all $\delta w \in \mathcal{K}$, and hence $\mathcal{K} \subset \mathcal{R}_1$.

Similarly, assume that there exists a linear functional $w_2^* \in Q$ which has an integral representation on the cone $\mathcal{R}_1$: $\langle w_2^*, \delta w\rangle = -\int_0^T a^2 g'(u)\delta u\, \mathrm{d}t$ $\forall\, \delta w \in \mathcal{R}_1$, where $a^2 \in L^\infty(0,T,\mathbb{R}^{q*})$, $a^2 \geq 0$, $a^2 g(u) = 0$. Obviously, $\langle w_2^*, \delta w\rangle \geq 0\ \forall\, \delta w \in \mathcal{R}_1$, i.e. $w_2^* \in \mathcal{R}_1^*$. Assume that there exists $\delta w_2 \in \mathcal{R}_1$ such that $\langle w_2^*, \delta w_2\rangle > 0$. Set $\mathcal{R}_2 = \{\delta w \in \mathcal{R}_1 \mid \langle w_2^*, \delta w\rangle = 0\}$. Then $\mathcal{R}_2 = \{\delta w \in \mathcal{R}_1 \mid a^2 g'(u)\delta u = 0\}$. On the other hand, since $w_2^* \in Q$, $w_2^* \in \mathcal{R}_1^*$ and $\mathcal{K} \subset \mathcal{R}_1$,we have $\langle w_2^*, \delta w\rangle = 0\ \forall\, \delta w \in \mathcal{K}$. Hence $\mathcal{K} \subset \mathcal{R}_2$.

Continuing this process, we obtain a set of functionals $w_1^*, w_2^*, \dots, w_s^*$ and a set of cones $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_s$ such that $\mathcal{K} \subset \mathcal{R}_s \subset \ldots \subset \mathcal{R}_1 \subset \mathcal{R}_0$. This process will be finished on some finite step s, because the functionals $w_1^*, w_2^*, \dots, w_s^*$ are linearly independent (this can be easily proved) and the cone Q (containing these functionals) is finite generated. Set $\mathcal{S} = \mathcal{R}_s$, $a^0 = \hat{a}$. Then $\mathcal{S}$ is the set of variations $\delta w \in \mathcal{W}$ such that

$$\begin{aligned} &\delta\dot{y} = f'(w)\delta w,\ (g_j'(u)\delta u)\chi_{\{g_j(u)=0\}} \leq 0,\ j = 1, \dots, q, \\ &a^i g'(u)\delta u = 0,\ i = 0, 1, \dots, s. \end{aligned} \tag{79}$$

Moreover,

$$\mathcal{K} = \left\{\delta w \in \mathcal{S} \mid \phi_i'(\eta)\delta\eta \leq 0,\ i \in I_\phi \cup \{0\},\ \phi_j'(\eta)\delta\eta = 0,\ j = r_1 + 1, \dots, r\right\}.$$

We call $\mathcal{S}$ the *support* of the critical cone $\mathcal{K}$. Let us note an important property of the cone $\mathcal{S}$ which follows from the maximality of the system $w_1^*, \dots, w_s^*$.

Proposition 8.1. *If a linear functional $w^* \in Q$ has an integral representation $\langle w^*, \delta w\rangle = \int_0^T a g'(u)\delta u\, \mathrm{d}t$ on the cone $\mathcal{S}$ such that $a \in L^\infty(0,T;\mathbb{R}^{q*})$, $a \geq 0$, and $ag(u) = 0$, then $\langle w^*, \delta w\rangle = 0$ for all $\delta w \in \mathcal{S}$.*

Let us define the cone $\mathcal{S}_u$ as the set of $\delta u \in \mathcal{U}$ such that

$$(g_j'(u)\delta u)\chi_{\{g_j=0\}} \leq 0,\ j = 1, \ldots, q,\ a^i g'(u)\delta u = 0,\ i = 0, \ldots, s. \tag{80}$$

Then

$$\mathcal{S} = \{\delta w \in \mathcal{W} \mid \delta \dot{y} = f'(w)\delta w,\ \delta u \in \mathcal{S}_u\}. \tag{81}$$

Set $\mathcal{M}_j = \{t \in [0,T] \mid \sum\limits_{i=0}^{s} a_j^i(t) > 0\}$, $\mathcal{N}_j = \{t \in [0,T] \mid g_j(u(t)) = 0\} \setminus \mathcal{M}_j$, $j = 1, \ldots, q$. Then we obviously have

$$\mathcal{S}_u = \left\{\delta u \in \mathcal{U} \mid g_j'(u)\delta u \chi_{\mathcal{N}_j} \leq 0,\ g_j'(u)\delta u \chi_{\mathcal{M}_j} = 0,\ j = 1, \ldots, q\right\}, \tag{82}$$

where $\chi_{\mathcal{N}_j}$ and $\chi_{\mathcal{M}_j}$ are the characteristic functions of the sets $\mathcal{N}_j$ and $\mathcal{M}_j$, respectively, $j = 1, \ldots, q$. In the space $L^1(0,T;\mathbb{R}^{m*})$, consider the cone $\mathcal{C}$ of functions h of the form $h = -ag'(u)$, where $a \in L^1(0,T;\mathbb{R}^{q*})$ is such that

$$a_j(t) = 0 \text{ if } t \notin \mathcal{M}_j \cup \mathcal{N}_j; \quad a_j(t) \geq 0 \text{ if } t \in \mathcal{N}_j,\ j = 1, \ldots, q. \tag{83}$$

Lemma 8.2. $\mathcal{C}^* = \mathcal{S}_u$.

Proof. (a) Let us show that $\mathcal{S}_u \subset \mathcal{C}^*$. Indeed, if $\delta u \in \mathcal{S}_u$ and $h \in \mathcal{C}$, then from (82) and (83) we get $\langle \delta u, h\rangle = \int_0^T h\delta u\, dt = -\int_0^T \sum_{j=1}^q a_j g_j'(u)\delta u\, dt \geq 0$. (b) Let us show that $\mathcal{C}^* \subset \mathcal{S}_u$. Let $\delta u \in \mathcal{C}^*$, i.e., $\delta u \in L^\infty(0,T;\mathbb{R}^m)$ and for any $h \in \mathcal{C}$ we have $\int_0^T h\delta u\, dt = -\sum_{j=1}^q \int_0^T a_j g_j'(u)\delta u\, dt \geq 0$, where $a \in L^1(0,T;\mathbb{R}^{q*})$ satisfies (83). This implies that $g_j'(u)\delta u\chi_{\mathcal{N}_j} \leq 0$ and $g_j'(u)\delta u \chi_{\mathcal{M}_j} = 0$, $j = 1, \ldots, q$, i.e., $\delta u \in \mathcal{S}_u$. □

Set $\Sigma = \{(\delta u, \delta y_0) \in \mathcal{U} \times \mathbb{R}^n \mid \delta u \in \mathcal{S}_u\}$. Then there is one-to-one correspondence between the sets $\mathcal{S}$ and Σ, given by the mapping $(\delta u, \delta y) \in \mathcal{S} \to (\delta u, \delta y(0)) \in \Sigma$. Let $\mu = (\mu_0, \ldots \mu_r) \in \mathbb{R}^{(r+1)*}$, $\varphi^\mu = \sum_{i=0}^r \mu_i \phi_i$. In what follows, it will be convenient to consider the functional $\varphi^\mu_\eta(\eta)\delta\eta$ as a functional defined on Σ. Namely, define a linear functional

$$(\delta u, \delta y_0) \in \mathcal{U} \times \mathbb{R}^n \to \varphi^\mu_{y_0}(\eta)\delta y_0 + \varphi^\mu_{y_T}(\eta)\delta y(T), \tag{84}$$

where δy is a solution to the equation $\delta\dot{y} = f_u(w)\delta u + f_y(w)\delta y$, $\delta y(0) = \delta y_0$. For this functional, we will use the same notation $\varphi^\mu_\eta(\eta)\delta\eta$. This also concerns elements of the set Q.

Let p be the solution to the adjoint equation $-\dot{p} = pf_y(w)$, $p(T) = \varphi^\mu_{y_T}(\eta)$. As is well-known, functional (84) can be written as

$$\varphi^\mu_\eta(\eta)\delta\eta = \Big(\varphi^\mu_{y_0}(\eta) + p(0)\Big)\delta y_0 + \int_0^T pf_u(w)\delta u\, dt.$$

This representation is a 'pure integral' iff $\varphi^\mu_{y_0}(\eta) + p(0) = 0$. Set

$$P = \{(h, z) \in L^1(0,T;\mathbb{R}^{m*}) \times \mathbb{R}^{n*} \mid h \in \mathcal{C},\ z = 0\}.$$

Since $\mathcal{C}^* = \mathcal{S}_u$, we obviously have $P^* = \Sigma$. The following property will play an important role for the existence of corresponding Hoffman's error bounds.

Proposition 8.3. *We have: $P \cap Q \subset L \subset P$, where L is a subspace in the cone P.*

Proof. Let the linear functional $l(\delta u, \delta y_0)$ belong to the intersection $P \cap Q$. Then l has a pure integral representation: $l(\delta u, \delta y_0) = \int_0^T p f_u(w) \delta u \, dt$, where $-\dot{p} = p f_y(w)$, $p(T) = \varphi^{\mu}_{yT}(\eta)$, $p(0) = -\varphi^{\mu}_{y0}(\eta)$, and moreover, there exists $a \in L^{\infty}(0,T;\mathbb{R}^{q*})$ such that $p f_u(w) = -a g'(w)$, $a g(u) = 0$, $a_j \chi_{\mathcal{N}_j} \geq 0$, $j = 1, \ldots, q$. Since $w_1^*, \ldots, w_s^*$ is a maximal system of functionals, we have $a_j \chi_{\mathcal{N}_j} = 0$, $j = 1, \ldots, q$. These relations define a subspace in the cone P. □

The following lemma will play an important role at the end of the proof of Theorem 3.1.

Lemma 8.4. *There exists $N > 0$ such that for any $\delta w \in \mathcal{S}$ there exists $\bar{w} \in \mathcal{S}$ such that the following relations hold*

$$\phi_i'(\eta)(\delta\eta + \bar{\eta}) \leq 0, \; i \in I_\phi \cup \{0\}, \; \phi_j'(\eta)(\delta\eta + \bar{\eta}) = 0, \; j = r_1 + 1, \ldots, r, \quad (85)$$

$$\|\bar{w}\| \leq N \left(\sum_{i \in I_\phi \cup \{0\}} (\phi_i'(\eta)\delta\eta)_+ + \sum_{j=r_1+1}^{r} |\phi_j'(\eta)\delta\eta| \right). \quad (86)$$

Proof. From Proposition 8.3 and abstract Lemma 11.6 (proved in Appendix), we deduce the existence of Hoffman's error bound (86) for system (85) considered on the cone Σ. Consequently, there exists Hoffman's error bound for same system considered on the cone $\mathcal{S}$. □

In Sec. 10, along with this lemma, we will also use two lemmas proved in the next section.

9 Auxiliary assertions

Lemma 9.1. *Let $r \in L^{\infty}_{+}(0,T,\mathbb{R})$, $\mathcal{M} = \{t \mid r(t) > 0\}$, $\{v\}$ be a sequence in $L^2_{+}(0,T;\mathbb{R})$, $\{\alpha\}$ be a sequence in $\mathbb{R}_+$ such that $\alpha \to 0$, $\|v\|_2 \leq \alpha$, and $\int_0^T r v \, dt = o(\alpha)$. Then there exists a sequence $\{B\}$ of the sets $B \subset \mathcal{M}$ such that $\operatorname{meas} B \to 0$ and for the sequence of the sets $\{A\} = \{\mathcal{M} \setminus B\}$ we have $\|v \chi_A\|_{\infty} = o(\alpha)$, where χ_A is a characteristic function of the set A.*

To prove this lemma, we will need the following proposition.

Proposition 9.2. *Let $r \in L^{\infty}_{+}(0,T;\mathbb{R})$, $\mathcal{M} = \{t \mid r(t) > 0\}$, $v \in L^1_{+}(0,T;\mathbb{R})$, $\mathcal{N} = \{t \mid v(t) > 0\}$. Assume that $\mathcal{N} \subset \mathcal{M}$. For $\delta > 0$, set*

$$r_\delta(t) = \begin{cases} r(t) & \text{if } \; r(t) < \delta \\ 0 & \text{otherwise,} \end{cases} \qquad r^\delta(t) = r(t) - r_\delta(t),$$

and similarly, for $\varepsilon > 0$, define $v_\varepsilon(t)$ and $v^\varepsilon(t)$. Set $\mathcal{M}_\delta = \{t \mid r_\delta(t) > 0\}$, $\mathcal{M}^\delta = \{t \mid r^\delta(t) > 0\}$, $\mathcal{N}_\varepsilon = \{t \mid v_\varepsilon(t) > 0\}$, and $\mathcal{N}^\varepsilon = \{t \mid v^\varepsilon(t) > 0\}$. Then the following estimate holds

$$\operatorname{meas} \mathcal{N}^\varepsilon \leq \frac{1}{\delta\varepsilon}\left(\int_0^T rv\,\mathrm{dt} + \delta \int_{\mathcal{M}_\delta} v\,\mathrm{dt}\right).$$

Proof. Since $\mathcal{N}^\varepsilon \subset \mathcal{N} \subset \mathcal{M} = \mathcal{M}_\delta \cup \mathcal{M}^\delta$, we have $\mathcal{N}^\varepsilon = (\mathcal{N}^\varepsilon \cap \mathcal{M}^\delta) \cup (\mathcal{N}^\varepsilon \cap \mathcal{M}_\delta)$. Moreover, $\mathcal{M}_\delta \cap \mathcal{M}^\delta = \emptyset$. Consequently,

$$\delta\varepsilon \operatorname{meas} \mathcal{N}^\varepsilon \leq \int_{\mathcal{N}^\varepsilon \cap \mathcal{M}^\delta} \delta\varepsilon\,\mathrm{dt} + \int_{\mathcal{N}^\varepsilon \cap \mathcal{M}_\delta} \delta\varepsilon\,\mathrm{dt} \leq \int_0^T rv\,\mathrm{dt} + \delta \int_{\mathcal{M}_\delta} v\,\mathrm{dt}.$$

The required estimate follows. □

Proof of Lemma 9.1. Without loss of generality, assume that $\{t \mid v(t) > 0\} \subset \mathcal{M}$ for all members of the sequence $\{v\}$. According to the assumptions of the lemma, we have $\|v\|_2 \leq \alpha$ and $\int_0^T rv\,dt \leq \rho\alpha$, where $\rho \to +0$. Set $\delta = \sqrt{\rho}$ and $\varepsilon = \alpha(\sqrt{\rho} + \sqrt{\operatorname{meas} \mathcal{M}_\delta})^{\frac{1}{2}}$, where $\mathcal{M}_\delta$ was defined in Proposition 9.2. Obviously, $\operatorname{meas} \mathcal{M}_\delta \to 0$ and hence $\|v_\varepsilon\|_\infty \leq \varepsilon = o(\alpha)$. According to Proposition 9.2, we have

$$\begin{aligned} \operatorname{meas}\{t \mid v^\varepsilon(t) > 0\} &\leq \tfrac{1}{\delta\varepsilon}\left(\textstyle\int_0^T rv\,\mathrm{dt} + \delta \int_{\mathcal{M}_\delta} v\,\mathrm{dt}\right) \\ &\leq \tfrac{1}{\delta\varepsilon}\left(\rho\alpha + \delta\sqrt{\operatorname{meas} \mathcal{M}_\delta}\|v\|_2\right) \leq (\sqrt{\rho} + \sqrt{\operatorname{meas} \mathcal{M}_\delta})^{\frac{1}{2}} \to 0. \end{aligned}$$

Thus, it suffices to set $B = \{t \mid v^\varepsilon(t) > 0\}$. The lemma is proved. □

Denote by S' the set of sequences $\{u'\}$ in $\mathcal{U}$ satisfying the relations

$$\begin{aligned} &\|u'\|_\infty \to 0, \quad \operatorname{meas}\{t \mid u'(t) \neq 0\} \to 0, \quad g(u) + g'(u)u' \leq 0, \\ &g_j'(u)u'\chi_{\{\hat{a}_j \geq \varepsilon\}} = 0, \ j = 1, \ldots, q, \quad \varepsilon \to +0. \end{aligned} \tag{87}$$

The following assertion holds.

Lemma 9.3. *For any $\{u'\} \in S'$, $C' < C$, and $\lambda \in M(C\Gamma)$ we have*

$$\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)u', u'\rangle + a(g'(u)u')_-\right)\,\mathrm{dt} \geq C' \int_0^T |u'|^2\,\mathrm{dt}, \tag{88}$$

starting from a certain number of the sequence.

Proof. Assume the contrary: there exist a sequence $\{u'\} \in S'$, a number $C' < C$, and an element $\lambda \in M(C\Gamma)$ such that

$$\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)u', u'\rangle + a(g'(u)u')_-\right)\,\mathrm{dt} \leq C' \int_0^T |u'|^2\,\mathrm{dt}. \tag{89}$$

Since $\{u'\}$ satisfies (87), the following estimates hold

$$g(u+u') \le k_1|u'|^2, \quad |g_j(u+u')|\chi_{\{\hat{a}_j\ge\varepsilon\}} \le k_1|u'|^2, \quad j=1,\dots,q,$$

where $k_1>0$ does not depend on the member of the sequence. Due to the hypothesis of linear independence of gradients $g_j'(u)$, there exists a sequence of corrections $\{v\}$ such that

$$g(u+u'+v)\le 0, \quad g_j(u+u'+v)\chi_{\{\hat{a}_j\ge\varepsilon\}}=0, \quad |v|\le k_2|u'|^2,$$

where $k_2>0$ does not depend on the member of the sequence. Set $\{\delta u\} = \{u'+v\}$. We, obviously, have

$$|\delta u|^2 = |u'|^2 + r_1, \quad \langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle = \langle \overline{H}_{uu}(w,p,a)u', u'\rangle + r_2,$$

where $|r_i|\le k_3|u'|^3$, $i=1,2$, and $k_3>0$ does not depend on the member of the sequence. Moreover, in virtue of (51), for any $j=1,\dots,q$, we have

$$a_j g_j(u+\delta u)_- = a_j g_j(u+\delta u)_-\chi_{\{\hat{a}_j<\varepsilon\}} \le a_j(g_j'(u)u')_- + r_{4j},$$

where $|r_{4j}|\le \varepsilon k_{4j}(|\delta u|^2+|v|)\le \varepsilon k_{5j}|u'|^2$, and $k_{4j}>0$ and $k_{5j}>0$ does not depend on the member of the sequence. Consequently,

$$\int_0^T |\delta u|^2\,\mathrm{d}t = \int_0^T |u'|^2\,\mathrm{d}t + o(\gamma'), \tag{90}$$

$$\int_0^T \langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle\,\mathrm{d}t = \int_0^T \langle \overline{H}_{uu}(w,p,a)u', u'\rangle\,\mathrm{d}t + o(\gamma'), \tag{91}$$

$$\int_0^T a_j g_j(u+\delta u)_-\,\mathrm{d}t \le \int_0^T a_j(g_j'(u)u')_-\,\mathrm{d}t + o(\gamma'), \tag{92}$$

where $\gamma' = \int\limits_0^T |u'|^2\,\mathrm{d}t$. These relations imply the inequality

$$\begin{array}{rl} & \int_0^T (\langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle + a\delta g_-)\,\mathrm{d}t \\ \le & \int_0^T (\langle \overline{H}_{uu}(w,p,a)u', u'\rangle + a(g'(u)u')_-)\,\mathrm{d}t + o(\gamma'), \end{array} \tag{93}$$

where $\delta g_- = g_-(u+\delta u) - g_-(u)$. Since $\overline{H}_u(w,p,a)=0$, $\delta_u\overline{H} = \delta_u H + a\delta g$, and $\delta g + \delta g_- = 0$, where $\delta_u\overline{H} = \overline{H}(u+\delta u, y,p,a) - \overline{H}(u,y,p,a)$, etc., we obtain

$$\begin{array}{rl} & \int_0^T (\langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle + a\delta g_-)\,\mathrm{d}t \\ = & \int_0^T (\delta_u\overline{H} + a\delta g_-)\,\mathrm{d}t + o(\int_0^T |\delta u|^2\,\mathrm{d}t) \\ = & \int_0^T \delta_u H\,\mathrm{d}t + o(\int_0^T |\delta u|^2\,\mathrm{d}t). \end{array} \tag{94}$$

Combining relations (89) and (94), we get

$$\int_0^T \delta_u H\,\mathrm{d}t + o\left(\int_0^T |\delta u|^2\,\mathrm{d}t\right) \le C'\int_0^T |\delta u|^2\,\mathrm{d}t. \tag{95}$$

But since the sequence $\{\delta u\}$ satisfies the condition $g(u+\delta u)\le 0$ and $\lambda\in M(C\Gamma)$, we have $\delta_u H \ge C|\delta u|^2$, and hence we get $\int_0^T \delta_u H\,\mathrm{d}t \ge C\int_0^T |\delta u|^2\,\mathrm{d}t$, contradicting to (95). $\square$

10 Passage to the set of critical variations.

Recall that S_1 was defined as the set of sequences $\{\delta w\}$ in $\mathcal{W}$ satisfying relations (52) and (53). Equivalently, S_1 is the set of sequences such that $\|\delta w\|_\infty \to 0$ and relations (53), (57), and (58) hold. For the set of functionals $w_1^*, \dots, w_s^*$, defined in Sec. 8, relations (58) imply the upper estimates $\langle w_i^*, \delta w\rangle \le o(\sqrt{\gamma})$, $i = 1, \dots, s$. This easily follows from the fact that each of the functionals $\langle w_i^*, \delta w\rangle$ belongs to the set Q and therefore has the form $\sum_{i\in I_\phi\cup\{0\}} \mu_i \phi_i'(\eta)\delta\eta + \sum_{i=r_1+1}^{r} \mu_i\phi_i'(\eta)\delta\eta$, where $\mu_i \ge 0$ for all $i \in I_\phi \cup \{0\}$. Further, recall that each functional w_i^* has the integral representation $\langle w_i^*, \delta w\rangle = -\int_0^T a^i g'(u)\delta u\, dt$ on the cone $\mathcal{R}_{i-1}$, $i = 1, \dots, s$. Let us show that, for each of these representations, the same upper estimate holds on any sequence from S_1.

Lemma 10.1. *For any sequence $\{\delta w\} \in S_1$ we have*

$$-\int_0^T a^i g'(u)\delta u\, dt \le o(\sqrt{\gamma(\delta w)}), \quad i = 1, \dots, s. \tag{96}$$

Proof. Let $\{\delta w\} \in S_1$ and $\{\varepsilon\}$ be the corresponding sequence of positive numbers converging to zero (see (53)). For any $j = 1, \dots, q$, define a sequence of sets $B_j^0 = \{t \mid 0 < \hat a_j(t) < \varepsilon\}$. Set $B^0 = \bigcup_{j=1}^q B_j^0$, $\delta u_{B^0} = \delta u \chi_{B^0}$ and $\delta u^0 = \delta u - \delta u_{B^0} = \delta u(1-\chi_{B^0})$, where χ_{B^0} is the characteristic function of the set B^0. Since $\operatorname{meas} B^0 \to 0$, we have $\|\delta u_{B^0}\|_1 \le \sqrt{\operatorname{meas} B^0}\|\delta u\|_2 = o(\sqrt{\gamma(\delta w)})$. Moreover, we obviously have $\hat a_j g_j'(u)\delta u^0 = 0$, $j = 1, \dots, q$. Let us define a sequence $\{\delta y^0\}$ such that all members of the sequence $\{\delta w^0\} = \{(\delta u^0, \delta y^0)\}$ belong to the cone $\mathcal{R}_0$. From (57), we get $\delta\dot y = f_y(w)\delta y + f_u(w)\delta u + \zeta$, where $\|\zeta\|_1 = o(\sqrt{\gamma(\delta w)})$. Define δy^0 as the solution to the equation $\delta\dot y^0 = f_y(w)\delta y^0 + f_u(w)\delta u^0$, $\delta y^0(0) = \delta y(0)$. Set $\bar y = \delta y - \delta y^0$. Then $\bar y$ satisfies the equation $\dot{\bar y} = f_y(w)\bar y + f_u(w)\delta u_{B^0} + \zeta$, $\bar y(0) = 0$, which in view of estimates $\|\delta u_{B^0}\|_1 = o(\sqrt{\gamma(\delta w)})$ and $\|\zeta\|_1 = o(\sqrt{\gamma(\delta w)})$ implies that $\|\bar y\|_\infty = o(\sqrt{\gamma(\delta w)})$. This and the estimate $\langle w_1^*, \delta w\rangle \le o(\sqrt{\gamma(\delta w)})$ imply that $\langle w_1^*, \delta w^0\rangle \le o(\sqrt{\gamma(\delta w)})$, where $\delta w^0 = (\delta u^0, \delta y^0)$. But now $\delta w^0 \in \mathcal{R}_0$, therefore we get $-\int_0^T a^1 g'(u)\delta u^0\, dt \le o(\sqrt{\gamma(\delta w)})$. Since $\|\delta u_{B^0}\|_1 = o(\sqrt{\gamma(\delta w)})$, we also get $-\int_0^T a^1 g'(u)\delta u\, dt \le o(\sqrt{\gamma(\delta w)})$.

Now, using Lemma 9.1, let us change the sequence $\{\delta u^0\}$ on the sets $\mathcal{M}_j^1 := \{t \mid a_j^1(t) > 0\}$, $j = 1, \dots, q$, and define a new sequence $\{\delta w^1\}$ whose members belong to the cone $\mathcal{R}_1$. Set $v_j^0 = (g_j'(u)\delta u^0)_-$. Then we have $\int_0^T a^1 v_j^0\, dt = o(\sqrt{\gamma(\delta w)})$. According to Lemma 9.1, there exist two sequences of sets $\{A_j^1\}$ and $\{B_j^1\}$ such that $A_j^1 \cup B_j^1 = \mathcal{M}_j^1$, $A_j^1 \cap B_j^1 = \emptyset$, $\operatorname{meas} B_j^1 \to 0$, and $\|v_j^0 \chi_{A_j^1}\|_\infty = o(\sqrt{\gamma(\delta w)})$. Set $B^1 = \bigcup_{j=1}^q B_j^1$ and $\delta u_{B^1} = \delta u^0 \chi_{B^1}$. Then $\operatorname{meas} B^1 \to 0$ and therefore $\|\delta u_{B^1}\|_1 = o(\sqrt{\gamma(\delta w)})$. In view of the hypothesis of linear independence of gradients $g_j'(u)$, from the estimate

$\|v_j^0\chi_{A_j^1}\|_\infty = o(\sqrt{\gamma(\delta w)})$ it follows that there exists a sequence $\{\bar{u}^0\}$ such that $\|\bar{u}^0\|_\infty = o(\sqrt{\gamma(\delta w)})$, $\bar{u}^0\chi_{B^1} = 0$, $g_j'(u)\bar{u}^0 = g_j'(u)\delta u^0$ a. e. on the set A_j^1, and $g_j'(u)\bar{u}^0 = 0$ a. e. on the set $\{t \mid g_j(u(t)) = 0\} \setminus A_j^1$, $j = 1, \ldots, q$. Set $\delta u^1 = \delta u^0 - \delta u_{B^1} - \bar{u}^0$ and define δy^1 as the solution to the equation $\delta\dot{y}^1 = f_y(w)\delta y^1 + f_u(w)\delta u^1$, $\delta y^1(0) = \delta y(0)$. Then it is easy to see that all members of the sequence $\{\delta w^1\} = \{(\delta u^1, \delta y^1)\}$ belong to the cone $\mathcal{R}_1$, and, for this sequence, we have $\langle w_2^*, \delta w^1\rangle \leq o(\sqrt{\gamma(\delta w)})$. This implies that $-\int_0^T a^2 g'(u)\delta u^1\,dt \leq o(\sqrt{\gamma(\delta w)})$, and then $-\int_0^T a^2 g'(u)\delta u\,dt \leq o(\sqrt{\gamma(\delta w)})$. Similarly, we transform the sequence $\{\delta u^1\}$ on the sets $\mathcal{M}_j^2 = \{t \mid a_j^2(t) > 0\}$, $j = 1, \ldots, q$, and deduce that $-\int_0^T a^2 g'(u)\delta u\,dt \leq o(\sqrt{\gamma(\delta w)})$, etc. □

Let us estimate $C_\gamma(\Phi_C^1, S_1)$ from below (see (55)). To this end, chose any sequence $\{\delta w\} \in S_1$. Since $a^i(g'(u)\delta u)_- = \sum_j a_j^i(g_j'(u)\delta u)_-$, it follows from Lemma 10.1 that $\int_0^T a_j^i(g_j'(u)\delta u)_-\,dt \leq o(\sqrt{\gamma})$, $i = 1, \ldots, s$, $j = 1, \ldots, q$. Set $v_j = (g_j'(u)\delta u)_-\chi_{\{g_j=0\}}$, $j = 1, \ldots, q$. Then $\int_0^T a_j^i v_j\,dt \leq o(\sqrt{\gamma})$, $i = 1, \ldots, s$, $j = 1, \ldots, q$, where $a_j^i \geq 0$, $v_j \geq 0$, $i = 1, \ldots, s$, $j = 1, \ldots, q$. Recall that, by definition, $\mathcal{M}_j^i = \{t \mid a_j^i(t) > 0\}$, $i = 0, \ldots, s$, $j = 1, \ldots, q$, and $\mathcal{M}_j = \bigcup_{i=0}^s \mathcal{M}_j^i$, $j = 1, \ldots, q$, where $a^0 = \hat{a}$. Let $i \in \{1, \ldots, s\}$, $j \in \{1, \ldots, q\}$. By Lemma 9.1, for the sequence $\{v_j\}$ and the set $\mathcal{M}_j^i$, there exist two sequences of sets $\{A_j^i\}$ and $\{B_j^i\}$ such that

$$A_j^i \cup B_j^i = \mathcal{M}_j^i, \quad A_j^i \cap B_j^i = \emptyset, \quad \operatorname{meas} B_j^i \to 0, \quad \|v_j\chi_{A_j^i}\|_\infty = o(\sqrt{\gamma}). \tag{97}$$

Set

$$B = \left(\bigcup_{i=1}^{s}\bigcup_{j=1}^{q} B_j^i\right) \bigcup \left(\bigcup_{j=1}^{q}\{t \mid 0 < \hat{a}_j(t) < \varepsilon\}\right). \tag{98}$$

Then in view of (97) and (98) we have: $\operatorname{meas} B \to 0$, $\|v_j\chi_{\mathcal{M}_j\setminus B}\|_\infty = o(\sqrt{\gamma})$. In view of the hypothesis of linear independence of gradients $g_j'(u)$, there exists a sequence of functions $\{\bar{u}\}$ in $\mathcal{U}$ such that

$$\begin{aligned}&\|\bar{u}\|_\infty = o(\sqrt{\gamma}), \quad \bar{u}\chi_B = 0,\\ &g_j'(u)\bar{u}\chi_{\mathcal{M}_j\setminus B} = v_j\chi_{\mathcal{M}_j\setminus B}, \quad g_j'(u)\bar{u}\chi_{\{g_j(u)=0\}\setminus\mathcal{M}_j} = 0, \quad j = 1, \ldots, q.\end{aligned}$$

Define a sequence $\{u'\}$ by the relation $u' = \delta u\chi_B$. Obviously, $\{u'\} \in S'$ (see (87)). Set $\delta u_1 = \delta u - \bar{u} - u'$, $\delta w_1 = (\delta u_1, \delta y)$. Then, for the sequences $\{\delta w\}$, $\{\delta w_1\}$, and $\{u'\}$, the following relations hold

$$\gamma(\delta w) = \gamma(\delta w_1) + \int_0^T |u'|^2\,dt + o(\gamma(\delta w)), \tag{99}$$

$$\Omega(\delta w, \lambda) = \Omega(\delta w_1, \lambda) + \int_0^T \langle \overline{H}_{uu}(w, p, a)u', u'\rangle\,dt + o(\gamma(\delta w)), \tag{100}$$

$$\int_0^T a(g'(u)\delta u)_- \,\mathrm{d}t \geq \int_B a(g'(u)\delta u)_- \,\mathrm{d}t = \int_0^T a(g'(u)u')_- \,\mathrm{d}t, \tag{101}$$

uniformly on Λ_0. Consequently,

$$\begin{aligned} \Phi_C^1(\delta w) &:= \max_{\lambda\in M(C\Gamma)} \left\{\Omega(\delta w,\lambda) + \textstyle\int_0^T a(g'(u)\delta u)^- \,\mathrm{d}t\right\} \\ \geq & \max_{\lambda\in M(C\Gamma)} \left\{\Omega(\delta w_1,\lambda) + \textstyle\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)u',u'\rangle + a(g'(u)u')_-\right) \,\mathrm{d}t\right\} \\ & +o(\gamma(\delta w)). \end{aligned} \tag{102}$$

Moreover, $\|\delta w_1\|_\infty \to 0$, $\delta u_1 \in \mathcal{S}_u$, $\|\delta\dot{y} - f_u(w)\delta u_1 - f_y(w)\delta y\|_1 = o(\sqrt{\gamma(\delta w)})$, and relations (58) hold. Choose any $C' < C$. Set $w' = (0,u')$, $\gamma(w') = \int_0^T |u'|^2 \,\mathrm{d}t$. Then in view of Lemma 9.3 we get

$$\begin{aligned} & \max_{\lambda\in M(C\Gamma)} \left\{\Omega(\delta w_1,\lambda) + \textstyle\int_0^T \langle \overline{H}_{uu}(w,p,a)u',u'\rangle + \int_0^T a(g'(u)\delta u')_- \,\mathrm{d}t\right\} \\ \geq & \max_{\lambda\in M(C\Gamma)} \Omega(\delta w_1,\lambda) + C'\gamma(w') = \Omega_{M(C)}(\delta w_1) + C'\gamma(w'). \end{aligned}$$

Combining this inequality with (102) we get

$$\Phi_C^1(\delta w) \geq \Omega_{M(C)}(\delta w_1) + C'\gamma(w') + o(\gamma(\delta w)). \tag{103}$$

Furthermore, the relation $\|\delta\dot{y} - f'(w)\delta w_1\|_1 = o(\sqrt{\gamma(\delta w)})$ may be written as $\delta\dot{y} - f_y(w)\delta x - f_u(w)\delta u_1 = \zeta$, $\|\zeta\|_1 = o(\sqrt{\gamma(\delta w)})$. Let us define a sequence $\{y_\zeta\}$ by the equation $\dot{y}_\zeta = f_y(w)y_\zeta + \zeta$, $y_\zeta(0) = 0$. Then $\|y_\zeta\|_\infty = o(\sqrt{\gamma(\delta w)})$. Set $\delta y_S = \delta y - y_\zeta$, $\delta w_S = (\delta u_1, \delta y_S)$. Then $\delta\dot{y}_S - f'(w)\delta w_S = 0$. Thus, we get

$$\begin{aligned} &\phi_i'(\eta)\delta\eta_S \leq o(\sqrt{\gamma(\delta w)}),\ i \in I_\phi(\eta)\cup\{0\}, \\ &|\phi_i'(\eta)\delta\eta_S| = o(\sqrt{\gamma(\delta w)}),\ i = r_1+1\dots,r, \\ &\delta w_S \in \mathcal{S}, \quad \|\delta w_S\|_\infty \to 0. \end{aligned}$$

By Lemma 8.4 the system

$$\begin{aligned} &\phi_i'(\eta)(\delta\eta_S + \bar{\eta}) \leq 0,\ i \in I_\phi(\eta)\cup\{0\}, \\ &\phi_j'(\eta)(\delta\eta_S + \bar{\eta}) = 0,\ j = r_1+1,\dots,r, \quad \bar{w} \in \mathcal{S} \end{aligned}$$

is compatible, and there exists a solution $\bar{w}$ to this system such that $\|\bar{w}\|_\infty = o(\sqrt{\gamma(\delta w)})$. Set $\delta w_{\mathcal{K}} = \delta w_S + \bar{w}$. Then for the sequence $\{\delta w_{\mathcal{K}}\}$ we obviously have

$$\delta w_{\mathcal{K}} \in \mathcal{K}, \quad \|\delta w_{\mathcal{K}}\|_\infty \to 0, \tag{104}$$

$$\gamma(\delta w_1) = \gamma(\delta w_{\mathcal{K}}) + o(\gamma(\delta w)), \tag{105}$$

$$\Omega_{M(C)}(\delta w_1) = \Omega_{M(C)}(\delta w_{\mathcal{K}}) + o(\gamma(\delta w)). \tag{106}$$

Assume that there exists $C_{\mathcal{K}} > 0$ such that $\Omega_C(\delta w) \geq C_{\mathcal{K}}\gamma(\delta w)$ for any $\delta w \in \mathcal{K}$ (cf. (20)). Then we have

$$\Omega_{M(C)}(\delta w_{\mathcal{K}}) \geq C_{\mathcal{K}}\gamma(\delta w_{\mathcal{K}}), \tag{107}$$

since $\delta w_{\mathcal{K}} \in \mathcal{K}$. Relations (103), (106), and (107) imply

$$\Phi^1_C(\delta w) \geq C_{\mathcal{K}}\gamma(\delta w_{\mathcal{K}}) + C'\gamma(w') + o(\gamma(\delta w)), \tag{108}$$

and combining (99) with (105) we get

$$\gamma(\delta w) = \gamma(\delta w_{\mathcal{K}}) + \gamma(w') + o(\gamma(\delta w)). \tag{109}$$

Consequently, the following relations hold

$$\liminf \frac{\Phi^1_C(\delta w)}{\gamma(\delta w)} \geq \liminf \frac{C_{\mathcal{K}}\gamma(\delta w_{\mathcal{K}}) + C'\gamma(w')}{\gamma(\delta w_{\mathcal{K}}) + \gamma(w')} \geq \min\{C_{\mathcal{K}}, C'\},$$

which imply that $C_\gamma(\Phi^1_C, S_1) \geq \min\{C_{\mathcal{K}}, C'\}$. This inequality holds for any $C' < C$. Hence $C_\gamma(\Phi^1_C, S_1) \geq \min\{C_{\mathcal{K}}, C\}$. Combining this inequality with (49) and (55) we get $C_\gamma(\Psi_{\Lambda_0}, S_{\sigma\gamma}) \geq \min\{C_{\mathcal{K}}, C\}$. In view of (31) and Proposition 2.1 this completes the proof of Theorem 3.1.

Acknowledgments I obtained Theorem 3.1 about 30 years ago. At that time I had permanent contacts with my teacher A.A. Milyutin, and some of his ideas were used in the proof of Theorem 3.1. Unfortunately, this proof (contained in [7]) was never published, while the formulation of the theorem (for a more general problem) was published in [3], Appendix. A new interest to similar results and recent publications in the field made me to recall and revise the proof. The revision (with partial modification) was mainly done during my stay at Ecole Polytechnique, France, in 2007. I would like to express my gratitude to F. Bonnans for arranging this visit and our numerous fruitful discussions, initiated this publication.

11 Appendix

In sec. 11.1-11.3, following [7], we consider a system of linear inequalities on a convex cone, and study two questions: (a) the existence of a solution to the system; (b) the existence of Hoffman's type [2] upper bounds for the distance from the origin to the set of the solutions to the system. In Sec. 11.4, we formulate an abstract notion of support of critical cone and prove that the system of inequalities defining this cone possesses a Hoffman's error bound on the support. The main results of sec. 11.1-11.3 were published in [6] without proofs. Sec. 11.4 was written in 2007.

11.1 On the compatibility of a linear system on a cone.

Let X be a linear space, $l_i : X \to \mathbb{R}$, $i = 1, \ldots, k$ a set of linear functionals, K a convex cone in X. Consider the following system:

$$\langle l_i, x\rangle + \xi_i \leq 0, \quad i = 1, \ldots, k, \quad x \in K. \tag{110}$$

We will write it briefly as

$$l(x) + \xi \le 0, \quad x \in K \tag{111}$$

where $l : X \to \mathbb{R}^k$ is a linear operator which corresponds to the set of linear functionals $l_i : X \to \mathbb{R}$, $i = 1, \dots, k$, and $\xi = (\xi_1, \dots, \xi_k)^* \in \mathbb{R}^k$.

Set $\Omega = l(K) + \mathbb{R}^k_+$, where $l(K)$ is the image of the cone K under the mapping $l : K \to \mathbb{R}^k$. It is clear that Ω is a convex cone in $\mathbb{R}^k$, and system (111) is compatible iff $(-\xi) \in \Omega$.

Lemma 11.1. *Assume that the cone Ω is closed. Then system (111) is compatible iff the relations*[1] $\alpha \in \mathbb{R}^{k*}_+$, $\alpha l(K) \ge 0$ *imply that* $\alpha\xi \le 0$.

Proof. Let system (111) be incompatible, i.e., $-\xi \notin \Omega$. Since Ω is convex and closed cone, there exits a vector $\alpha \in \mathbb{R}^{k*}$ such that $\alpha(-\xi) < 0 \le \alpha\Omega$, where $\alpha\Omega = \{\alpha\omega \mid \omega \in \Omega\}$. The inequality $\alpha\Omega \ge 0$ implies that $\alpha l(K) \ge 0$ and $\alpha\mathbb{R}^k_+ \ge 0$. Consequently, $\alpha \ge 0$. Moreover, $\alpha\xi > 0$. Thus, we obtain: if the relations $\alpha \in \mathbb{R}^{k*}_+$, $\alpha l(K) \ge 0$ imply that $\alpha\xi \le 0$, then system (111) is compatible.

Vice versa, let system (111) be compatible, and let x_0 be a solution to this system. Let $\alpha \in \mathbb{R}^{k*}$ and $\alpha l(K) \ge 0$. Since $x_0 \in K$, we have $\alpha l(x_0) \ge 0$. Consequently, $\alpha\xi \le \alpha l(x_0) + \alpha\xi = \alpha(l(x_0) + \xi) \le 0$. The lemma is proved. □

If $K = L$, where L is a subspace in X, then the cone $\Omega = l(L) + \mathbb{R}^k_+$ is closed. It follows from the fact that, in this case, Ω is a finite faced cone, since Ω is the sum of the subspace $l(L)$ and a finite faced cone $\mathbb{R}^k_+$. Moreover, the inequality $\alpha l(L) \ge 0$ is equivalent to the equality $\alpha l(L) = 0$. Thus we obtain the following consequence from Lemma 11.1.

Corollary 11.2. *Let $K = L$, where L is a subspace in X. Then system (111) is compatible iff the relations $\alpha \in \mathbb{R}^{k*}_+$, $\alpha l(L) = 0$ imply that $\alpha\xi \le 0$.*

11.2 On the existence of Hoffman's error bound on a cone.

Now consider a second question: how to estimate the distance from the origin to the set of the solutions to the system of linear inequalities on a cone.

Definition 11.3. We say that system (111) has the *Hoffman's error bound* if there exists a finite dimensional subspace $H \subset X$ such that for any norm $\|\cdot\|$ in H there exists a constant $N = N(l, K, H, \|\cdot\|) > 0$ such that the following condition holds: if system (111) is compatible, then the system

$$l(x) + \xi \le 0, \quad x \in K \cap H \tag{112}$$

[1]For $A \subset \mathbb{R}$ the inequality $A \ge 0$ means that $\inf A \ge 0$.

is compatible too, and there exists a solution x_0 to system (112) such that

$$\|x_0\| \leq N|\xi_+|, \tag{113}$$

where $|\xi_+| = \max_{1\leq i\leq k} \xi_{i+}$, and $\xi_{i+} = \max\{\xi_i, 0\}$.

The well-known Hoffman's lemma state that if $K = X$ and X is finite dimensional, then system (111) has the Hoffman's error bound.

Below we will study the question: when system (111) has the Hoffman's error bound? We will not assume that the space X is finite dimensional.

Recall that a cone C is called *finite generated* if there exists a finite set of its elements $a_1, \ldots, a_s$ (*generators od the cone*) such that each element $x \in C$ may be represented as $x = \lambda_1 a_1 + \ldots + \lambda_s a_s$ with $\lambda_1 \geq 0, \ldots, \lambda_s \geq 0$. In a finite dimensional space, a cone is finite faced iff this cone is finite generated.

Hoffman's lemma has the following simple generalization.

Lemma 11.4. *Assume that $\Omega = l(K) + \mathbb{R}^k_+$ is a finite faced cone. Then system (111) has the Hoffman's error bound.*

Proof. Since Ω is a finite faced cone in $\mathbb{R}^k$, then Ω is finite generated. Let $\xi^i = l(x^i) + \eta^i$, $i = 1, \ldots, r$ be the generators of the cone Ω, where $x^i \in K$, $\eta^i \in \mathbb{R}^k_+$, $i = 1, \ldots, r$. Let H be a linear span of the set $\{x^1, x^2, \ldots, x^r\}$. Then it is easy to see that $\Omega = l(K \cap H) + \mathbb{R}^k_+$. The cone $K \cap H$ is finite generated and hence finite faced in H. Consequently, there exist linear functionals $m_j : H \to \mathbb{R}$, $j = 1, \ldots, s$ such that $K \cap H = \{x \in H \mid \langle m_j, x\rangle \leq 0,\ j = 1, \ldots, s\}$. In the finite dimensional subspace H, let us consider the system

$$\langle l_i, x\rangle + \xi_i \leq 0, \quad i = 1, \ldots, k, \quad \langle m_j, x\rangle \leq 0,\ j = 1, \ldots, s. \tag{114}$$

The set of solutions to system (114) contains in the set of solutions to system (114); moreover, both sets are empty or nonempty simultaneously. Now, using Hoffman's lemma for system (114), we obtain needed error bound. The lemma is proved. □

Corollary 11.5. *If $K = L$ is a subspace in X, then $\Omega = l(K) + \mathbb{R}^k_+$ is a finite faced cone, and hence system (111) has the Hoffman's error bound.*

Thus, Hoffman's lemma has an analog in any linear space.

11.3 Special case, where Hoffman's error bound holds.

Lemma 11.6. *Let Y be a locally convex topological space, $X = Y^*$ a dual space, K_Y a closed convex cone in Y, $K = K_Y^*$ a dual cone in X, l_i, $i = 1, \ldots, k$ a set of elements of the space Y considered as linear functionals on X. Denote by K_l the cone in Y generated by the elements l_i, $i = 1, \ldots, k$, i.e. $K_l = \{y = \alpha l \mid \alpha \in \mathbb{R}^{k*}_+\}$. Assume that $K_l \cap K_Y \subset S \subset K_Y$, where S is a* *subspace in Y. Then $\Omega = l(K)+\mathbb{R}^k_+$ is a finite faced cone, and hence system (111) has the Hoffman's error bound. Moreover, system (111) is compatible iff the conditions $\alpha \in \mathbb{R}^{k*}_+$, $\alpha l \in K_Y$ imply that $\alpha\xi \leq 0$. Particularly, for any $x_0 \in K$ the system $l(x+x_0) \leq 0$, $x \in K$ is always compatible.*

The proof of this lemma will be based on the following theorem.

Theorem 11.7. *Let Y be a locally convex topological space, $K_1 \subset Y$ a finite generated cone, $K_2 \subset Y$ a convex closed cone. Assume that $K_1 \cap K_2 \subset S \subset K_2$, where S is a subspace in Y. Then $(K_1 \cap K_2)^* = K_1^* + K_2^*$.*

Proof. (a) At first, consider the case $K_1 \cap K_2 = S \subset K_2$, where S is a subspace in Y. Set $Q = K_1^* + K_2^*$. We have to show that $Q = S^*$.

It is obvious, that $S^* = (K_1 \cap K_2)^* \supset K_1^* + K_2^* = Q$. Next, it is easy to check that, for two nonempty cones, the dual cone to their sum is equal to the intersection of their dual cones. Therefore, $Q^* = (K_1^* + K_2^*)^* = K_1^{**} \cap K_2^{**}$. Since K_1 is a finite generated cone, we have $K_1^{**} = K_1$. Moreover, $K_1 \subset Y$. Consequently, $Q^* = K_1 \cap K_2^{**} = K_1 \cap (K_2^{**} \cap Y)$. Since K_2 is a convex closed cone and Y is a locally convex topological space, we have that $K_2^{**} \cap Y = K_2$ (here, local convexity of Y is important). Thus we obtain $Q^* = K_1 \cap K_2 = S$. Moreover, $Q \subset S^*$. Let us show that this implies the equality $Q = S^*$.

Assume that $Q \subset S^*$ and $Q \neq S^*$. Then there exists $x \in S^*$ such that $x \notin Q$. Consequently, there exists $x^* \in Q^*$ such that $\langle x^*, x\rangle < 0$. Since $x^* \in Q^*$, $Q^* = S$, and $x \in S^*$, we have $\langle x^*, x\rangle \geq 0$. We arrive at contradiction. Therefore our assumption $Q \neq S^*$ is not true, i.e., $Q = S^*$.

(b) Now, consider the general case $K_1 \cap K_2 \subset S \subset K_2$, where S is a subspace in Y. Since the cone K_1 is finite dimensional, without loss of generality, we may assume that the subspace S is finite dimensional.

Let us set $\hat{K}_1 = K_1 + S$. Let us show that $\hat{K}_1 \cap K_2 = S$. Since $0 \in K_1$, then $S \subset \hat{K}_1$. Moreover, $S \subset K_2$. Therefore, $S \subset \hat{K}_1 \cap K_2$.

Vice versa, let $y_2 \in K_2$ and $y_2 \in \hat{K}_1 = K_1 + S$, i.e. $y_2 = y_1 + s$, where $y_1 \in K_1$ and $s \in S$. Since S is a subspace, we have $-s \in S$. Therefore, $-s \in K_2$. Since K_2 is a convex cone, we have $y_1 = y_2 - s \in K_2$. Consequently, $y_1 \in K_1 \cap K_2 \subset S$. Therefore, $y_2 = y_1 + s \in S$. This implies that $\hat{K}_1 \cap K_2 \subset S$. Consequently, $\hat{K}_1 \cap K_2 = S$.

Since $K_1 \cap K_2 \subset S \subset K_2$, we have $K_1 \cap K_2 = K_1 \cap S$. The cone K_1 is finite generated, and S is a finite dimensional subspace, hence S may also be considered as a finite generated cone. As is well-known for finite generated cones, the dual cone to their intersection is equal to the sum of their dual cones. Therefore, $(K_1 \cap K_2)^* = (K_1 \cap S)^* = K_1^* + S^*$. Further, since $\hat{K}_1 \cap K_2 = S \subset K_2$ and $\hat{K}_1 = K_1 + S$ is a finite generated cone, in virtue of (a), we have $S^* = (\hat{K}_1 \cap K_2)^* = \hat{K}_1^* + K_2^*$. But $\hat{K}_1^* = (K_1 + S)^* = K_1^* \cap S^*$. Consequently, $S^* = K_1^* \cap S^* + K_2^*$. Thus, we get $(K_1 \cap K_2)^* = K_1^* + S^* = K_1^* + (K_1^* \cap S^*) + K_2^*$. The theorem is proved. □

Proof of Lemma 11.6. Set $C = (K_l \cap K_Y)^*$. Let us show that $\Omega = l(C) + \mathbb{R}^k_+$. Indeed, in virtue of Theorem 11.7, $C = K_l^* + K_Y^*$. Consequently, $l(C) = l(K_l^*) + l(K_Y^*)$. But, obviously, $l(K_l^*) \subset \mathbb{R}^k_+$, and hence $l(K_l^*) + \mathbb{R}^k_+ = \mathbb{R}^k_+$. Consequently, $l(C) + \mathbb{R}^k_+ = l(K_Y^*) + \mathbb{R}^k_+ = l(K) + \mathbb{R}^k_+ = \Omega$. Further, it is easy to see that $K_l \cap K_Y = K_l \cap S$. Since K_l is a finite generated cone and S is a subspace, $K_l \cap S$ is a finite generated cone too. Consequently, $C = (K_l \cap K_Y)^*$ is a finite faced cone in X. This implies that $\Omega = l(C) + \mathbb{R}^k_+$ is a finite faced cone too.

Since each finite faced cone is closed, by Lemma 11.1, the compatibility of the system $l(x) + \xi \leq 0$, $x \in K$ is equivalent to the condition: if $\alpha \in \mathbb{R}^k_+$, $\alpha l(K) \geq 0$, then $\alpha\xi \leq 0$. But since $l_i \in Y$ for all i, the condition $\alpha l(K) \geq 0$ is equivalent to the condition $\alpha l \in K^* \cap Y = K_Y$. This implies the last assertion of the lemma on the compatibility of the system. In particular, if $x_0 \in K$ and $\xi_0 = l(x_0)$, then the conditions $\alpha l \in K_Y$, $\alpha \in \mathbb{R}^k_+$ imply $\alpha l \in S \subset K_Y$. Since S is a subspace, $K = K_Y^*$, and $x_0 \in K$, we have $\alpha l(x_0) = 0$, i.e. $\alpha\xi_0 = 0$. Thus, in this case the system is compatible. The lemma is proved. $\square$

11.4 Abstract notion of support of critical cone

Let Y be a Banach space, $X = Y^*$ a dual space, l_i, $i = 1, \dots, k$ a set of elements in the space Y considered as linear functionals on X, C a closed convex cone in Y, $\Omega = C^*$ a dual cone in X. Set $K = \{x \in \Omega \mid l_i(x) \leq 0,\ i = 1, \dots, k\}$. The cone K will serve us as an abstract analog of the critical cone (see, eg., [1]). We will define the notion of support of the cone K. Set $Q = \{l = \sum_{i=1}^k \alpha_i l_i \mid \alpha_i \geq 0, \quad i = 1, \dots, k\}$. For any $y \in Q$ and for any $x \in K$ we obviously have: $\langle y, x \rangle \leq 0$.

Assume that there exists an element $y_1 \in Q$ such that $y_1(\Omega) \geq 0$, i.e. $y_1 \in C$, and let there exists $x_1 \in \Omega$ such that $\langle y_1, x_1 \rangle > 0$. Set $\Omega_1 = \{x \in \Omega \mid \langle y_1, x \rangle = 0\}$. Then $K \subset \Omega_1$. Moreover, the cone Ω_1 is dual to the cone $C_1 := C + \mathrm{Span}\{y_1\}$, where $\mathrm{Span}\{y_1\}$ is a one dimensional subspace generated by vector y_1. Indeed, $(C + \mathrm{Span}\{y_1\})^* = C^* \cap \mathrm{Span}\{y_1\}^* = \Omega \cap \{x \in X \mid \langle y_1, x \rangle = 0\} = \Omega_1$.

Assume that there exists an element $y_2 \in Q$ such that $y_2(\Omega_1) \geq 0$, i.e. $y_2 \in C_1$, and let there exists $x_2 \in \Omega_1$ such that $\langle y_2, x_2 \rangle > 0$. Set $\Omega_2 = \{x \in \Omega_1 \mid \langle y_2, x \rangle = 0\}$. Then the cone Ω_2 is dual to the cone $C_2 := C_1 + \mathrm{Span}\{y_2\}$.

Assume that we get the cones Ω_1, Ω_2,..., Ω_{s-1}, dual to the cones C_1, C_2, ...,C_{s-1}, respectively, and functionals $y_1 \in Q \cap C$, $y_2 \in Q \cap C_1$, ... , $y_{s-1} \in Q \cap C_{s-2}$. Let there exists an element $y_s \in Q$ such that $y_s(\Omega_{s-1}) \geq 0$, i.e. $y_s \in C_{s-1}$, and there exists $x_s \in \Omega_{s-1}$ such that $\langle y_s, x_s \rangle > 0$. Set $\Omega_s = \{x \in \Omega_{s-1} \mid \langle y_s, x \rangle = 0\}$. Then the cone Ω_s is dual to the cone $C_s := C_{s-1} + \mathrm{Span}\{y_s\}$. The vectors y_1, ... y_s are linear independent, therefore this process will be finished on a final step s. Let us show that the vectors y_1, ... y_s are linearly independent. Let there exist numbers λ_1, ...,

λ_{s-1} such that $y_s = \lambda_1 y_1 + \ldots + \lambda_{s-1} y_{s-1}$. Since $y_k(\Omega_k) \equiv 0$, $k = 1, \ldots, s-1$ and $\Omega_s \subset \Omega_{s-1} \subset \ldots \subset \Omega_1 \subset \Omega$, we get $y_k(\Omega_{s-1}) \equiv 0, \quad k = 1, \ldots, s-1$,, and hence $y_s(\Omega_{s-1}) \equiv 0$. We arrive at contradiction. Consequently, $y_s \notin \mathrm{Span}\{y_1, \ldots, y_{s-1}\}$.

Let the system y_1, ... y_s be maximal. Then for the cone Ω_s we have: if $y \in Q$ be such that $y(\Omega_s) \geq 0$, i.e. $y \in C_s$, then $y(\Omega_s) \equiv 0$. In this case we have $(-y) \in C_s$. It means that the linear span of the intersection $Q \cap C_s$ is a subspace in C_s. Denote this subspace by H. Set $\mathcal{S} = \Omega_s, \quad \mathcal{S}_Y = C_s$. Then $\mathcal{S}_Y^* = \mathcal{S}$, $\mathcal{S}_Y \cap Q \subset H \subset \mathcal{S}_Y$, where H is a subspace in the cone $\mathcal{S}_Y$. Thus, for the system of functionals $l_1, \ldots, l_k$ on the cone $\mathcal{S}$, the assumptions of Lemma 11.6 are satisfied . By this lemma, for any $x_0 \in \mathcal{S}$, the system $l(x + x_0) \leq 0$, $x \in \mathcal{S}$ is compatible and has the Hoffman's error bound. We call $\mathcal{S}$ the *support* of the cone K.

Remark 11.8. Denote by $\hat{l}_i$, $i = 1, \ldots, \hat{k}$ functionals of the set $l_1, \ldots, l_k$ such that $\hat{l}_i(K) \equiv 0$. The functionals l_i which do not posses this property we denote by $\tilde{l}_i$, $i = 1, \ldots, \tilde{k}$. By definition, for each $\tilde{l}_i$ there exits an element $\tilde{x}_i \in K$ such that $\tilde{l}_i(\tilde{x}_i) < 0$. Set $\tilde{x} = \sum \tilde{x}_i$. Then $\tilde{x} \in K$, and $\tilde{l}_i(\tilde{x}) < 0$ for all $i = 1, \ldots, \tilde{k}$. Set $\hat{\Omega} = \{x \in \Omega \mid \hat{l}_i(x) = 0 \ \forall i\}$. One can show that $K \subset \hat{\Omega} \subset \mathcal{S}$, and, for any $x_0 \in \hat{\Omega}$, the system $l_i(x_0 + x) \leq 0$, $i = 1, \ldots, k$, $x \in \hat{\Omega}$ is compatible and has the Hoffman's error bound. Question: is it true that $\hat{\Omega} = \mathcal{S}$? A simple example shows that this is not true. Let $k = 2$, l_1 be such that $l_1(\Omega) = \mathbb{R}$, and $l_2 = -l_1$. Then $\hat{l}_1 = l_1$, $\hat{l}_2 = l_2$ and $K = \hat{\Omega} = \{x \in \Omega \mid l_1(x) = 0\} \neq \Omega = \mathcal{S}$.

References

[1] A. Ya. Dubovitski and A. A. Milyutin "Extremum problems with constraints," [in Russian], *J. Comp. Sc. and Math. Phys.* **5**, No. 3 (1965), 395-453.

[2] A. J. Hoffman , "On approximate solutions of systems of linear inequalities," *J. Res. Nat'l Bur. Standarts*, No. 49 (1952), 263-265.

[3] E. S. Levitin, A. A. Milyutin, and N. P. Osmolovskii, "Higher-order local minimum conditions in problems with constraints," *Usp. Mat. Nauk*, **33**, No. 6 (1978), 85–148 .

[4] A. A. Milyutin, *Maximum Principle in the General Optimal Control Problem* [in Russian], Fizmatlit, Moscow, 2001.

[5] Milyutin, A.A. and Osmolovskii, N.P. , *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs, Vol. **180**, American Mathematical Society, Providence, 1998.

[6] N. P. Osmolovskii, "On a system of linear inequalities on a convex set" [in Russian], *Usp. Mat. Nauk*, **32**, No. 2 (1977), 223–224.

[7] N. P. Osmolovskii, 'Theory of higher order conditions in optimal control', Doctoral Thesis, MISI (Moscow Civil Engineering Institute), Moscow, 1988.