ESAIM: COCV 18 (2012) 452–482
DOI: 10.1051/cocv/2011101

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

SECOND-ORDER SUFFICIENT OPTIMALITY CONDITIONS FOR CONTROL PROBLEMS WITH LINEARLY INDEPENDENT GRADIENTS OF CONTROL CONSTRAINTS

Nikolai P. Osmolovskii[1,2,3]

Abstract. Second-order sufficient conditions of a bounded strong minimum are derived for optimal control problems of ordinary differential equations with initial-final state constraints of equality and inequality type and control constraints of inequality type. The conditions are stated in terms of quadratic forms associated with certain tuples of Lagrange multipliers. Under the assumption of linear independence of gradients of active control constraints they guarantee the bounded strong quadratic growth of the so-called "violation function". Together with corresponding necessary conditions they constitute a no-gap pair of conditions.

Mathematics Subject Classification. 49K15.

Received November 28, 2009. Revised November 27, 2010.
Published online June 22, 2011.

1. Introduction

In this paper we discuss sufficient second order conditions for bounded strong minimum in optimal control problems of ordinary differential equations with control constraints and constraints on the initial-final state. There exists an extensive literature on this subject, *cf.* Bonnans and Hermant [1], Bonnans and Shapiro [3], Bonnans and Osmolovskii [2], Levitin *et al.* [8], Malanowski [9,10], Maurer [11], Maurer and Pickenhain [12], Milyutin and Osmolovskii [17], Osmolovskii [19–22], Zeidan [23,24] and further literature cited in these papers. The no-gap second order conditions of Pontryagin and bounded strong minima in optimal control problems with initial-final state constraints and mixed state-control constraints, satisfying the hypothesis of linear independence of gradients (LIG) w.r.t. control of active mixed constraints, were formulated by the author in [8], pp. 155–156, but the proofs (written later in [19,20]) were not published. The aim of this paper is to publish the proofs of sufficient conditions [8] (partially modified in 2007–2008). For simplicity, we consider a less general problem than in [8], replacing mixed constraints by control constraints.

Recently, the same problem, as in the present paper, was analyzed by Bonnans and Osmolovskii in [2] under the hypothesis of positive linear independence (*i.e.*, linear independence with nonnegative coefficients) of gradients (PLIG) of active control constraints. The no-gap second order conditions of Pontryagin and bounded

Keywords and phrases. Pontryagin's principle, critical cone, quadratic form, second order sufficient condition, quadratic growth, Hoffman's error bound.

[1] Systems Research Institute, ul. Newelska 6, 01-447 Warszawa, Poland.
[2] Politechnika Radomska, ul. Malczewskiego 20A, 26-600 Radom, Poland.
[3] University of Natural Sciences and Humanities in Siedlce, ul. 3 Maja 54, 08-110 Siedlce, Poland. osmolovski@uph.edu.pl

Article published by EDP Sciences
© EDP Sciences, SMAI 2011

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

strong minima, similar to conditions [8], were obtained for this problem. Although the assumption of PLIG is weaker than the assumption of LIG, the results, obtained in [2], cannot be considered as a generalization of results [8], since in [2] we additionally assumed the Mangasarian-Fromovitz type condition for endpoint constraints and the so-called "restoration property" for these constraints. Hopefully, it would be possible to prove the same results, as in [2], under the PLIG hypothesis, but without two mentioned additional assumptions. The present publication could be an important step in this direction.

The paper is organized as follows. Section 2 sets the problem, recalls the concepts of bounded strong, Pontryagin and weak minima and gives the formulations of the first order necessary conditions of Pontryagin and weak minima, namely Pontryagin's principle and local Pontryagin's principle. In Section 3, for the cost functional, we define the "bounded strong growth condition of the order γ", where γ is a continuous functional (in the space of variations $\delta w = (\delta u, \delta y)$) equal to $\|\delta y\|_\infty^2 + \int_0^T |\delta u|^2 \, \mathrm{d}t$ in some neighborhood of zero and positive outside of this neighborhood. This condition implies a bounded strong minimum. We also define the violation function of the optimal control problem and the notion of the bounded strong γ-sufficiency (as the bounded strong growth condition of the order γ for the violation function). The latter implies the bounded strong growth of the order γ for the cost functional (and hence a bounded strong minimum). In Section 4, for the reference point $w = (x, u)$, we introduce the quadratic form Ω (the second variation of the Lagrange function) and the critical cone $\mathcal{K}$. We formulate a sufficient condition for a bounded strong γ-sufficiency in terms of Ω and $\mathcal{K}$. It should be noted that this condition implies the quadratic growth of the Hamiltonian, from which the continuity of the optimal control follows [2]. Sections 5–11 are devoted to the proofs. In Section 5 we define the set of Pontryagin's sequences Π related to the notion of Pontryagin minimum and the so-called basic constant C_γ for Π. We prove that the inequality $C_\gamma > 0$ implies the bounded strong γ-sufficiency. Hence the positivity of any lower bound for C_γ also implies the bounded strong γ-sufficiency. In the next sections, we obtain several successive lower bounds for C_γ, pursuing the goal to find the lower bound defined in the most simple way. In Section 7 we introduce a concept of support of the critical cone and show that the first order approximations of the endpoint constraints possesses Hoffman's error bound on the support. This fact plays a crucial role in the proof of sufficient optimality conditions and allows us to obtain in Section 9 (which completes the proof) a lower bound for C_γ formulated in terms of the quadratic form Ω and the critical cone $\mathcal{K}$. The proofs of several basic lemmas are gathered in Section 10. In Section 11 we prove abstract lemmas (used earlier in Sect. 10) concerning compatibility of a system of linear inequalities on a convex cone and Hoffman's type [7] upper bounds for the distance from the origin to the set of solutions of such system on the cone. We also discuss an abstract notion of support of critical cone.

The paper is self-contained and formally does not use results of other papers. But it should be noted that the main idea of the proof (of sufficient optimality conditions), related to the role of the basic constant C_γ, first appeared in the abstract theory of higher order conditions published in [8,14–16]. We use notation introduced in these papers.

2. Pontryagin and bounded strong minima. First order necessary conditions

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

$$\dot{y}(t) = f(u(t), y(t)) \quad \text{for a.a. } t \in [0, T], \tag{2.1}$$

$$u(t) \in U, \quad \text{for a.a. } t \in [0, T], \tag{2.2}$$

$$\phi_i(y(0), y(T)) \le 0, \ i = 1, \dots, r_1, \tag{2.3}$$

$$\phi_i(y(0), y(T)) = 0, \ i = r_1 + 1, \dots, r, \tag{2.4}$$

$$J(u, y) := \phi_0(y(0), y(T)) \to \min, \tag{2.5}$$

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 (where $W^{1,1}(0, T; \mathbb{R}^n)$ is the space of absolutely continuous functions from $[0, T]$ to $\mathbb{R}^n$). We

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

consider problem (2.1)–(2.5) 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 $\|w\|_{\mathcal{W}} := \|u\|_\infty + \|y\|_{1,1} = \operatorname{ess\,sup}_{[0,T]} |u(t)| + |y(0)| + \int_0^T |\dot{y}(t)|\,\mathrm{d}t$. Elements of $\mathcal{W}$ satisfying (2.1)–(2.4) are said to be feasible. The set of feasible points is denoted by $F(P)$. We shall use abbreviations $y_0 := y(0)$, $y_T := y(T)$, $\eta := (y_0, y_T)$.

It is well known that any control problem with a cost functional in the integral form $J = \int_0^T \mathcal{F}(u, y)\,\mathrm{d}t$ can be represented in the endpoint form by introducing a new state variable z defined by the state equation $\dot{z} = \mathcal{F}(u, y)$, $z(0) = 0$. 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 [17], p. 290, is as follows. The state variable y_i, *i.e.*, the ith component of the state vector y is called *unessential* if the function f does not depend on y_i and the functions ϕ_j, $j = 0, 1, \ldots, r$ 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. We introduce two concepts of minimum.

Definition 2.1. Following [8], p. 156, and [17], p. 291, we say that $w^0 = (u^0, y^0) \in F(P)$ is a *bounded strong minimum* if $J(w^0) \le J(w_k)$ 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 $y_k(0) \to y^0(0)$.

Definition 2.2. We say that $w^0 \in F(P)$ is a *Pontryagin minimum* (see [8], p. 156, and [17], pp. 2–3) if $J(w^0) \le J(w_k)$ 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 $\|u\|_1 = \int_0^T |u(t)|\,\mathrm{d}t$.

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

The *strict bounded strong* (*Pontryagin*) *minimum* is defined in a similar way. The only two distinctions are that the nonstrict inequality $J(w^0) \le J(w_k)$ (in Defs. 2.1 and 2.2) is replaced by the corresponding strict inequality and the sequence $w_k \in F(P)$ is assumed to contain no terms equal to w^0.

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

$$\varphi^\mu(y_0, y_T) = \varphi(y_0, y_T, \mu) := \sum_{i=0}^{r} \mu_i \phi_i(y_0, y_T), \tag{2.6}$$

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

$$H(u, y, p) = pf(u, y). \tag{2.7}$$

Let $W^{1,\infty}(0, T, \mathbb{R}^{n*})$ denote the space of Lipschitz continuous functions from $[0, T]$ to $\mathbb{R}^{n*}$. By the *costate associated with* $\mu \in \mathbb{R}^{(r+1)*}$ we understand the solution $p = p^\mu \in W^{1,\infty}(0, T, \mathbb{R}^{n*})$ (whenever it exists) of

$$\begin{aligned} -\dot{p}(t) &= H_y(u(t), y(t), p(t)), \quad \text{a.a. } t \in [0, T]; \\ p(0) &= -\varphi^\mu_{y_0}(y(0), y(T)); \quad p(T) = \varphi^\mu_{y_T}(y(0), y(T)). \end{aligned} \tag{2.8}$$

Definition 2.3. We say that $w = (u, y) \in F(P)$ satisfies Pontryagin'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 (2.8) holds and

$$\mu_i \ge 0, \quad i = 0, \ldots, r_1, \quad \mu_i \phi_i(y(0), y(T)) = 0, \quad i = 1, \ldots, r_1, \tag{2.9}$$

$$H(u(t), y(t), p(t)) \le H(v, y(t), p(t)), \quad \text{for all } v \in U, \text{ a.a. } t \in (0, T). \tag{2.10}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

The following theorem holds (see, *e.g.*, [13], pp. 17–19, and [17], pp. 24–25, 150).

Theorem 2.4. *A Pontryagin minimum satisfies Pontryagin's principle.*

In the sequel, we assume that the set U is given in the form

$$U = \{u \in \mathbb{R}^m \mid g(u) \leq 0\}, \tag{2.11}$$

where $g : \mathbb{R}^m \to \mathbb{R}^q$ is a C^2 mapping. In other words, the control constraints are defined by

$$g_j(u(t)) \leq 0, \quad \text{for a.a.} \quad t \in [0,T], \quad j = 1, \ldots, q. \tag{2.12}$$

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 $I_g(u) = \big\{i \in \{1, \ldots, q\} \mid g_i(u) = 0\big\}$ is the set of active indices.

Let us recall the 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

$$\overline{H}(u, y, p, a) = H(u, y, p) + ag(u). \tag{2.13}$$

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

$$\begin{aligned}
&\mu_i \geq 0,\ i = 0, \ldots, r_1,\ \mu_i \phi_i(\eta) = 0,\ i = 1, \ldots, r_1,\ |\mu| = 1,\\
&a(t) \geq 0,\ a(t)g(u(t)) = 0,\ \text{a.a. } t \in (0,T),\\
&-\dot{p}(t) = H_y(w(t), p(t)),\ \text{a.a. } t \in (0,T),\\
&p(0) = -\varphi^\mu_{y_0}(\eta),\ p(T) = \varphi^\mu_{y_T}(\eta),\\
&\overline{H}_u(w(t), p(t), a(t)) = 0,\ \text{a.a. } t \in (0,T),
\end{aligned} \tag{2.14}$$

where $\eta = (y(0), y(T))$. The following result is well-known (see, *e.g.*, [6], pp. 422–429, [13], pp. 148–149, and [17], p. 13).

Theorem 2.5. *Let w be a weak minimum. Then the set Λ_0 is nonempty and bounded. Moreover, the projector $(\mu, p, a) \to \mu$ is injective on Λ_0, and hence Λ_0 is a finite dimensional compact set.*

Denote by M_0 the set of all $\lambda = (\mu, p, a) \in \Lambda_0$ such that inequality (2.10) 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.

3. 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) which corresponds to the variation δw. 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.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Definition 3.1. 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

$$\gamma(\delta w) = \|\delta y\|_\infty^2 + \int_0^T \Gamma(\delta u(t))\,\mathrm{d}t. \tag{3.1}$$

Following [8], p. 123, we call the functional $\gamma : \mathcal{W} \to \mathbb{R}$ the *order* (of minimum).

Further, following [8], p. 112, we define the *violation function*

$$\sigma(\delta w) = (\delta\phi_0)_+ + \sum_{i=1}^{r_1} \phi_i(\eta + \delta\eta)_+ + \sum_{i=r_1+1}^{r} |\phi_i(\eta + \delta\eta)| + \|\delta\dot{y} - \delta f\|_1, \tag{3.2}$$

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

Definition 3.2. We say that $\{\delta w_k\}$ is a *bounded strong sequence* if

$$\sup_k \|\delta u_k\|_\infty < \infty, \quad |\delta y_k(0)| + \|\delta\underline{y}_k\|_\infty \to 0 \quad (k \to \infty). \tag{3.3}$$

Denote by S the set of all bounded strong sequences satisfying:

$$\begin{aligned} &\text{(a) } u(t) + \delta u_k(t) \in U \text{ for a.a. } t \in [0, T] \text{ for all } k;\\ &\text{(b) } \sigma(\delta w_k) \to 0 \ (k \to \infty). \end{aligned} \tag{3.4}$$

Definition 3.3. We say that a *bounded strong γ-sufficiency* holds (at a point $w \in F(P)$) (*cf.* [8], p. 126) 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 $\sigma(\delta w) \geq C\gamma(\delta w)$.

Definition 3.4. We say that a *bounded strong γ-growth condition holds for the cost function J* (at a point $w \in F(P)$) if there exists $C > 0$ such that, for any sequence $\{\delta w_k\} \in S$ satisfying the condition $w + \delta w_k \in F(P)$ for all k, we have $\delta_k J \geq C\gamma(\delta w_k)$ for all sufficiently large k, where $\delta_k J = J(\eta + \delta\eta_k) - J(\eta)$.

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

Theorem 3.5. *For any order function $\Gamma(u)$:*
(a) a bounded strong γ-sufficiency implies a bounded strong γ-growth condition for the cost function; and
(b) a bounded strong γ-growth condition for the cost function implies a strict bounded strong minimum.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Proof. (*a*) Assume that a bounded strong γ-sufficiency holds, *i.e.*, 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. Then for any sequence $\{\delta w_k\} \in S$ satisfying the condition $w + \delta w_k \in F(P)$ for all k, we have $\sigma(\delta w_k) = (\delta_k J)_+ \geq C\gamma(\delta w_k)$ for all sufficiently large k. The latter means that the γ-growth condition for the cost function holds at w.

(*b*) Now assume that w is not a point of a strict bounded strong minimum, *i.e.*, there exists a sequence $w_k \in F(P)$, bounded in $\mathcal{W}$, such that $\underline{y}_k \to \underline{y}$ uniformly, $y_k(0) \to y(0)$, $J(w_k) \leq J(w)$ and $w_k \neq w$ for all k. Set $\delta w_k = w_k - w$ for all k. Then, $\delta w_k \neq 0$, $\sigma(\delta w_k) = 0$, $u(t) + \delta u_k(t) \in U$ for a.a. $t \in [0, T]$, for all k, and conditions (3.3) hold. Consequently, $\{\delta w_k\} \in S$, $\gamma(\delta w_k) > 0$ and $\delta_k J \leq 0$ for all k. This implies that a bounded strong γ-growth condition for the cost function does not hold at w. □

Corollary 3.6. *For any order function $\Gamma(u)$, a bounded strong γ-sufficiency implies a strict bounded strong minimum.*

Set (*cf.* [8], p. 126)

$$C_\gamma(\sigma, S) := \inf_{\{\delta w_k\} \in S} \left(\liminf_k \frac{\sigma(\delta w_k)}{\gamma(\delta w_k)} \right), \tag{3.5}$$

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

Proposition 3.7. *The inequality $C_\gamma(\sigma, S) > 0$ is equivalent to the bounded strong γ-sufficiency.*

Our goal is to obtain conditions which guarantee the inequality $C_\gamma(\sigma, S) > 0$. To this end, we will estimate $C_\gamma(\sigma, S)$ from below.

In order to formulate second order sufficient conditions for a bounded strong γ-sufficiency we shall need the following strengthening of condition (2.10) of the maximum principle.

Definition 3.8. Let $\lambda = (\mu, p, a) \in M_0$. We say that the function $H(v, y(t), p(t))$ satisfies the *growth condition of the order* Γ if there exists $C > 0$ such that

$$\begin{aligned} &H(v, y(t), p(t)) - H(u(t), y(t), p(t)) \geq C\Gamma(v - u(t)) \\ &\text{for all } v \in U \text{ and for a.a. } t \in [0, T]. \end{aligned} \tag{3.6}$$

For a given $C > 0$, denote by $M(C\Gamma)$ the set of all $\lambda \in M_0$ such that condition (3.6) holds.

Remark 3.9. It was shown in [19], pp. 372–374, and [20], pp. 275–276, that a bounded strong γ-sufficiency implies the existence of $C > 0$ such that $M(C\Gamma) \neq \emptyset$. The latter implies the continuity of the control $u(t)$ [2], Lemma 2.7. The case of discontinuous (piecewise continuous) control $u(t)$ corresponds to a certain finer order than (3.1). The no-gap conditions of this order were obtained in [19,20], and their proofs will be published elsewhere. The proofs of sufficient optimality conditions for a piecewise continuous control appeared in [22]. Also note that condition (3.6) is never fulfilled in the cases of singular or bang-bang controls, in problems linear in control. These cases were studied, for example, in [4,5,17]. Finally, note that in this paper we do not discuss an important question of a characterization of condition (3.6) in terms of (strengthened) Legendre-type conditions. This will be the subject of another paper.

Remark 3.10. From Theorem 2.5 it follows that $M(C\Gamma)$ is a finite dimensional compact set.

4. Second order sufficient conditions

A direction (variation) $\delta w = (\delta u, \delta y) \in \mathcal{W}$ is said to be *critical* (*cf.* [6], Sect. 2) at a point $w \in F(P)$ if the following relations hold

$$\phi_i'(\eta)\delta\eta \leq 0,\ i \in I_\phi(\eta) \cup \{0\};\ \phi_j'(\eta)\delta\eta = 0,\ j = r_1 + 1, \ldots, r, \tag{4.1}$$

$$\delta\dot{y} = f'(w)\delta w, \tag{4.2}$$

$$(g_j'(u)\delta u)\chi_{\{g_j(u)=0\}} \leq 0,\ \text{a.a. } t \in [0, T],\ j = 1, \ldots, q, \tag{4.3}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

where $I_\phi(\eta) = \{i \in \{1,\dots,r_1\} \mid \phi_i(\eta) = 0\}$ is the set of active indices, $\chi_{\{g_j(u)=0\}}$ is the characteristic function of the set $\{t \in [0,T] \mid g_j(u(t)) = 0\}$, $j = 1,\dots,q$, $\eta = (y(0), y(T))$, 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 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 w) by the relation

$$\Omega(\delta w, \lambda) := \frac{1}{2}\langle \varphi_{\eta\eta}(\eta,\mu)\delta\eta, \delta\eta\rangle + \frac{1}{2}\int_0^T \langle \overline{H}_{ww}(w,p,a)\delta w, \delta w\rangle \,\mathrm{d}t, \tag{4.4}$$

where $\varphi_{\eta\eta}(\eta,\mu)$ is the second partial derivative (w.r.t. η) of the function $\varphi(\eta,\mu)$ defined by (2.6) and $\overline{H}_{ww}(w,p,a)$ is the second partial derivative (w.r.t. w) of the function $\overline{H}(w,p,a)$ defined by (2.13). Obviously, $\Omega(\delta w, \lambda)$ is quadratic in δw and linear in λ. For any $C \geq 0$ such that $M(C\Gamma) \neq \emptyset$, set

$$\Omega_{M(C\Gamma)}(\delta w) := \max_{\lambda \in M(C\Gamma)} \Omega(\delta w, \lambda). \tag{4.5}$$

In (4.5), the maximum exists, since $M(C\Gamma)$ is a finite dimensional compact set. Our main result is the following.

Theorem 4.1. *Let $w = (u, y)$ be a feasible point of problem (2.1)–(2.5) (with the set U defined by (2.11)), where $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$, $\phi_i : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ $(i = 0,\dots,r)$, and $g : \mathbb{R}^m \to \mathbb{R}^q$ are C^2 mappings. Assume that the mapping g satisfies the qualification hypothesis of linear independence (see Sect. 2). Let there exist an order function Γ and a number $C > 0$ such that the set $M(C\Gamma)$ is nonempty and let there exist a number $C_{\mathcal{K}} > 0$ such that*

$$\Omega_{M(C\Gamma)}(\delta w) \geq C_{\mathcal{K}} \left(|\delta y(0)|^2 + \int_0^T |\delta u|^2 \,\mathrm{d}t \right) \quad \text{for all} \quad \delta w \in \mathcal{K}. \tag{4.6}$$

Then, for the higher order γ defined by (3.1), a bounded strong γ-sufficiency holds at w.

Remark 4.2. It was proved in [19], Chapter 5, and [20], Chapter 4, that if w is a Pontryagin minimum, then the set M_0 is nonempty and $\max_{\lambda \in M_0} \Omega(\delta w, \lambda) \geq 0$ for all $\delta w \in \mathcal{K}$. Thus, the sufficient condition given by Theorem 4.1 is a natural strengthening of the necessary condition. Also, it was proved in [19], pp. 323–325, and [20], Supplement, pp. 154–156, that if, in the definition of critical cone $\mathcal{K}$, the condition $\delta w = (\delta u, \delta y) \in \mathcal{U} \times \mathcal{Y}$ is replaced by $\delta w = (\delta u, \delta y) \in \mathcal{U}_2 \times \mathcal{Y}_2$, where $\mathcal{U}_2 = L^2(0,T,\mathbb{R}^m)$ and $\mathcal{Y}_2 = W^{1,2}(0,T,\mathbb{R}^n)$, the new condition of the form (4.6) is equivalent to the original one (this equivalence is not obvious).

Remark 4.3. In the literature, there are sufficient conditions for strong local optimality that require (i) one tuple of Lagrange multipliers at which the corresponding quadratic form is positive definite on a larger critical cone; (ii) the strengthened Legendre-Clebsch condition; (iii) the existence of a "unique" control maximizing the Hamiltonian (the strengthened condition of Pontryagin's principle) which means that $\operatorname{argmin}_u H(x(t), u, \psi(t))$ is a singleton equal to $\{u(t)\}$ a.a., and (iv) a certain growth condition (w.r.t. u) of the Hamiltonian at the infinity (*cf.*, *e.g.*, [21], Sect. 10.4, Thm. 10.5). Since (i) and (ii) form a pair of sufficient conditions for weak local optimality, (iii) and (iv) complement them to a set of conditions being sufficient for strong local optimality. Moreover, (i) can be equivalently represented in the form of the strengthened Jacobi-type condition (*cf.*, *e.g.*, [23]) or in the form of the existence of a solution to the corresponding Riccati equation (*cf.*, *e.g.*, [24]). Let us briefly compare Theorem 4.1 of the present paper with a known result by Vera Zeidan for strong local minimum given in [24], p. 1308, Theorem 6.1. In [24], the optimal control problem, referred as problem (C), has the form: minimize $J(x,u) := l(x(b)) + \int_a^b g(t, x(t)u(t))\,\mathrm{d}t$ subject to $\dot{x}(t) = f(t, x(t), u(t))$, $x(a) = A$, $\psi(x(b)) = 0$, and $G_i(t, x(t), u(t)) \leq 0$, $i = 1,\dots,k$. So, in comparison with (P), problem (C) has mixed state-control inequality constraints (instead of pure control inequality constraints), but has no endpoint inequality constraints. Let $(\hat{x}, \hat{u})$ be an extremal such that $\hat{u} \in L^\infty$. It is assumed that the gradients $G_{iu}(t,x,u)$ of active constraints are uniformly linearly independent along the trajectory $(t, \hat{x}(t), \hat{u}(t))$ and moreover, there exists a tuple of Lagrange multipliers with $\lambda_0 \neq 0$, where λ_0 is the Lagrange multiplier of J. Let $q_i(t)$ be the Lagrange multiplier of the constraint $G_i(t,x,u) \leq 0$. For any $\gamma \geq 0$, set $J_\gamma(t) = \{i \in \{1,\dots,k\} : q_i(t) > \gamma\}$. In [24],

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

it is assumed that there exists $\gamma > 0$ such that $J_\gamma(t) = J_0(t)$ a.e. This is a very strong assumption removing certain difficulties in the proofs. Consider the case, where the control $\hat{u}(t)$ is continuous and f, g, and G are independent of t. In this case, all $q_i(t)$ are continuous and the assumption $J_\gamma(t) \equiv J_0(t)$ ($\gamma > 0$) means that for any $i = 1, \dots, k$ we have: (a) $q_i(t) > 0$ for all $t \in [a, b]$ or (b) $q_i(t) \equiv 0$ on $[a, b]$. In [24], the critical cone (corresponding to the Riccati equation) is enlarged up to a subspace in such a way that among all $G_i \le 0$ only the constraints of type (a) are taken into account in the sufficient conditions, while the constraints of type (b) are simply ignored. Therefore, in the case of continuous control, sufficient conditions of weak minimum in [24] turned out to be much stronger than those in the present paper. Now let us turn to sufficient conditions for strong minimum. We can prove that the strengthened Legendre-Clebsch condition (ii) together with the strengthened condition of Pontryagin's principle (iii) guarantee the existence of order function Γ (Def. 3.1) such that the growth condition of order Γ holds for the Hamiltonian (Def. 3.8). Hence (according to Thm. 4.1), (iii) complements (i) and (ii) to the set of sufficient conditions for bounded strong minimum. To guarantee a strong minimum, one may add some kind of growth condition (iv) for the Hamiltonian at the infinity (or assume that the inequality $g(u) \le 0$ define a compact set in $\mathbb{R}^m$). In [24], the following condition complements (i) and (ii) to the set of sufficient conditions for strong minimum (and hence play the same role as pair (iii) and (iv)): there exists a mapping $\tilde{u}(t, x, p)$ (defined in some neighborhood of the trajectory $(\hat{x}(t), \hat{p}(t))$, where $\hat{p}(t)$ is the adjoint variable) such that (a) $\tilde{u}(t, \cdot, \cdot, \cdot)$ is continuous uniformly in t, (b) $\tilde{u}(t, \hat{x}(t), \hat{p}(t)) = \hat{u}(t)$ a.e., and (c) $\tilde{u}(t, x, p) \in \operatorname{argmin}_u \{pf(t, x, u) + g(t, x, u) : G(t, x, u) \le 0\}$ (where p corresponds to the minimum principle). Obviously, the latter condition strengthens the Pontryagin principle, but a detailed comparison of this condition with (iii) and (iv) is not so simple and could be done elsewhere.

The proof of Theorem 4.1 will require preparatory sections, namely, Sections 5–8. This proof will be completed in Section 9. The proofs of several important lemmas will be gathered in Section 10; one of these lemmas (namely, Lem. 7.2) will be based on the results of Section 11 concerning Hoffman's error bounds.

5. Bounded strong and Pontryagin's sequences

Let us introduce a set of Pontryagin's sequences Π related to the notion of Pontryagin minimum. In what follows, we omit the subscript k in the notation of sequences.

Definition 5.1. We 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 $\|\delta y\|_{1,1} \to 0$. Denote by Π the set of all Pontryagin's sequences.

5.1. Passage to Pontryagin's sequences, the basic constant

Given any $\lambda = (\mu, p, a) \in \Lambda_0$ and $\delta w \in \mathcal{W}$, we define the following Lagrange function for the cost function (2.5) and constraints (2.1), (2.3) and (2.4):

$$\Psi(\delta w, \lambda) := \delta\varphi^\mu - \int_0^T p(\delta\dot{y} - \delta f)\,\mathrm{d}t = \delta\varphi^\mu - \int_0^T (p\delta\dot{y} - \delta H)\,\mathrm{d}t, \tag{5.1}$$

where $\delta\varphi^\mu = \varphi^\mu(\eta + \delta\eta) - \varphi^\mu(\eta)$ (recall that φ^μ was defined in (2.6)) and $\delta H = p\delta f$. Below we shall need another representation of Ψ. Using (2.8), we get

$$\begin{aligned} \int_0^T p\delta\dot{y}\,\mathrm{d}t &= p\delta y\,|_0^T - \int_0^T \dot{p}\delta y\,\mathrm{d}t \\ &= \varphi^\mu_{y_0}(\eta)\delta y(0) + \varphi^\mu_{y_T}(\eta)\delta y(T) + \int_0^T H_y(w, p)\delta y\,\mathrm{d}t \end{aligned} \tag{5.2}$$

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

$$\Psi(\delta w, \lambda) = \delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T (\delta H - H_y(w, p)\delta y)\,\mathrm{d}t, \quad \forall \lambda \in \Lambda_0. \tag{5.3}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Set

$$\Psi_{\Lambda_0}(\delta w) = \max_{\lambda \in \Lambda_0} \Psi(\delta w, \lambda). \tag{5.4}$$

Since Λ_0 is a bounded set, using (3.2) and (5.1) it is easy to show that there exists $k_0 > 0$ such that

$$\Psi_{\Lambda_0}(\delta w) \le k_0 \sigma(\delta w). \tag{5.5}$$

In what follows we assume that, *for a given order function* Γ *and a number* $C > 0$, *the set* $M(C\Gamma)$ *is nonempty.* Using estimate (5.5), let us show that any sequence from S is a Pontryagin's sequence.

Proposition 5.2. *If* $\{\delta w\} \in S$, *then* $\|\delta u\|_1 \to 0$, $\|\delta y\|_{1,1} \to 0$, *and* $\gamma(\delta w) \to 0$.

The proof of this proposition will be based on two lemmas.

Lemma 5.3. *If* $\{\delta w\} \in S$ *and* $(\mu, p, a) \in \Lambda_0$, *then* $\left(\int_0^T \delta H \, \mathrm{d}t\right)_+ \to 0$.

Proof. According to (5.3)–(5.5), $\delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T (\delta H - H_y(w,p)\delta y)\,\mathrm{d}t \le k_0\sigma(\delta w)$. Since $\|\underline{\delta y}\|_\infty \to 0$, the condition $\sigma(\delta w) \to 0$ implies $\left(\int_0^T \delta H \, \mathrm{d}t\right)_+ \to 0$. □

Lemma 5.4. *If* $\{\delta w\} \in S$, *then* $\int_0^T \Gamma(\delta u)\,\mathrm{d}t \to 0$ *and* $\int_0^T |\delta u|^2\,\mathrm{d}t \to 0$.

Proof. Let $\{\delta w\} \in S$ and $(\mu, p, a) \in M(C\Gamma)$ $(C > 0)$. We have $\delta H := H(w + \delta w, p) - H(w, p) = \bar\delta_y H + \delta_u H$, 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 Lem. 5.3) implies $\left(\int_0^T \delta_u H \, \mathrm{d}t\right)_+ \to 0$. But $\delta_u H \ge C\Gamma(\delta u)$, since $(\mu, p, a) \in M(C\Gamma)$ and $u(t) + \delta u(t) \in U$ a.e. Consequently, $\int_0^T \Gamma(\delta u)\,\mathrm{d}t \to 0$. This implies that $\int_0^T |\delta u|^2\,\mathrm{d}t \to 0$. □

Proof of Proposition 5.2. Let $\{\delta w\} \in S$. Then by Cauchy-Schwartz inequality and by Lemma 5.4 we have $\|\delta u\|_1 \le \sqrt{T}\|\delta u\|_2 \to 0$. 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 $\|\delta y\|_\infty \to 0$. Also, by Lemma 5.4, $\int_0^T \Gamma(\delta u)\,\mathrm{d}t \to 0$. Consequently, $\gamma(\delta w) \to 0$. □

Corollary 5.5. *We have:* $S \subset \Pi$ *and moreover,*

$$S = \{\{\delta w\} \in \Pi \mid g(u(t) + \delta u(t)) \le 0\}, \tag{5.6}$$

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

Proof. The inclusion $S \subset \Pi$ follows from Proposition 5.2. Further, for any $\{\delta w\} \in \Pi$, we obviously have $\sigma(\delta w) \to 0$. Therefore, $\{\{\delta w\} \in \Pi \mid g(u(t) + \delta u(t)) \le 0\} \subset S$. This and the inclusion $S \subset \Pi$ implies equality (5.6), since any sequence from S satisfies $g(u(t) + \delta u(t)) \le 0$. □

Set

$$\Pi_{\sigma\gamma} = \{\{\delta w\} \in \Pi \mid g(u(t) + \delta u(t)) \le 0,\ \sigma(\delta w) \le O(\gamma(\delta w))\}. \tag{5.7}$$

From equality (5.6) and definition (5.7) we easily deduce that

$$C_\gamma(\sigma, S) = C_\gamma(\sigma, \Pi_{\sigma\gamma}), \tag{5.8}$$

where the definition of $C_\gamma(\sigma, \Pi_{\sigma\gamma})$ is similar to the definition (3.5) of $C_\gamma(\sigma, S)$. Using (5.5), we get

$$C_\gamma(\sigma, \Pi_{\sigma\gamma}) \ge k_0^{-1} C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}), \tag{5.9}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

where, by definition

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) := \inf_{\{\delta w\}\in\Pi_{\sigma\gamma}} \left(\liminf \frac{\Psi_{\Lambda_0}(\delta w)}{\gamma(\delta w)} \right) \tag{5.10}$$

(the lower bound in this formula is taken over the set of sequences from $\Pi_{\sigma\gamma}$ that do not vanish). We call $C_\gamma := C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma})$ the *basic constant* (on the set of Pontryagin's sequences). From (5.8) and (5.9) we obtain the following inequality

$$C_\gamma(\sigma, S) \geq k_0^{-1} C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}). \tag{5.11}$$

In the sequel, we shall estimate the basic constant from below. Under conditions of Theorem 4.1, we shall show that this constant is positive. Then in view of (5.11), Theorem 4.1 will follow from Proposition 3.7.

5.2. **Extension of the set $\Pi_{\sigma\gamma}$**

In the sequel, we shall use one more representation for Ψ. 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 (5.3) combined with the equalities $a\delta g + a\delta g_- = 0$ (where $\alpha_- = \max\{-\alpha, 0\} \geq 0$ and $g_- = (g_{1-}, \dots, g_{q-})$) and $\delta\overline{H} = \delta H + a\delta g$ implies

$$\Psi(\delta w, \lambda) = \delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T \left(\delta\overline{H} - H_y(w, p)\delta y \right) \mathrm{d}t + \int_0^T a\delta g_- \, \mathrm{d}t. \tag{5.12}$$

Assume in addition that $\{\delta w\} \in \Pi_{\sigma\gamma}$ and hence $\sigma(\delta w) \leq O(\gamma(\delta w))$. For given sequence, we deduce from (5.5) and (5.12) that

$$\max_{\lambda\in\Lambda_0} \left\{ \delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta + \int_0^T \left(\delta\overline{H} - H_y(w, p)\delta y \right) \mathrm{d}t + \int_0^T a\delta g_- \, \mathrm{d}t \right\} \leq O(\gamma), \tag{5.13}$$

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: $\int_0^T |\delta\overline{H} - H_y(w, p)\delta y| \, \mathrm{d}t \leq O(\gamma)$. Moreover, $|\delta\varphi^\mu - \varphi^\mu_\eta(\eta)\delta\eta| \leq O(\gamma)$ uniformly on Λ_0. Therefore, condition (5.13) implies

$$\max_{\lambda\in\Lambda_0} \int_0^T a\delta g_- \, \mathrm{d}t \leq O(\gamma). \tag{5.14}$$

This estimate is satisfied for any $\{\delta w\} \in \Pi_{\sigma\gamma}$. Hence we can rewrite (5.7) as

$$\Pi_{\sigma\gamma} = \left\{ \{\delta w\} \in \Pi \mid g(u + \delta u) \leq 0, \ \sigma \leq O(\gamma), \ \max_{\lambda\in\Lambda_0} \int_0^T a\delta g_- \, \mathrm{d}t \leq O(\gamma) \right\}. \tag{5.15}$$

Since the set Λ_0 is finite dimensional, so is its convex hull $\operatorname{co}\Lambda_0$. Hence the relative interior $\operatorname{ri}\operatorname{co}\Lambda_0$ is nonempty. The following proposition will allow us to simplify the last relation in (5.15).

Proposition 5.6. *Let $\hat\lambda = (\hat\mu, \hat p, \hat a) \in \operatorname{ri}\operatorname{co}\Lambda_0$. Then there exists $\hat C > 0$ such that, for any $\lambda = (\mu, p, a) \in \operatorname{co}\Lambda_0$, the following inequalities hold*

$$\mu_i \leq \hat C\hat\mu_i, \ i = 0, \dots, r_1; \quad a_j(t) \leq \hat C\hat a_j(t) \ \text{a.e. on } [0, T], \ j = 1, \dots, q. \tag{5.16}$$

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

$$\frac{1+\rho}{\rho}\hat\mu_i \geq \mu_i, \quad i = 0, \dots, r_1; \quad \frac{1+\rho}{\rho}\hat a_j(t) \geq a_j(t), \quad j = 1, \dots, q.$$

Thus, it suffices to set $\hat C = (1 + \rho)/\rho$. □

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Let us fix an element $\hat\lambda = (\hat\mu, \hat p, \hat a) \in \operatorname{ri}\operatorname{co}\Lambda_0$. It follows from Proposition 5.6, that condition (5.14) is equivalent to the condition $\int_0^T \hat a(\delta g)_-\,\mathrm{d}t \le O(\gamma)$. Then, from (5.15) it follows that

$$\Pi_{\sigma\gamma} = \left\{ \{\delta w\} \in \Pi \mid g(u+\delta u) \le 0,\ \sigma \le O(\gamma),\ \int_0^T \hat a \delta g_-\,\mathrm{d}t \le O(\gamma) \right\}. \tag{5.17}$$

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

$$\Pi_{o(\sqrt\gamma)} = \left\{ \{\delta w\} \in \Pi \mid g(u+\delta u) \le 0,\ \sigma \le o(\sqrt\gamma),\ \int_0^T \hat a \delta g_-\,\mathrm{d}t \le O(\gamma) \right\}. \tag{5.18}$$

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

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) \ge C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt\gamma)}), \tag{5.19}$$

where the r.h.s. of inequality (5.19) is defined similarly to the l.h.s. of this inequality. In what follows, we will estimate $C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt\gamma)})$ from below.

6. Local sequences

Our final goal is to obtain a lower bound of C_γ defined by the set of sequences of critical variations (satisfying (4.1)–(4.3)) and by the quadratic form Ω (defined in (4.4)). It will be done in several steps. In each step lower bound of C_γ is defined on a different set of sequences and a modification of the function Ψ is used. In this section, we shall make one more step in this direction, namely, in the definition of $C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt\gamma)})$ we shall replace Pontryagin's sequences by the sequences satisfying the condition $\|\delta w\|_{\mathcal W} \to 0$. After that the obtained lower bound of C_γ will be simplified.

6.1. Passage to the sequences of local variations

Set

$$S^{\mathrm{loc}} = \{\{\delta w\} \mid \|\delta w\|_{\mathcal W} := \|\delta u\|_\infty + \|\delta y\|_{1,1} \to 0\}. \tag{6.1}$$

Sequences from S^{loc} will be called *local*. Note that $S^{\mathrm{loc}} \subset \Pi$. Bellow, we shall pass to the set of sequences

$$S^{\mathrm{loc}}_{o(\sqrt\gamma)} := \Pi_{o(\sqrt\gamma)} \cap S^{\mathrm{loc}}. \tag{6.2}$$

In other words, $S^{\mathrm{loc}}_{o(\sqrt\gamma)}$ is the set of the sequences $\{\delta w\}$ satisfying the relations

$$\|\delta w\|_{\mathcal W} \to 0,\ g(u+\delta u) \le 0, \tag{6.3}$$

$$\sigma = o(\sqrt\gamma),\ \int_0^T \hat a(\delta g)_-\,\mathrm{d}t \le O(\gamma). \tag{6.4}$$

By (3.2), the relation $\sigma = o(\sqrt\gamma)$ is equivalent to the set of relations

$$\|\delta\dot y - f'(w)\delta w\|_1 = o(\sqrt\gamma), \tag{6.5}$$

$$\phi_i'(\eta)\delta\eta \le o(\sqrt\gamma),\ i \in I_\phi(\eta) \cup \{0\}, \tag{6.6}$$

$$|\phi_i'(\eta)\delta\eta| = o(\sqrt\gamma),\ i = r_1+1, \dots, r. \tag{6.7}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

The second relation in (6.4) is equivalent to

$$\int_0^T \hat{a}_j (g_j'(u)\delta u)_- \, \mathrm{d}t \le O(\gamma), \ j = 1, \dots, q. \tag{6.8}$$

Thus, the set $S^{\mathrm{loc}}_{o(\sqrt{\gamma})}$ can be characterizes by (6.3) and (6.5)–(6.8).

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

$$\Psi(\delta w, \lambda) = \Omega(\delta w, \lambda) + \int_0^T a\delta g_- \, \mathrm{d}t + o(\gamma(\delta w)) \tag{6.9}$$

uniformly on Λ_0, where Ω is defined by (4.4).

Proof. Formula (6.9) follows from (5.12) and the relations: $H_y = \overline{H}_y$ and $\overline{H}_u(w, p, a) = 0$ for all $\lambda \in \Lambda_0$. □

For any $C > 0$ such that the set $M(C\Gamma)$ is nonempty, we set

$$\Psi_{M(C\Gamma)}(\delta w) := \max_{\lambda \in M(C\Gamma)} \Psi(\delta w, \lambda). \tag{6.10}$$

Lemma 6.2. *For any $C > 0$ such that the set $M(C\Gamma)$ is nonempty, the following inequality holds*

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{o(\sqrt{\gamma})}) \ge \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right), C \right\}, \tag{6.11}$$

where the constant $C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right)$ is defined as in (5.10).

The proof is given in Section 10.1.

Recall that we have fixed $C > 0$ such that the set $M(C\Gamma)$ is nonempty. Lemma 6.2 along with inequality (5.19) implies

$$C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) \ge \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right), C \right\}. \tag{6.12}$$

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

6.2. **Passage to the set of sequences S_1**

Proposition 6.3. *Let $\lambda = (\mu, p, a) \in \Lambda_0$. Then the critical cone $\mathcal{K}$, defined by (4.1)–(4.3), has the following equivalent representation:*

$$\begin{aligned} &\phi_i'(\eta)\delta\eta \le 0, \ \mu_i \phi_i'(\eta)\delta\eta = 0, \ i \in I_\phi(\eta) \cup \{0\}, \\ &\phi_j'(\eta)\delta\eta = 0, \ j = r_1 + 1, \dots, r, \ \delta\dot{y} = f'(w)\delta w, \\ &(g_j'(u)\delta u)\chi_{\{g_j(u)=0\}} \le 0, \ a_j g_{ju}(u)\delta u = 0, \ j = 1, \dots, q. \end{aligned} \tag{6.13}$$

Proof. Indeed, from definition (2.14) of the set Λ_0 it easily follows that, for any $\lambda = (\mu, p, a) \in \Lambda_0$ and any $\delta w \in \mathcal{W}$, we have

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

Therefore, relations (4.1)–(4.3) imply that $\mu_i \phi_i'(\eta)\delta\eta = 0$, $i \in I_\phi(\eta) \cup \{0\}$, $a g'(u)\delta u = 0$. Thus, relations (6.13) follows from (4.1)–(4.3). *Vice versa*, relations (4.1)–(4.3) obviously follows from (6.13). □

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Define a new set S_1 of sequences $\{\delta w\} \subset \mathcal{W}$ by the relations

$$\|\delta u\|_\infty + \|\delta y\|_{1,1} \to 0,\ \sigma(\delta w) = o\left(\sqrt{\gamma(\delta w)}\right), \tag{6.14}$$

$$g(u) + g'(u)\delta u \le 0,\ g'_j(u)\delta u \chi_{\{\hat a_j \ge \varepsilon\}} = 0,\ j = 1, \dots, q,\ \varepsilon \to +0, \tag{6.15}$$

where $\chi_{\{\hat a_i \ge \varepsilon\}}(t)$ is the characteristic function of the set $\{t \mid \hat a_i(t) \ge \varepsilon\}$. The equalities in (6.15) 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 $j = 1, \dots, q$ and for all $k = 1, 2, \dots$ Set

$$\Phi^1_C(\delta w) := \max_{\lambda \in M(C\Gamma)} \left\{ \Omega(\delta w, \lambda) + \int_0^T a(g'(u)\delta u)_- \, \mathrm{d}t \right\}. \tag{6.16}$$

Lemma 6.4. *The following inequality holds*

$$C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right) \ge C_\gamma \left(\Phi^1_C, S_1 \right), \tag{6.17}$$

where

$$C_\gamma\left(\Phi^1_C, S_1\right) := \inf_{\{\delta w\} \in S_1} \left(\liminf \frac{\Phi^1_C(\delta w)}{\gamma(\delta w)} \right). \tag{6.18}$$

The proof is given in Section 10.2.

7. Support of the critical cone

Here we define a notion of support of the critical cone and formulate an important property of the support (Lem. 7.2) which will play a crucial role at the end of the proof of Theorem 4.1 in Section 9. This property will mean that the first order approximations of the endpoint functionals possess the so-called "Hoffman's error bound" [7] on the support.

Set

$$\mathcal{W}_0 = \{\delta w = (\delta u, \delta y) \in \mathcal{W} \mid \delta \dot y = f_w(w)\delta w\}. \tag{7.1}$$

Consider two sets of linear functionals

$$l_i : \delta w \in \mathcal{W}_0 \ \to \phi'_i(\eta)\delta\eta, \quad i \in I_\phi(\eta) \cup \{0\}, \tag{7.2}$$

$$l_i : \delta w \in \mathcal{W}_0 \ \to \phi'_i(\eta)\delta\eta, \quad i = r_1 + 1, \dots, r, \tag{7.3}$$

where $\delta w = (\delta u, \delta y)$, $\delta\eta = (\delta y(0), \delta y(T))$. Let Q_0 be the cone generated by functionals (7.2), and let Q_1 be the subspace generated by functionals (7.3). Set

$$Q := Q_0 + Q_1. \tag{7.4}$$

Then Q is a convex and finitely generated cone. Note that $w^* \in Q$ iff there is a vector $\mu = (\mu_0, \dots, \mu_r) \in \mathbb{R}^{(r+1)*}$ such that $\mu_i \ge 0$, $i = 0, \dots, r_1$, $\mu_i\phi_i(\eta) = 0$, $i = 1, \dots, r_1$ and $\langle w^*, \delta w\rangle = \varphi^\mu_\eta(\eta)\delta\eta$ for all $\delta w \in \mathcal{W}_0$, where $\varphi^\mu = \sum_{i=0}^r \mu_i\phi_i$, as in (2.6). Obviously, $\langle w^*, \delta w\rangle \le 0$ for any $w^* \in Q$ and any $\delta w \in \mathcal{K}$. Moreover, let $p = p^\mu$ be the solution to the adjoint equation $-\dot p = pf_y(w)$, $p(T) = \varphi^\mu_{y_T}(\eta)$. As it is well-known, the functional $\varphi^\mu_\eta(\eta)\delta\eta$ has the following representation on the subspace $\mathcal{W}_0$: $\left(\varphi^\mu_{y_0}(\eta) + p(0)\right)\delta y_0 + \int_0^T pf_u(w)\delta u \, \mathrm{d}t$. This representation is "pure integral" iff $\varphi^\mu_{y_0}(\eta) + p(0) = 0$.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Let us define the set

$$\mathcal{R}_0 = \{\delta w = (\delta u, \delta y) \in \mathcal{W}_0 \mid \begin{array}{l} g_j'(u)\delta u \chi_{\{g_j(u)=0\}} \le 0, \\ \hat{a}_j g_j'(u)\delta u = 0,\ j = 1, \dots, q\}, \end{array} \tag{7.5}$$

where $\hat{a}$ is given in Proposition 5.6. Obviously, $\mathcal{R}_0$ is a closed convex cone, and by Proposition 6.3, $\mathcal{K} \subset \mathcal{R}_0$. Let $w_1^* \in Q$ be a linear functional which has the following integral representation on the cone $\mathcal{R}_0$: $\langle w_1^*, \delta w\rangle = -\int_0^T a^1 g'(u)\delta u \,\mathrm{d}t \ \forall \delta w \in \mathcal{R}_0$, where $a^1 \in L^1(0,T;\mathbb{R}^{q*})$, $a^1 \ge 0$, and $a^1 g(u) = 0$. Clearly, $\langle w_1^*, \delta w\rangle \ge 0$ 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 $\mathcal{R}_1 = \{\delta w \in \mathcal{R}_0 \mid \langle w_1^*, \delta w\rangle = 0\}$. Conditions (7.5) along with $a^1 \ge 0$ and $a^1 g(u) = 0$ imply $\mathcal{R}_1 = \{\delta w \in \mathcal{R}_0 \mid a^1 g'(u)\delta u = 0\}$. Moreover, since $w_1^* \in Q$, we have $\langle w_1^*, \delta w\rangle \le 0$ for all $\delta w \in \mathcal{K}$. But $\mathcal{K} \subset \mathcal{R}_0$ and $w_1^* \in \mathcal{R}_0^*$, consequently $\langle w_1^*, \delta w\rangle = 0$ for all $\delta w \in \mathcal{K}$. This implies that $\mathcal{K} \subset \mathcal{R}_1$.

Similarly, let $w_2^* \in Q$ be a linear functional 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 $a^2 \in L^1(0,T,\mathbb{R}^{q*})$, $a^2 \ge 0$, $a^2 g(u) = 0$. Obviously, $\langle w_2^*, \delta w\rangle \ge 0 \ \forall \delta w \in \mathcal{R}_1$, *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 $\mathcal{R}_2 = \{\delta w \in \mathcal{R}_1 \mid a^2 g'(u)\delta u = 0\}$. From conditions $w_2^* \in Q$, $w_2^* \in \mathcal{R}_1^*$, and $\mathcal{K} \subset \mathcal{R}_1$ we deduce that $\langle w_2^*, \delta w\rangle = 0 \ \forall \delta w \in \mathcal{K}$. 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 $\mathcal{R}_0, \mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_s$ such that $\mathcal{K} \subset \mathcal{R}_s \subset \dots \subset \mathcal{R}_1 \subset \mathcal{R}_0$ and $w_k^*(\mathcal{R}_k) = \{0\}$, $w_k^*(\mathcal{R}_{k-1}) \ne \{0\}$, $k = 1, \dots, s$. Due to these properties, the functionals $w_1^*, w_2^*, \dots, w_s^*$ are linearly independent. Since the cone Q (containing these functionals) is finite generated, this process will be finished on some finite step s. Set $\mathcal{S} = \mathcal{R}_s$, $a^0 = \hat{a}$. Thus

$$\mathcal{S} = \{\delta w \in \mathcal{W}_0 \mid \delta u \in \mathcal{S}_u\}, \tag{7.6}$$

where

$$\begin{array}{c} \mathcal{S}_u := \{\delta u \in \mathcal{U} \mid (g_j'(u)\delta u)\chi_{\{g_j(u)=0\}} \le 0,\ j = 1, \dots, q, \\ a^i g'(u)\delta u = 0,\ i = 0, 1, \dots, s\}. \end{array} \tag{7.7}$$

We call $\mathcal{S}$ the *support* of the critical cone $\mathcal{K}$.

Remark 7.1. Without loss of generality we can assume that $a^i \in L^\infty(0,T;\mathbb{R}^m)$, $i = 0, 1, \dots, s$. Indeed, for $a^0 = \hat{a}$ this condition holds. Further, we can choose a^1 so that $a^1 \ge 0$, $a^1 g(u) = 0$, and $a_j^1(t) = 0$ a.e. on the set $\{t \mid a_j^0(t) > 0\}$, $j = 1, \dots, q$. Also we have $\langle w_1^*, \delta w\rangle := \varphi_\eta^{\mu^1}(\eta)\delta\eta = \int_0^T p^1 f_u(w)\delta u \,\mathrm{d}t = -\int_0^T a^1 g'(u)\delta u \,\mathrm{d}t$ for all $\delta w \in \mathcal{R}_0$, where p^1 is the solution to the adjoint equation $-\dot{p}^1 = p^1 f_y(w)$, $p^1(T) = \varphi_{y_T}^{\mu^1}(\eta)$. Using these conditions and the linear independence hypothesis for the gradients g_{ju}, we easily deduce that $a^1 \in L^\infty(0,T;\mathbb{R}^m)$. Similarly, we can choose $a^i \in L^\infty(0,T;\mathbb{R}^m)$ for $i > 1$.

The following important lemma will be used in Section 9.

Lemma 7.2. *There exists $N > 0$ such that for any $\delta w = (\delta u, \delta y) \in \mathcal{S}$ there exists $\bar{w} = (\bar{u}, \bar{y}) \in \mathcal{S}$ such that the following relations hold*

$$\phi_i'(\eta)(\delta\eta + \bar{\eta}) \le 0,\ i \in I_\phi \cup \{0\},\ \phi_j'(\eta)(\delta\eta + \bar{\eta}) = 0,\ j = r_1 + 1, \dots, r, \tag{7.8}$$

$$\|\bar{w}\|_{\mathcal{W}} \le N \left(\sum_{i \in I_\phi \cup \{0\}} (\phi_i'(\eta)\delta\eta)_+ + \sum_{j=r_1+1}^{r} |\phi_j'(\eta)\delta\eta| \right), \tag{7.9}$$

where $\delta\eta = (\delta y(0), \delta y(T))$, $\bar{\eta} = (\bar{y}(0), \bar{y}(T))$.

The proof is given in Section 10.3.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

8. Auxiliary assertions

Along with Lemma 7.2, the following three lemmas will be used in Section 9. Let $L^p_+(0,T;\mathbb{R})$ denote the cone of nonnegative functions from $L^p(0,T;\mathbb{R})$ $(1 \le p \le \infty)$, and let $\mathbb{R}_+$ be the set of nonnegative numbers.

Lemma 8.1. *Let $r \in L^\infty_+(0,T,\mathbb{R})$, set $\mathcal{M} = \{t \mid r(t) > 0\}$. Let $\{(v,\alpha)\}$ be a sequence in $L^2_+(0,T;\mathbb{R}) \times \mathbb{R}_+$ such that $\alpha \to 0$, $\|v\|_2 \le \alpha$, and $\int_0^T rv\,\mathrm{d}t = o(\alpha)$. Then there exists a sequence $\{B\}$ of subsets $B \subset \mathcal{M}$ such that $\operatorname{meas} B \to 0$ and for the sequence of sets $\{A\} := \{\mathcal{M} \setminus B\}$ we have $\|v\chi_A\|_\infty = o(\alpha)$, where χ_A is the characteristic function of the set A.*

The proof is given in Section 10.4.

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

$$\begin{aligned} &\|u'\|_\infty \to 0, \quad \operatorname{meas}\{t \mid u'(t) \ne 0\} \to 0, \quad g(u) + g'(u)u' \le 0, \\ &g'_j(u)u'\chi_{\{\hat a_j \ge \varepsilon\}} = 0,\ j = 1,\dots,q, \quad \varepsilon \to +0. \end{aligned} \tag{8.1}$$

The equalities in (8.1) mean that for the sequence $\{u'_k\}$ there exists a sequence $\varepsilon_k \to +0$ such that $g'_j(u(t))u'_k(t) = 0$ a.e. on the set $\{t \mid \hat a_j(t) \ge \varepsilon_k\}$, $j = 1,\dots,q$, $k = 1,2,\dots$

Recall that by the assumption $M(C\Gamma) \ne \emptyset$, where $C > 0$. The following assertion holds.

Lemma 8.2. *For any $\{u'\} \in S'$, $C' < C$, and $\lambda \in M(C\Gamma)$ we have*

$$\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)u', u'\rangle + a(g'(u)u')_-\right) \mathrm{d}t \ge C' \int_0^T |u'|^2\,\mathrm{d}t, \tag{8.2}$$

starting from a certain member of the sequence $\{u'\}$.

The proof is given in Section 10.5.

Recall that S_1 was defined in Section 6.2 as the set of sequences $\{\delta w\}$ in $\mathcal{W}$ satisfying relations (6.14) and (6.15). Equivalently, S_1 is the set of sequences such that $\|\delta w\|_{\mathcal{W}} \to 0$ and relations (6.15) and (6.5)–(6.7) hold. For the set of functionals $w_1^*,\dots,w_s^*$, introduced in Section 7, and for any $\{\delta w\} \in S_1$ relations (6.6) and (6.7) imply the upper bounds $\langle w_i^*, \delta w\rangle \le o(\sqrt{\gamma})$, $i = 1,\dots,s$. This easily follows from the fact that each functional $\langle w_i^*, \delta w\rangle$ belongs to the set Q (defined by (7.4)) 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 an integral representation $\langle w_i^*, \delta w\rangle = -\int_0^T a^i g'(u)\delta u\,\mathrm{d}t$ on the cone $\mathcal{R}_{i-1}$, $i = 1,\dots,s$, where $a^i \in L^\infty(0,T;\mathbb{R}^m)$, $a^i \ge 0$, $a^i g(u) = 0$, $i = 1,\dots,s$ (see Rem. 7.1).

Lemma 8.3. *For any sequence $\{\delta w\} \in S_1$ we have*

$$-\int_0^T a^i g'(u)\delta u\,\mathrm{d}t \le o\left(\sqrt{\gamma(\delta w)}\right), \quad i = 1,\dots,s. \tag{8.3}$$

The proof is given in Section 10.6.

Remark 8.4. Although the upper bound $\langle w_i^*, \delta w\rangle \le o(\sqrt{\gamma})$ holds for any $\{\delta w\} \in S_1$ and any $i = 1,\dots,s$, it does not automatically imply upper bound (8.3), since generally speaking elements of $\{\delta w\} \in S_1$ do not belong to the cone $\mathcal{R}_{i-1}$ on which w_i^* has the representation $\langle w_i^*, \delta w\rangle = -\int_0^T a^i g'(u)\delta u\,\mathrm{d}t$.

9. Proof of Theorem 4.1

We are going to estimate $C_\gamma(\Phi^1_C, S_1)$ (given in (6.18)) from below. To this end, choose any sequence $\{\delta w\} \in S_1$ that does not vanish. Using Lemmas 8.1 and 8.3 let us construct sequences $\{u'\} \in S'$ (where S' is given in (8.1))

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

and $\{\delta u_1\} \subset \mathcal{S}_u$ (where $\mathcal{S}_u$ is given in (7.7)) as follows. Set $v_j := -g_j'(u)\delta u \chi_{\{g_j(u)=0\}}$, $j = 1, \dots, q$. Then $v_j \geq 0$ and $\|v_j\|_2 \leq O(\sqrt{\gamma})$, $j = 1, \dots, q$. By Lemma 8.3 $\int_0^T a_j^i v_j \, \mathrm{d}t \leq o(\sqrt{\gamma})$, $i = 1, \dots, s$, $j = 1, \dots, q$, where $\gamma = \gamma(\delta w)$. Set

$$\mathcal{M}_j^i = \{t \mid a_j^i(t) > 0\}, \quad i = 0, \dots, s, \ j = 1, \dots, q, \tag{9.1}$$

where $a^0 = \hat{a}$ (recall that $\hat{a}$ is given in Prop. 5.6), and

$$\mathcal{M}_j = \bigcup_{i=0}^{s} \mathcal{M}_j^i, \quad j = 1, \dots, q. \tag{9.2}$$

Let $i \in \{1, \dots, s\}$, $j \in \{1, \dots, q\}$. By Lemma 8.1, 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

$$A_j^i \cup B_j^i = \mathcal{M}_j^i, \quad A_j^i \cap B_j^i = \emptyset, \quad \operatorname{meas} B_j^i \to 0, \quad \|v_j \chi_{A_j^i}\|_\infty = o(\sqrt{\gamma}) \tag{9.3}$$

(so, in Lem. 8.1, we can set $\alpha := O(\sqrt{\gamma})$). Let $\{\varepsilon\}$ be the sequence corresponding to $\{\delta w\}$ such that (6.15) holds. Set

$$B = \left(\bigcup_{i=1}^{s} \bigcup_{j=1}^{q} B_j^i \right) \bigcup \left(\bigcup_{j=1}^{q} \{t \mid 0 < \hat{a}_j(t) < \varepsilon\} \right). \tag{9.4}$$

Then in view of (9.3) and (9.4) we have: $\operatorname{meas} B \to 0$, $\|v_j \chi_{\mathcal{M}_j \setminus B}\|_\infty = o(\sqrt{\gamma})$. Due to the hypothesis of linear independence of gradients $g_j'(u)$, there exists a sequence of functions $\{\tilde{u}\}$ in $\mathcal{U}$ such that

$$\begin{aligned} &\|\tilde{u}\|_\infty = o(\sqrt{\gamma}), \ \tilde{u}\chi_B = 0, \\ &g_j'(u)\tilde{u}\chi_{\mathcal{M}_j \setminus B} = -v_j \chi_{\mathcal{M}_j \setminus B}, \\ &g_j'(u)\tilde{u}\chi_{\{g_j(u)=0\} \setminus \mathcal{M}_j} = 0, \ j = 1, \dots, q. \end{aligned} \tag{9.5}$$

Define a sequence $\{u'\}$ by the relation $u' = \delta u \chi_B$. Obviously, $\{u'\} \in S'$. Set $\delta u_1 = \delta u - \tilde{u} - u'$, $\delta w_1 = (\delta u_1, \delta y)$. Then $\delta u_1 \in \mathcal{S}_u$. Moreover, using (9.5), we deduce that, for the sequences $\{\delta w\}$, $\{\delta w_1\}$, and $\{u'\}$, the following relations hold

$$\gamma(\delta w) = \gamma(\delta w_1) + \int_0^T |u'|^2 \, \mathrm{d}t + o(\gamma(\delta w)), \tag{9.6}$$

$$\Omega(\delta w, \lambda) = \Omega(\delta w_1, \lambda) + \int_0^T \langle \overline{H}_{uu}(w, p, a)u', u' \rangle \, \mathrm{d}t + o(\gamma(\delta w)), \tag{9.7}$$

$$\int_0^T a(g'(u)\delta u)_- \, \mathrm{d}t \geq \int_B a(g'(u)\delta u)_- \, \mathrm{d}t = \int_0^T a(g'(u)u')_- \, \mathrm{d}t \tag{9.8}$$

uniformly on Λ_0. Consequently,

$$\begin{aligned} \Phi_C^1(\delta w) &:= \max_{\lambda \in M(C\Gamma)} \left\{ \Omega(\delta w, \lambda) + \int_0^T a(g'(u)\delta u)_- \, \mathrm{d}t \right\} \\ &\geq \max_{\lambda \in M(C\Gamma)} \left\{ \Omega(\delta w_1, \lambda) \right. \\ &\quad \left. + \int_0^T \left(\langle \overline{H}_{uu}(w, p, a)u', u' \rangle + a(g'(u)u')_- \right) \mathrm{d}t \right\} + o(\gamma(\delta w)). \end{aligned} \tag{9.9}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Moreover, $\|\delta w_1\|_{\mathcal{W}} \to 0$, $\|\delta\dot y - f_u(w)\delta u_1 - f_y(w)\delta y\|_1 = o(\sqrt{\gamma(\delta w)})$, and relations (6.6)–(6.7) hold. Choose any $C' < C$. Set $w' = (0, u')$, $\gamma(w') = \int_0^T |u'|^2\,\mathrm{d}t$. Then in view of Lemma 8.2 and definition (4.5) (of the functional $\Omega_{M(C\Gamma)}$) we get

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

starting from a certain number of the sequences. Combining this inequality with (9.9) we obtain

$$\Phi^1_C(\delta w) \geq \Omega_{M(C\Gamma)}(\delta w_1) + C'\gamma(w') + o(\gamma(\delta w)). \tag{9.10}$$

Furthermore, the relation $\|\delta\dot y - f'(w)\delta w_1\|_1 = o(\sqrt{\gamma(\delta w)})$ can be written as $\delta\dot y - f_y(w)\delta y - f_u(w)\delta u_1 = \zeta$, $\|\zeta\|_1 = o(\sqrt{\gamma(\delta w)})$. Let us define a sequence $\{y_\zeta\}$ by the equation $\dot y_\zeta = f_y(w)y_\zeta + \zeta$, $y_\zeta(0) = 0$. Then $\|y_\zeta\|_{1,1} = o(\sqrt{\gamma(\delta w)})$. Set $\delta y_{\mathcal{S}} = \delta y - y_\zeta$, $\delta w_{\mathcal{S}} = (\delta u_1, \delta y_{\mathcal{S}})$. Then $\delta\dot y_{\mathcal{S}} - f'(w)\delta w_{\mathcal{S}} = 0$, and hence $\delta w_{\mathcal{S}} \in \mathcal{S}$ (see Def. (7.6) of the set $\mathcal{S}$). Since $\{\delta w_1\}$ satisfies relations (6.6)–(6.7), the same is true for $\{\delta w_{\mathcal{S}}\}$. Thus, we get

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

where $\delta\eta_{\mathcal{S}} = (\delta y_{\mathcal{S}}(0), \delta y_{\mathcal{S}}(T))$. By Lemma 7.2, for each member $\delta w_{\mathcal{S}}$ of the sequence $\{\delta w_{\mathcal{S}}\}$ there exists a solution $\bar w = (\bar u, \bar y)$ to the system

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

such that $\|\bar w\|_{\mathcal{W}} = o(\sqrt{\gamma(\delta w)})$, where $\bar\eta = (\bar y(0), \bar y(T))$. Set $\delta w_{\mathcal{K}} = \delta w_{\mathcal{S}} + \bar w$. Then for the sequence $\{\delta w_{\mathcal{K}}\}$ we obviously have

$$\delta w_{\mathcal{K}} \in \mathcal{K},\quad \|\delta w_{\mathcal{K}}\|_{\mathcal{W}} \to 0, \tag{9.11}$$

$$\gamma(\delta w_1) = \gamma(\delta w_{\mathcal{K}}) + o(\gamma(\delta w)), \tag{9.12}$$

$$\Omega_{M(C\Gamma)}(\delta w_1) = \Omega_{M(C\Gamma)}(\delta w_{\mathcal{K}}) + o(\gamma(\delta w)). \tag{9.13}$$

Now note that on the subspace $\mathcal{W}_0$ (specified by (7.1)) the functional $|\delta y(0)|^2 + \int_0^T |\delta u(t)|^2\,\mathrm{d}t$ is equivalent to the functional $\|\delta y\|^2_\infty + \int_0^T |\delta u(t)|^2\,\mathrm{d}t$ (in the sense that the first functional estimates the second one on $\mathcal{W}_0$ from below and from above with certain positive multipliers) and by (3.1) the latter is equal to $\gamma(\delta w)$ in a small neighborhood of zero in $\mathcal{W}$. Since $\delta w_{\mathcal{K}} \in \mathcal{K}$, condition (4.6) of Theorem 4.1 implies that there exists $C'_{\mathcal{K}} > 0$ such that

$$\Omega_{M(C\Gamma)}(\delta w_{\mathcal{K}}) \geq C'_{\mathcal{K}}\gamma(\delta w_{\mathcal{K}}) \tag{9.14}$$

starting from a certain member of the sequence $\{\delta w_{\mathcal{K}}\}$. Relations (9.10), (9.13), and (9.14) imply

$$\Phi^1_C(\delta w) \geq C'_{\mathcal{K}}\gamma(\delta w_{\mathcal{K}}) + C'\gamma(w') + o(\gamma(\delta w)). \tag{9.15}$$

Recall that $\gamma(w') = \int_0^T |u'|^2\,\mathrm{d}t$, hence combining (9.6) with (9.12) we get

$$\gamma(\delta w) = \gamma(\delta w_{\mathcal{K}}) + \gamma(w') + o(\gamma(\delta w)). \tag{9.16}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Consequently, the following relations hold

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

Since $\{\delta w\}$ is an arbitrary sequence of S_1, the above inequalities together with (6.18) implies $C_\gamma\big(\Phi^1_C, S_1\big) \geq \min\{C'_{\mathcal{K}}, C'\}$. The latter holds for any $C' < C$. Hence $C_\gamma\big(\Phi^1_C, S_1\big) \geq \min\{C'_{\mathcal{K}}, C\}$. Combining this inequality with (6.12) and (6.17) we get $C_\gamma(\Psi_{\Lambda_0}, \Pi_{\sigma\gamma}) \geq \min\{C'_{\mathcal{K}}, C\} > 0$. In view of (5.11) and Proposition 3.7 this completes the proof of Theorem 4.1.

10. Proofs of basic lemmas

Here we give the proofs of Lemmas 6.2, 6.4, 7.2, 8.1, 8.2, and 8.3 used in the proof of Theorem 4.1.

10.1. **Proof of Lemma 6.2**

Consider the sequence $\{(\delta w, \varepsilon)\}$, where $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$ (see (5.18)) and $\varepsilon \to +0$. For any member $(\delta w, \varepsilon) = ((\delta u, \delta y), \varepsilon)$ of the sequence $\{(\delta w, \varepsilon)\}$, we set

$$\delta u_\varepsilon(t) = \begin{cases} \delta u(t) & \text{if} \quad |\delta u(t)| < \varepsilon, \\ 0 & \text{otherwise,} \end{cases}$$
$$\delta u^\varepsilon = \delta u - \delta u_\varepsilon, \quad \delta w_\varepsilon = (\delta u_\varepsilon, \delta y), \quad \delta w^\varepsilon = (\delta u^\varepsilon, 0).$$

Then $\|\delta u_\varepsilon\|_\infty \to 0$ and hence $\{\delta w_\varepsilon\} \in S^{\mathrm{loc}}$. Moreover, $\{\delta w\} = \{\delta w_\varepsilon\} + \{\delta w^\varepsilon\}$. To continue the proof of the lemma, we need the following proposition.

Proposition 10.1. *For the sequences $\{(\delta w, \varepsilon)\}$, $\{\delta w_\varepsilon\}$, and $\{\delta w^\varepsilon\}$, the following formula holds*

$$\delta f = \delta_\varepsilon f + \delta^\varepsilon f + r_f, \tag{10.1}$$

where

$$\begin{aligned} &\delta f = f(w + \delta w) - f(w),\ \delta_\varepsilon f = f(w + \delta w_\varepsilon) - f(w), \\ &\delta^\varepsilon f = f(w + \delta w^\varepsilon) - f(w),\ r_f = (\bar\delta_y f - \delta_y f)\chi^\varepsilon, \\ &\delta_y f = f(u, y + \delta y) - f(u, y), \\ &\bar\delta_y f = f(u + \delta u, y + \delta y) - f(u + \delta u, y), \end{aligned} \tag{10.2}$$

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

Proof. We have

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

□

Let us continue the proof of the lemma. Note that from the definition of r_f in (10.2) it follows that

$$\|r_f\|_\infty \leq \|\delta_y f\|_\infty + \|\bar\delta_y f\|_\infty \leq \text{const.}\, \|\delta y\|_\infty \to 0,$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

where const. > 0 does not depend on the member of the sequence $\{(\delta w, \varepsilon)\}$. Obviously, for a given sequence $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$, one can choose a sequence $\varepsilon \to +0$ such that

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

(Note that $(b) \Rightarrow (a)$, since $\|\delta y\|_\infty \leq \sqrt{\gamma(\delta w)}$.) Then we have

$$\|r_f\|_1 \leq \frac{\|r_f\|_\infty}{\varepsilon^2} \int_{\mathcal{M}^\varepsilon} \varepsilon^2 \,\mathrm{d}t \leq \frac{\|r_f\|_\infty}{\varepsilon^2} \int_{\mathcal{M}^\varepsilon} |\delta u|^2 \,\mathrm{d}t = o(\gamma^\varepsilon), \tag{10.3}$$

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 10.1 and estimate (10.3) 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 $\delta_\varepsilon H = H(u+\delta u_\varepsilon, y+\delta y, p) - H(u,y,p)$, $\delta^\varepsilon H = H(u+\delta u^\varepsilon, y, p) - H(u,y,p)$. Hence, from definition (5.1) of the function Ψ, we obtain: $\Psi(\delta w, \lambda) = \Psi(\delta w_\varepsilon, \lambda) + \int_0^T \delta^\varepsilon H \,\mathrm{d}t + o(\gamma^\varepsilon)$ uniformly on Λ_0, and then

$$\Psi_{\Lambda_0}(\delta w) = \max_{\lambda \in \Lambda_0} \left\{ \Psi(\delta w_\varepsilon, \lambda) + \int_0^T \delta^\varepsilon H \,\mathrm{d}t \right\} + o(\gamma^\varepsilon). \tag{10.4}$$

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

$$g(u+\delta u_\varepsilon) \leq 0, \quad g(u+\delta u^\varepsilon) \leq 0. \tag{10.5}$$

Consequently, by (3.6), $\delta^\varepsilon H \geq C\Gamma(\delta u^\varepsilon)$ 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

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

(see Def. (6.10) of $\Psi_{M(C\Gamma)}$). This along with (10.4) implies

$$\Psi_{\Lambda_0}(\delta w) \geq \Psi_{M(C\Gamma)}(\delta w_\varepsilon) + C\gamma^\varepsilon + o(\gamma^\varepsilon). \tag{10.6}$$

Moreover, we obviously have $\gamma = \gamma_\varepsilon + \gamma^\varepsilon$, where $\gamma := \gamma(\delta w)$, $\gamma_\varepsilon := \gamma(\delta w_\varepsilon)$, $\gamma^\varepsilon := \gamma(\delta w^\varepsilon)$. For the sequence of pairs $\{(\delta w, \varepsilon)\}$, let us consider two cases.

Case 1. $\liminf \gamma_\varepsilon/\gamma = 0$. In this case, we choose a subsequence such that $\gamma_\varepsilon = o(\gamma)$, and hence $\gamma^\varepsilon/\gamma \to 1$ for this subsequence. Without loss of generality assume that this condition is satisfied for the whole sequence $\{\delta w\}$. Then using Lemma 6.1 and the inequality $a\delta_\varepsilon g_- \geq 0$ we get

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

(see Def. (4.5) of $\Omega_{M(C\Gamma)}$). Combining (10.6) and (10.7) we obtain

$$\Psi_{\Lambda_0}(\delta w) \geq \Omega_{M(C\Gamma)}(\delta w_\varepsilon) + C\gamma^\varepsilon + o(\gamma). \tag{10.7}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Since $|\Omega_{M(C\Gamma)}(\delta w_\varepsilon)| \le O(\gamma_\varepsilon) = o(\gamma)$ and $\gamma^\varepsilon/\gamma \to 1$, we get

$$\liminf \frac{\Psi_{\Lambda_0}(\delta w)}{\gamma(\delta w)} \ge C. \tag{10.8}$$

Case 2. $\liminf \gamma_\varepsilon/\gamma > 0$, and hence $\gamma \le O(\gamma_\varepsilon)$ and $\gamma^\varepsilon \le O(\gamma_\varepsilon)$. Let us show, in this case, that

$$\{\delta w_\varepsilon\} \in S^{\mathrm{loc}}_{o(\sqrt{\gamma})}. \tag{10.9}$$

Indeed, the sequence $\{\delta w_\varepsilon\} = \{(\delta u_\varepsilon, \delta y)\}$ satisfies the conditions $\|\delta u_\varepsilon\|_\infty + \|\delta y\|_{1,1} \to 0$, *i.e.*, $\{\delta w_\varepsilon\} \in S^{\mathrm{loc}}$, and moreover, $\varphi(u + \delta u_\varepsilon) \le 0$. Furthermore, since $\sqrt{\gamma}/\varepsilon^2 \to 0$, the following estimate holds

$$\operatorname{meas} \mathcal{M}^\varepsilon = \frac{1}{\varepsilon^2} \int_{\mathcal{M}^\varepsilon} \varepsilon^2 \, \mathrm{d}t \le \frac{1}{\varepsilon^2}\gamma^\varepsilon \cdot O(1) \le \frac{\sqrt{\gamma}}{\varepsilon^2}\sqrt{\gamma^\varepsilon} \cdot O(1) = o(\sqrt{\gamma^\varepsilon}). \tag{10.10}$$

In virtue of Proposition 10.1 $\delta f = \delta_\varepsilon f + \delta^\varepsilon f + r_f$, where $\|r_f\|_1 = o(\gamma^\varepsilon)$ (see (10.3)) and $\|\delta^\varepsilon f\|_1 = O(\operatorname{meas} \mathcal{M}^\varepsilon) = o(\sqrt{\gamma^\varepsilon})$ (see (10.10)). This and the relations $\|\delta\dot{y} - \delta f\|_1 = o(\sqrt{\gamma})$, $\gamma = O(\gamma_\varepsilon)$ imply $\|\delta\dot{y} - \delta_\varepsilon f\|_1 = o(\sqrt{\gamma_\varepsilon})$. Note that in the definition of $\delta w_\varepsilon = (\delta u_\varepsilon, \delta y)$ the component δy is the same as in δw. Therefore, conditions $\sigma(\delta w) = o(\sqrt{\gamma})$ and $\gamma \le O(\gamma_\varepsilon)$ imply $\sigma(\delta w_\varepsilon) = o(\sqrt{\gamma_\varepsilon})$. Finally, $\int_0^T \hat{a}\delta_\varepsilon g_- \, \mathrm{d}t \le \int_0^T \hat{a}\delta g_- \, \mathrm{d}t \le O(\gamma) = O(\gamma_\varepsilon)$. Thus, (10.9) is proved.

It follows from (10.9) and inequality (10.6) that

$$\begin{aligned}
\liminf \frac{\Psi_{\Lambda_0}(\delta w)}{\gamma(\delta w)} &\ge \liminf \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon) + C\gamma^\varepsilon}{\gamma} \\
&= \liminf \left(\frac{\gamma_\varepsilon}{\gamma} \cdot \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon)}{\gamma_\varepsilon} + C \cdot \frac{\gamma^\varepsilon}{\gamma} \right) \ge \liminf \left(\min \left\{ \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon)}{\gamma_\varepsilon} \, , \, C \right\} \right) \\
&= \min \left\{ \liminf \frac{\Psi_{M(C\Gamma)}(\delta w_\varepsilon)}{\gamma_\varepsilon} \, , \, C \right\} \ge \min \left\{ \inf_{S^{\mathrm{loc}}_{o(\sqrt{\gamma})}} \left(\liminf \frac{\Psi_{M(C\Gamma)}}{\gamma} \right) \, , \, C \right\} \\
&= \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right) \, , \, C \right\},
\end{aligned}$$

where $\gamma = \gamma(\delta w)$.

Thus, we have proved that for any sequence $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$ there exists a subsequence such that for the subsequence we have:

$$\liminf \Psi_{\Lambda_0}/\gamma \ge \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right) \, , \, C \right\}. \tag{10.11}$$

This implies inequality (6.11). Indeed, in the opposite case there exists a sequence $\{\delta w\} \in \Pi_{o(\sqrt{\gamma})}$ such that

$$\liminf \Psi_{\Lambda_0}/\gamma < \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right) \, , \, C \right\}.$$

Consequently, there exists a subsequence such that

$$\lim \Psi_{\Lambda_0}/\gamma < \min \left\{ C_\gamma \left(\Psi_{M(C\Gamma)}, S^{\mathrm{loc}}_{o(\sqrt{\gamma})} \right) \, , \, C \right\}.$$

But then the choice of a subsequence satisfying (10.11) is impossible. The lemma is proved.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

10.2. **Proof of Lemma 6.4**

Take any sequence $\{\delta w\} \in S^{\mathrm{loc}}_{o(\sqrt{\gamma})}$. As it was noted in Section 6.1, this sequence satisfies relations (6.3) and (6.5)–(6.8). From the relation $g(u+\delta u) \le 0$ (see (6.3)) it follows that

$$g_j(u(t)) + g_j'(u(t))\delta u(t) \le k_1 |\delta u(t)|^2,\ j = 1, \dots, q, \tag{10.12}$$

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

$$|\bar u(t)| \le k_2 \left(|\delta u(t)|^2 + \sum_{j=1}^{q} \chi_{\{\hat a_j \ge \varepsilon\}}(t) (g_j'(u(t))\delta u(t))_- \right), \tag{10.13}$$

$$g_j(u(t)) + g_j'(u(t))(\delta u(t) + \bar u(t)) \le 0 \text{ a.e., } j = 1, \dots, q, \tag{10.14}$$

$$g_j'(u(t))(\delta u(t) + \bar u(t)) = 0 \ \text{ if } \ \hat a_j(t) \ge \varepsilon,\ j = 1, \dots, q, \tag{10.15}$$

$$g_j'(u(t))\bar u(t) = 0 \ \text{ if } \ 0 < \hat a_j(t) < \varepsilon,\ j = 1, \dots, q. \tag{10.16}$$

Relations (10.13) imply $\|\bar u\|_\infty \le O(\|\delta u\|_\infty) = o(1)$. Since

$$\int_{\{\hat a_j \ge \varepsilon\}} (g_j'(u)\delta u)_- \, \mathrm{d}t \le \frac{1}{\varepsilon} \int_0^T \hat a_j (g_j'(u)\delta u)_- \, \mathrm{d}t \le \frac{1}{\varepsilon} O(\gamma(\delta w))$$

(see (6.8)), relation (10.13) implies that $\|\bar u\|_1 \le O(\gamma(\delta w))/\varepsilon$. 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

$$\int_0^T |\bar u|^2 \, \mathrm{d}t \le \|\bar u\|_\infty \|\bar u\|_1 \le O(\|\delta u\|_\infty) \cdot O(\gamma(\delta w))/\varepsilon = o(\gamma(\delta w)), \tag{10.17}$$

$$\int_0^T |\delta u| \cdot |\bar u| \, \mathrm{d}t \le \|\delta u\|_\infty \|\bar u\|_1 \le \|\delta u\|_\infty O(\gamma(\delta w))/\varepsilon = o(\gamma(\delta w)), \tag{10.18}$$

$$\int_0^T |\delta y| \cdot |\bar u| \, \mathrm{d}t \le \|\delta y\|_\infty \|\bar u\|_1 \le \|\delta y\|_\infty O(\gamma(\delta w))/\varepsilon = o(\gamma(\delta w)). \tag{10.19}$$

Set $\{\delta w_1\} = \{(\delta u + \bar u, \delta y)\}$. Relations (10.17)–(10.19) imply that, for the sequences $\{\delta w_1\}$ and $\{\delta w\}$, we have

$$\Omega(\delta w_1, \lambda) = \Omega(\delta w, \lambda) + o(\gamma(\delta w)), \quad \gamma(\delta w_1) = \gamma(\delta w) + o(\gamma(\delta w)) \tag{10.20}$$

uniformly on Λ_0. Since $\|\bar u\|_\infty \to 0$, we have $\|\delta w_1\|_{\mathcal{W}} \to 0$, and using (6.5) and the estimate $\|\bar u\|_1 = o(\sqrt{\gamma(\delta w)})$ we get

$$\|\delta \dot y - f'(w)\delta w_1\|_1 = o(\sqrt{\gamma}). \tag{10.21}$$

Moreover, the sequence $\{\delta w_1\}$ satisfies relations (6.6)–(6.7) (with $\delta\eta = \delta\eta_1$ since $\delta y = \delta y_1$) and the relations

$$g_j(u(t)) + g_j'(u(t))\delta u_1(t) \le 0 \text{ a.e., } j = 1, \dots, q, \tag{10.22}$$

$$g_j'(u(t))\delta u_1(t) = 0 \ \text{ if } \ \hat a_j(t) \ge \varepsilon,\ j = 1, \dots, q, \tag{10.23}$$

$$g_j'(u(t))\delta u(t) = g_j'(u(t))\delta u_1(t) \ \text{ if } \ 0 < \hat a_j(t) < \varepsilon,\ j = 1, \dots, q. \tag{10.24}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

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

$$\int_0^T a_j(g_j'(u)\delta u_1)_- \,\mathrm{d}t \le \int_0^T a_j(g_j(u+\delta u))_- \,\mathrm{d}t + o(\gamma) \tag{10.25}$$

uniformly on Λ_0. Indeed, since $\varepsilon \to 0$ and in view of (5.16) $\{t \mid \hat{a}_j(t) \le \varepsilon\} \subset \{t \mid a_j(t) \le \hat{C}\varepsilon\}$, relations (10.23) and (10.24) imply

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

From relations (6.9), (6.10), (6.16), (10.20), and (10.25) it follows 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 (10.20) 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 $S^{\mathrm{loc}}_{o(\sqrt{\gamma})}$, inequality (6.17) follows. The lemma is proved.

10.3. **Proof of Lemma 7.2**

In order to prove Lemma 7.2 we need five auxiliary propositions, which will be proved bellow. Recall that the critical cone $\mathcal{K}$ was defined by relations (4.1)–(4.3) and the support $\mathcal{S}$ was defined by (7.6).

Proposition 10.2. *The following equality holds*

$$\mathcal{K} = \{\delta w \in \mathcal{S} \mid \textit{conditions (4.1) hold}\}. \tag{10.26}$$

Proof. Let $\mathcal{K}_{\mathcal{S}}$ denote the r.h.s. of equality (10.26). Since $\mathcal{K} \subset \mathcal{R}_s = \mathcal{S}$ and each element of $\mathcal{K}$ satisfies conditions (4.1), we have $\mathcal{K} \subset \mathcal{K}_{\mathcal{S}}$. *Vice versa*, each element of $\mathcal{K}_{\mathcal{S}}$ satisfies all equalities and inequalities in (7.7), (4.1), and (4.2), and hence belongs to $\mathcal{K}$. It means that $\mathcal{K}_{\mathcal{S}} \subset \mathcal{K}$. Consequently, $\mathcal{K} = \mathcal{K}_{\mathcal{S}}$. □

Recall that the cone Q (defined in (7.4)) and the system of functionals $w_1^*, \dots, w_s^*$ were introduced in Section 7. Let us note that from the maximality of this system we get the following important property of the cone $\mathcal{S}$.

Proposition 10.3. *If a linear functional $w^* \in Q$ has an integral representation $\langle w^*, \delta w\rangle = -\int_0^T ag'(u)\delta u \,\mathrm{d}t$ on the cone $\mathcal{S}$ such that $a \in L^1(0,T;\mathbb{R}^{q*})$, $a \ge 0$, and $ag(u) = 0$, then $\langle w^*, \delta w\rangle = 0$ for all $\delta w \in \mathcal{S}$.*

Recall that $\mathcal{S}_u$ was defined by (7.7), and the sets $\mathcal{M}_j$ were defined by (9.2). Obviously, $\mathcal{M}_j = \{t \mid \sum_{i=0}^s a_j^i(t) > 0\}$, $j = 1, \dots, q$. Set

$$\mathcal{N}_j = \{t \mid g_j(u(t)) = 0\} \setminus \mathcal{M}_j, \quad j = 1, \dots, q. \tag{10.27}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Proposition 10.4. *The following equality holds*

$$\mathcal{S}_u = \left\{\delta u \in \mathcal{U} \mid g_j'(u)\delta u \chi_{\mathcal{N}_j} \le 0,\ g_j'(u)\delta u \chi_{\mathcal{M}_j} = 0,\ j = 1, \dots, q\right\}, \tag{10.28}$$

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, $j = 1, \dots, q$.

Proof. Let $\tilde{\mathcal{S}}_u$ denote the r.h.s. of equality (10.28). If $\delta u \in \tilde{\mathcal{S}}_u$, then we have $g_j'(u)\delta u \chi_{\{g_j(u)=0\}} \le 0$, since $\mathcal{M}_j \cup \mathcal{N}_j = \{t \mid g_j(u(t)) = 0\}\ \forall j$. Moreover, for any j, the condition $g_j'(u)\delta u \chi_{\mathcal{M}_j} = 0$ implies $a_j^i g_j'(u)\delta u = 0$ $\forall i$ and hence $a^i g'(u)\delta u = 0\ \forall i$. Consequently, $\delta u \in \mathcal{S}_u$. *Vice versa*, let $\delta u \in \mathcal{S}_u$. For any i, the condition $a^i g'(u)\delta u = 0$ means that $\sum_j a_j^i g_j'(u)\delta u = 0$. Along with the conditions $g_j'(u)\delta u \chi_{\{g_j(u)=0\}} \le 0$, $a_j^i \ge 0$, and $a_j^i g_j(u) = 0\ \forall j$ this implies that $a_j^i g_j'(u)\delta u = 0$ and hence $\sum_i a_j^i g_j'(u)\delta u = 0\ \forall j$. Consequently, $g_j'(u)\delta u \chi_{\mathcal{M}_j} = 0$ $\forall j$. Then from the conditions $g_j'(u)\delta u \chi_{\{g_j(u)=0\}} \le 0\ \forall j$ it follows that $g_j'(u)\delta u \chi_{\mathcal{N}_j} \le 0\ \forall j$. Consequently, $\delta u \in \tilde{\mathcal{S}}_u$. Thus, $\mathcal{S}_u = \tilde{\mathcal{S}}_u$. □

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

$$a_j(t) = 0 \text{ if } t \notin \mathcal{M}_j \cup \mathcal{N}_j; \quad a_j(t) \ge 0 \text{ if } t \in \mathcal{N}_j,\ j = 1, \dots, q. \tag{10.29}$$

Proposition 10.5. *$\mathcal{C}^* = \mathcal{S}_u$, where $\mathcal{C}^*$ is the cone polar to C.*

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 (10.28) and (10.29) we get $\langle \delta u, h\rangle = \int_0^T h\delta u\, dt = -\int_0^T \sum_{j=1}^q a_j g_j'(u)\delta u\, dt \ge 0$. (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 $\int_0^T h\delta u\, dt = -\sum_{j=1}^q \int_0^T a_j g_j'(u)\delta u\, dt \ge 0$, where $a \in L^1(0, T; \mathbb{R}^{q*})$ satisfies (10.29). This implies that $g_j'(u)\delta u \chi_{\mathcal{N}_j} \le 0$ and $g_j'(u)\delta u \chi_{\mathcal{M}_j} = 0$, $j = 1, \dots, q$, *i.e.*, $\delta u \in \mathcal{S}_u$. □

Now set

$$\Sigma = \{(\delta u, \delta y_0) \in \mathcal{U} \times \mathbb{R}^n \mid \delta u \in \mathcal{S}_u\}. \tag{10.30}$$

Obviously, there is a one-to-one correspondence between the sets $\mathcal{S}$ (defined in (7.6)) and Σ given by the mapping $(\delta u, \delta y) \in \mathcal{S} \to (\delta u, \delta y(0)) \in \Sigma$.

Let $\mu = (\mu_0, \dots, \mu_r) \in \mathbb{R}^{(r+1)*}$, $\varphi^\mu = \sum_{i=0}^r \mu_i \phi_i$, and let $\varphi_\eta^\mu(\eta)\delta\eta$ be the functional (defined on the subspace $\mathcal{W}_0$) introduced in Section 7. In what follows, it will be convenient to treat $\varphi_\eta^\mu(\eta)\delta\eta$ as a functional defined on Σ. Namely, define a linear functional

$$(\delta u, \delta y_0) \in \mathcal{U} \times \mathbb{R}^n \to \varphi_{y_0}^\mu(\eta)\delta y_0 + \varphi_{y_T}^\mu(\eta)\delta y(T), \tag{10.31}$$

where δy is the solution to the equation $\delta\dot{y} = f_u(w)\delta u + f_y(w)\delta y$ with the initial condition $\delta y(0) = \delta y_0$. For this functional, we shall use the same notation $\varphi_\eta^\mu(\eta)\delta\eta$. This will also concern elements of Q. Set

$$P = \{(h, \xi) \in L^1(0, T; \mathbb{R}^{m*}) \times \mathbb{R}^{n*} \mid h \in \mathcal{C},\ \xi = 0\}. \tag{10.32}$$

Since by Proposition 10.5, $\mathcal{C}^* = \mathcal{S}_u$, we obviously have

$$P^* = \Sigma. \tag{10.33}$$

Moreover, each element of P can be treated as a functional on $\mathcal{U} \times \mathbb{R}^n$ with a pure integral representation. The following property will play a crucial role for the existence of corresponding Hoffman's error bounds.

Proposition 10.6. *There exists a subspace L in P such that $P \cap Q \subset L \subset P$.*

Proof. Let $l \in P \cap Q$. Then l has an integral representation: $l(\delta u, \delta y_0) = -\int_0^T ag'(u)\delta u\, dt$, where $a \in L^1(0, T; \mathbb{R}^{q*})$ satisfies (10.29). Set $\tilde{a}_j = a_j \chi_{\mathcal{N}_j}\ \forall j$, $\tilde{a} = (\tilde{a}_1, \dots, \tilde{a}_q)$. Then, obviously, we have $l(\delta w) = -\int_0^T \tilde{a}g'(u)\delta u\, dt$ on $\mathcal{S}$. Moreover, $\tilde{a} \ge 0$, $\tilde{a}g(u) = 0$, and $\tilde{a}_j \chi_{\mathcal{N}_j} = \tilde{a}_j\ \forall j$. Since $l \in Q$, by Proposition 10.3, l is

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

equal to zero on $\mathcal{S}$ and hence $\int_0^T \tilde{a} g'(u)\delta u \,\mathrm{d}t = 0$ on $\mathcal{S}_u$. This implies that $\tilde{a} = 0$. Consequently, $a_j \chi_{\mathcal{N}_j} = 0 \ \forall j$. The latter conditions define a subspace L in P (see (10.29) and (10.32)). □

Now we can prove Lemma 7.2. The proof will be based on abstract Lemma 11.9 proved in Section 11.

Proof of Lemma 7.2. For any $\delta w \in \mathcal{S}$ consider system (7.8) of linear inequalities and equalities on the cone Σ (*i.e.*, with $(\bar{u}, \bar{y}_0) \in \Sigma$). According to Proposition 10.6 and since $P^* = \Sigma$, this system on Σ satisfies all conditions of Lemma 11.9 (with $X = \mathcal{U} \times \mathbb{R}^n$, $Y = L^1(0,T;\mathbb{R}^{m*}) \times \mathbb{R}^{n*}$, $K = \Sigma$, $K_Y = P$, $K_{lm} = Q$). By this lemma, for any $\delta w \in \mathcal{S}$, system (7.8) is compatible and has Hoffman's error bound. Consequently, the same system considered on the cone $\mathcal{S}$ (*i.e.*, with $\bar{w} = (\bar{u}, \bar{y}) \in \mathcal{S}$) is always compatible and has Hoffman's error bound too (note that a finite dimensional subspace in Def. 11.8 can be equipped with the norm of the space $\mathcal{W}$). The lemma is proved.

10.4. **Proof of Lemma 8.1**

Recall that at the beginning of Section 8, $L^p_+(0,T;\mathbb{R})$ was defined as the cone of nonnegative functions in $L^p(0,T;\mathbb{R})$. The following proposition will be used in the proof of Lemma 8.1.

Proposition 10.7. *Let $r \in L^\infty_+(0,T;\mathbb{R})$, $\mathcal{M} = \{t \mid r(t) > 0\}$, $v \in L^1_+(0,T;\mathbb{R})$, and $\mathcal{N} = \{t \mid v(t) > 0\}$. Assume that $\mathcal{N} \subset \mathcal{M}$. For $\delta > 0$, set*

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

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

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

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

$$\delta\varepsilon \operatorname{meas} \mathcal{N}^\varepsilon = \int_{\mathcal{N}^\varepsilon \cap \mathcal{M}^\delta} \delta\varepsilon \,\mathrm{d}t + \int_{\mathcal{N}^\varepsilon \cap \mathcal{M}_\delta} \delta\varepsilon \,\mathrm{d}t \le \int_0^T rv \,\mathrm{d}t + \delta \int_{\mathcal{M}_\delta} v \,\mathrm{d}t.$$

The required estimate follows. □

Proof of Lemma 8.1. Without loss of generality, assume that $\{t \mid v(t) > 0\} \subset \mathcal{M}$ for all members of the sequence $\{v\}$. According to the conditions of the lemma, we have $\|v\|_2 \le \alpha$ and $\int_0^T rv \,\mathrm{d}t \le \rho\alpha$, 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 10.7. Obviously, $\operatorname{meas} \mathcal{M}_\delta \to 0$ and hence $\|v_\varepsilon\|_\infty \le \varepsilon = o(\alpha)$. Since $\|v\|_2 \le \alpha$, according to Proposition 10.7 we have

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

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

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

10.5. **Proof of Lemma 8.2**

Assume that the assertion of the lemma is not true, *i.e.*, there exist a sequence $\{u'\} \in S'$, a number $C' < C$, and an element $\lambda \in M(C\Gamma)$ such that

$$\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)u', u'\rangle + a(g'(u)u')_-\right) \mathrm{d}t \le C' \int_0^T |u'|^2 \,\mathrm{d}t \tag{10.34}$$

for all members of the sequence. Since $\{u'\}$ satisfies (8.1), 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 $\varepsilon \to +0$ and $k_1 > 0$ 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

$$g(u+u'+v) \le 0, \quad g_j(u+u'+v)\chi_{\{\hat{a}_j \ge \varepsilon\}} = 0, \quad |v| \le k_2|u'|^2, \tag{10.35}$$

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

$$|\delta u|^2 = |u'|^2 + r_1, \ \langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle = \langle \overline{H}_{uu}(w,p,a)u', u'\rangle + r_2, \tag{10.36}$$

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, using the second and the third relations in (10.35), as well as inequalities (5.16), we get for any $j = 1,\dots,q$

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

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,

$$\int_0^T |\delta u|^2 \,\mathrm{d}t = \int_0^T |u'|^2 \,\mathrm{d}t + o(\gamma'), \tag{10.37}$$

$$\int_0^T \langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle \,\mathrm{d}t = \int_0^T \langle \overline{H}_{uu}(w,p,a)u', u'\rangle \,\mathrm{d}t + o(\gamma'), \tag{10.38}$$

$$\int_0^T a_j g_j(u+\delta u)_- \,\mathrm{d}t \le \int_0^T a_j\big(g_j'(u)u'\big)_- \,\mathrm{d}t + o(\gamma'), \tag{10.39}$$

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

$$\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle + a\delta g_-\right) \mathrm{d}t \le \int_0^T \left(\langle \overline{H}_{uu}(w,p,a)u', u'\rangle + a\big(g'(u)u'\big)_-\right) \mathrm{d}t + o(\gamma'), \tag{10.40}$$

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

$$\begin{aligned}\int_0^T \left(\langle \overline{H}_{uu}(w,p,a)\delta u, \delta u\rangle + a\delta g_-\right) \mathrm{d}t &= \int_0^T \left(\delta_u \overline{H} + a\delta g_-\right) \mathrm{d}t + o\left(\int_0^T |\delta u|^2 \,\mathrm{d}t\right) \\ &= \int_0^T \delta_u H \,\mathrm{d}t + o\left(\int_0^T |\delta u|^2 \,\mathrm{d}t\right).\end{aligned} \tag{10.41}$$

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Combining relations (10.34), (10.37), (10.40), and (10.41), we get

$$\int_0^T \delta_u H \,\mathrm{d}t + o\left(\int_0^T |\delta u|^2 \,\mathrm{d}t\right) \le C' \int_0^T |\delta u|^2 \,\mathrm{d}t. \tag{10.42}$$

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 $\int_0^T \delta_u H \,\mathrm{d}t \ge C \int_0^T |\delta u|^2 \,\mathrm{d}t$. This together with estimate (10.42) contradicts the inequality $C > C'$. The lemma is proved.

10.6. **Proof of Lemma 8.3**

Let $\{\delta w\} \in S_1$ and let $\{\varepsilon\}$ be the corresponding sequence of positive numbers converging to zero (see (6.15)). For any $j = 1, \dots, q$, define a sequence of sets $B_j^0 = \{t \mid 0 < \hat{a}_j(t) < \varepsilon\}$. Let $B^0 = \bigcup_{j=1}^q B_j^0$, $\delta u_{B^0} = \delta u \chi_{B^0}$ 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. Note that $\|\delta u^0\|_2 \le \|\delta u\|_2 \le \sqrt{\gamma(\delta w)}$. 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 have $g_j'(u)\delta u^0 \chi_{\{g_j(u)=0\}} \le 0$, $\hat{a}_j g_j'(u)\delta u^0 = 0$, $j = 1, \dots, q$. 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 $\mathcal{R}_0$ (see (7.5)). From (6.5) 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 $\dot{\bar{y}} = f_y(w)\bar{y} + f_u(w)\delta u_{B^0} + \zeta$, $\bar{y}(0) = 0$, from which, in view of estimates $\|\delta u_{B^0}\|_1 = o(\sqrt{\gamma(\delta w)})$ and $\|\zeta\|_1 = o(\sqrt{\gamma(\delta w)})$, it follows that $\|\bar{y}\|_{1,1} = o(\sqrt{\gamma(\delta w)})$. This and the estimate $\langle w_1^*, \delta w\rangle := \varphi_\eta^{\mu^1}(\eta)\delta\eta \le o(\sqrt{\gamma(\delta w)})$ imply $\langle w_1^*, \delta w^0\rangle \le o(\sqrt{\gamma(\delta w)})$. But now $\delta w^0 \in \mathcal{R}_0$, therefore we get $-\int_0^T a^1 g'(u)\delta u^0 \,\mathrm{d}t \le o(\sqrt{\gamma(\delta w)})$. Since $\|\delta u_{B^0}\|_1 = o(\sqrt{\gamma(\delta w)})$, we also get $-\int_0^T a^1 g'(u)\delta u \,\mathrm{d}t \le o(\sqrt{\gamma(\delta w)})$.

Now, using Lemma 8.1, let us change the members of 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 $v_j^0 = -g_j'(u)\delta u^0 \chi_{\{g_j(u)=0\}}$, $j = 1, \dots, q$. Then we have $v_j^0 \ge 0$, $\|v_j^0\|_2 \le O(\sqrt{\gamma(\delta w)})$ and $\int_0^T a^1 v_j^0 \,\mathrm{d}t = o(\sqrt{\gamma(\delta w)})$, $j = 1, \dots, q$. According to Lemma 8.1, for any j 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 $\|v_j^0 \chi_{A_j^1}\|_\infty = o(\sqrt{\gamma(\delta w)})$. Let $B^1 = \bigcup_{j=1}^q B_j^1$ and $\delta u_{B^1}^0 = \delta u^0 \chi_{B^1}$. Since $\operatorname{meas} B^1 \to 0$ and $\|\delta u^0\|_2 \le \sqrt{\gamma(\delta w)}$ we have $\|\delta u_{B^1}^0\|_1 = o(\sqrt{\gamma(\delta w)})$. In view of the hypothesis of linear independence of the 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 $\|\bar{u}^0\|_\infty = o(\sqrt{\gamma(\delta w)})$, $\bar{u}^0 \chi_{B^1} = 0$, $g_j'(u)\bar{u}^0 = v_j^0$ a.e. on the set $A_j^1 \setminus B^1$, and $g_j'(u)\bar{u}^0 = 0$ a.e. on the set $\{t \mid g_j(u(t)) = 0\} \setminus \mathcal{M}_j^1$, $j = 1, \dots, q$. Set $\delta u^1 = \delta u^0 - \delta u_{B^1}^0 - \bar{u}^0$ 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, as it is easily seen, 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 $\langle w_2^*, \delta w^1\rangle \le o(\sqrt{\gamma(\delta w)})$. This implies that $-\int_0^T a^2 g'(u)\delta u^1 \,\mathrm{d}t \le o(\sqrt{\gamma(\delta w)})$, and then $-\int_0^T a^2 g'(u)\delta u \,\mathrm{d}t \le o(\sqrt{\gamma(\delta w)})$. Similarly, we change the sequence $\{\delta u^1\}$ on the sets $\mathcal{M}_j^2 = \{t \mid a_j^2(t) > 0\}$, $j = 1, \dots, q$, and deduce that $-\int_0^T a^2 g'(u)\delta u \,\mathrm{d}t \le o(\sqrt{\gamma(\delta w)})$, etc. The lemma is proved.

11. Hoffman's error bounds for a system of linear inequalities on a convex cone

Formally this section is independent from the others. Following [18,20], pp. 229–243, in Sections 11.1–11.3 we study two questions concerning a system of linear inequalities on a convex cone: (a) the existence of a solution to the system; (b) the existence of Hoffman's type [7] upper bounds for the distance from the origin to the set of solutions to the system. In Section 11.4 we consider the same questions for a system of linear inequalities and equalities on a cone. In Section 11.5, we formulate an abstract notion of support of the critical cone and prove that the system of inequalities defining this cone possesses a Hoffman's error bound on the support. The main

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

result of Sections 11.1–11.3 (Lem. 11.6) was published in [18] with a short sketch of the proof. Section 11.5 was written in 2007.

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

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

$$\langle l_i, x\rangle + \xi_i \le 0, \quad i = 1, \dots, k, \quad x \in K. \tag{11.1}$$

We will write it briefly as

$$l(x) + \xi \le 0, \quad x \in K, \tag{11.2}$$

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 (11.2) is compatible (*i.e.*, has a solution) iff $(-\xi) \in \Omega$. In what follows, for any $A \subset \mathbb{R}$ the inequality $A \ge 0$ means that $\inf A \ge 0$.

Lemma 11.1. *Assume that the cone Ω is closed. Then system (11.2) is compatible iff the relations $\alpha \in \mathbb{R}^{k*}_+$, $\alpha l(K) \ge 0$ imply $\alpha\xi \le 0$.*

Proof. Let system (11.2) be incompatible, *i.e.*, $-\xi \notin \Omega$. Since Ω is a 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 (11.2) is compatible.

Vice versa, let system (11.2) 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 the 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 corollary of Lemma 11.1.

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

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

Now let us study the second question: how one can estimate (from above) the distance from the origin to the set of solutions to system (11.1)? Although X is assumed to be a linear space, any finite dimensional subspace $H \subset X$ can be endowed with a norm $\|\cdot\|$ and any two norms on H are equivalent. The following definition pertains to this possibility.

Definition 11.3. We say that system (11.2) has *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 (11.2) is compatible, then the system

$$l(x) + \xi \le 0, \quad x \in K \cap H \tag{11.3}$$

is also compatible, and there exists a solution x_0 to system (11.3) such that

$$\|x_0\| \le N|\xi_+|, \tag{11.4}$$

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

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

The well-known Hoffman's lemma states that if $K = X$ and the space X is finite dimensional, then system (11.2) has Hoffman's error bound.

So, we shall study the question: under what conditions system (11.2) has Hoffman's error bound? In the sequel, we do 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 $a_1, \dots, a_s$ (*generators of the cone*) such that each element $x \in C$ can be represented as $x = \lambda_1 a_1 + \dots + \lambda_s a_s$ with $\lambda_1 \geq 0, \dots, \lambda_s \geq 0$. In a finite dimensional space, a cone is finite faced iff it 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 (11.2) has Hoffman's error bound.*

Proof. Since Ω is a finite faced cone in $\mathbb{R}^k$, then Ω is finite generated. Let $\xi^i = l(x^i) + \eta^i$, $i = 1, \dots, r$ be the generators of the cone Ω, where $x^i \in K$, $\eta^i \in \mathbb{R}^k_+$, $i = 1, \dots, r$. Let H be the linear span of the set $\{x^1, \dots, x^r\}$, and let K_0 be the cone generated by $x^1, \dots, x^r$. Then it is easy to see that $\Omega = l(K_0) + \mathbb{R}^k_+$. The cone K_0 is finite generated and hence finite faced in H. Consequently, there exist linear functionals $m_j : H \to \mathbb{R}$, $j = 1, \dots, s$ such that $K_0 = \{x \in H \mid \langle m_j, x\rangle \leq 0,\ j = 1, \dots, s\}$. In the finite dimensional subspace H, let us consider the system

$$\langle l_i, x\rangle + \xi_i \leq 0, \quad i = 1, \dots, k, \quad \langle m_j, x\rangle \leq 0, \quad j = 1, \dots, s, \quad x \in H. \tag{11.5}$$

The set of solutions to system (11.5) contains in the set of solutions to system (11.1); moreover, both sets are empty or nonempty simultaneously. Now, using Hoffman's lemma for system (11.5), we obtain the required 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 (11.2) has Hoffman's error bound.*

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

Lemma 11.6. *Let Y be a locally convex topological space, $X = Y^*$ the dual space, K_Y a closed convex cone in Y, $K = K_Y^*$ the dual cone in X, l_i, $i = 1, \dots, k$ a set of elements of the space Y considered as the linear functionals on X. Denote by K_l the cone in Y generated by the elements l_i, $i = 1, \dots, k$,* i.e., *$K_l = \{y = \alpha l \mid \alpha \in \mathbb{R}^{k*}_+\}$. Assume that there exists a subspace $S \subset Y$ such that $K_l \cap K_Y \subset S \subset K_Y$. Then system (11.2) has Hoffman's error bound. Moreover, system (11.2) is compatible iff the conditions $\alpha \in \mathbb{R}^{k*}_+$, $\alpha l \in K_Y$ imply that $\alpha\xi \leq 0$. In particular, 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 there exists a subspace $S \subset Y$ such that $K_1 \cap K_2 \subset S \subset K_2$. Then $(K_1 \cap K_2)^* = K_1^* + K_2^*$.*

Proof. (a) First, consider the case $K_1 \cap K_2 = S$, where S is a subspace in Y. Set $Q = K_1^* + K_2^*$. We must 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 see 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 $K_2^{**} \cap Y = K_2$ (here, local convexity of Y was used). 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 $\langle x^*, x\rangle < 0$. But since $x^* \in Q^*$, $Q^* = S$, and $x \in S^*$, we have $\langle x^*, x\rangle \geq 0$. This contradiction proves that the condition $Q \neq S^*$ does not hold, *i.e.*, $Q = S^*$.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

(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 can assume that the subspace S is finite dimensional.

Set $\hat{K}_1 = K_1 + S$. Let us show that $\hat{K}_1 \cap K_2 = S$. Since $0 \in K_1$, we have $S \subset \hat{K}_1$. 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 can be also considered as a finite generated cone. As it 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 the cone $\hat{K}_1 = K_1 + S$ is finite generated, 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 $(K_1 \cap K_2)^* = K_1^* + S^* = K_1^* + (K_1^* \cap S^*) + K_2^* = K_1^* + K_2^*$. The theorem is proved. □

This theorem was discussed and proved in the seminar of Milyutin.

Proof of Lemma 11.6.. Set $C = (K_l \cap K_Y)^*$, $\Omega = l(K) + \mathbb{R}^k_+$. 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, as it was mentioned in the proof of Theorem 11.7, $K_l \cap K_Y = K_l \cap S$. Since K_l is a finite generated cone and S is a subspace, the cone $K_l \cap S$ is finite generated 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 $l(x) + \xi \le 0$, $x \in K$ is equivalent to the condition: if $\alpha \in \mathbb{R}^k_+$, $\alpha l(K) \ge 0$, then $\alpha\xi \le 0$. But since $l_i \in Y$ for all i, the condition $\alpha l(K) \ge 0$ is equivalent to the condition $\alpha l \in K^* \cap Y = K_Y$. This implies the assertion of the lemma concerning 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. □

11.4. **System of linear inequalities and equalities on a cone**

Let Y be a locally convex topological space, $X = Y^*$ the dual space, K_Y a closed convex cone in Y, $K = K_Y^*$ the dual cone in X, l_i, $i = 1, \dots, k$ and m_j, $j = 1, \dots, s$ two sets of elements of the space Y considered as the linear functionals on X. Let $l : X \to \mathbb{R}^k$ and $m : X \to \mathbb{R}^s$ be the linear operators corresponding to the sets of functionals l_i, $i = 1, \dots, k$ and m_j, $j = 1, \dots, s$, respectively. Consider the system of linear inequalities and equalities on the cone K:

$$l(x) + \xi \le 0,\ m(x) + \eta = 0,\ x \in K, \tag{11.6}$$

where $\xi \in \mathbb{R}^k$ and $\eta \in \mathbb{R}^s$.

Definition 11.8. We say that system (11.6) has *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 $N > 0$ such that the following condition holds: if system (11.6) is compatible, then the system

$$l(x) + \xi \le 0,\ m(x) + \eta = 0,\ x \in K \cap H \tag{11.7}$$

is compatible too, and there exists a solution x to system (11.7) such that

$$\|x\| \le N(|\xi_+| + |\eta|). \tag{11.8}$$

Denote by K_l the cone in Y generated by the elements l_i, $i = 1, \dots, k$, and denote by L_m the subspace generated by the elements m_j, $j = 1, \dots, k$. Set $K_{lm} = K_l + L_m$, *i.e.*, $K_{lm} = \{y = \alpha l + \beta m \mid \alpha \in \mathbb{R}^{k*}_+,\ \beta \in \mathbb{R}^{s*}\}$.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Lemma 11.9. *Assume that there exists a subspace $S \subset Y$ such that $K_{lm} \cap K_Y \subset S \subset K_Y$. Then system (11.6) has Hoffman's error bound. Moreover, system (11.6) is compatible iff the conditions $\alpha \in \mathbb{R}^{k*}_+$, $\beta \in \mathbb{R}^{s*}$, $\alpha l + \beta m \in K_Y$ imply $\alpha\xi + \beta\eta \le 0$. In particular, for any $x_0 \in K$, the system $l(x + x_0) \le 0$, $m(x + x_0) = 0$, $x \in K$ is always compatible.*

Proof. System (11.6) can be represented as a system of inequalities

$$l(x) + \xi \le 0,\ m(x) + \eta \le 0,\ -m(x) - \eta \le 0,\ x \in K, \tag{11.9}$$

and hence one can apply Lemma 11.6 to the latter system. Obviously, system (11.6) is compatible iff system (11.9) is compatible; moreover system (11.6) has Hoffman's error bound iff system (11.9) has Hoffman's error bound. Thus we get an analog of Lemma 11.6 for system (11.6). □

11.5. **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 $K = \{x \in \Omega \mid l_i(x) \le 0, i = 1, \dots, k\}$. The cone K will serve us as an abstract model of the critical cone (see, *e.g.*, [6]). We shall define the notion of the support of the cone K. Set $Q = \{l = \sum_{i=1}^k \alpha_i l_i \mid \alpha_i \ge 0,\ i = 1, \dots, k\}$. For any $y \in Q$ and for any $x \in K$ we obviously have: $\langle y, x\rangle \le 0$.

Let $y_1 \in Q$ be such that $y_1(\Omega) \ge 0$ (*i.e.* $y_1 \in C$) and let $x_1 \in \Omega$ be 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, $(C + \mathrm{Span}\{y_1\})^* = C^* \cap \mathrm{Span}\{y_1\}^* = \Omega \cap \{x \in X \mid \langle y_1, x\rangle = 0\} = \Omega_1$.

Let $y_2 \in Q$ be such that $y_2(\Omega_1) \ge 0$ (*i.e.* $y_2 \in C_1$) and let $x_2 \in \Omega_1$ be 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 have already cones $\Omega_1, \Omega_2, \dots, \Omega_{s-1}$, dual to the cones $C_1, C_2, \dots, C_{s-1}$, respectively, and functionals $y_1 \in Q \cap C$, $y_2 \in Q \cap C_1, \dots, y_{s-1} \in Q \cap C_{s-2}$. Let $y_s \in Q$ be such that $y_s(\Omega_{s-1}) \ge 0$ (*i.e.* $y_s \in C_{s-1}$) and let $x_s \in \Omega_{s-1}$ be 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 $C_s := C_{s-1} + \mathrm{Span}\{y_s\}$. Let us show that the vectors $y_1, \dots, y_s$ are linearly independent. Let the numbers $\lambda_1, \dots, \lambda_{s-1}$ be such that $y_s = \lambda_1 y_1 + \dots + \lambda_{s-1} y_{s-1}$. Since $y_k(\Omega_k) \equiv 0$, $k = 1, \dots, s-1$ and $\Omega_s \subset \Omega_{s-1} \subset \dots \subset \Omega_1 \subset \Omega$, we have $y_k(\Omega_{s-1}) \equiv 0$, $k = 1, \dots, s-1$, and hence $y_s(\Omega_{s-1}) \equiv 0$. This is a contradiction with condition $\langle y_s, x_s\rangle > 0$. Consequently, $y_s \notin \mathrm{Span}\{y_1, \dots, y_{s-1}\}$. Since the vectors $y_1, \dots, y_s \in Q$ are linearly independent and the set Q is finite dimensional, this process will be finished on a certain step s.

Let the system $y_1, \dots, y_s$ be maximal. Then for the cone Ω_s we have: if $y \in Q$ is such that $y(\Omega_s) \ge 0$ (*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 $\mathcal{S} = \Omega_s$, $\mathcal{S}_Y = C_s$. 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 $l_1, \dots, l_k$ considered on the cone $\mathcal{S}$, the conditions of Lemma 11.6 are satisfied. By this lemma, for any $x_0 \in \mathcal{S}$, the system $l(x + x_0) \le 0$, $x \in \mathcal{S}$ is compatible and has Hoffman's error bound. We call $\mathcal{S}$ the *support* of the cone K.

Remark 11.10. As before, let $K = \{x \in \Omega \mid l_i(x) \le 0,\ i = 1, \dots, k\}$. Denote by $\hat{l}_i$, $i = 1, \dots, \hat{k}$ all functionals of the set $l_1, \dots, l_k$ such that $\hat{l}_i(K) \equiv 0$. All functionals l_i which do not posses this property we denote by $\tilde{l}_i$, $i = 1, \dots, \tilde{k}$, where $\tilde{k} := k - \hat{k}$. 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 $i = 1, \dots, \tilde{k}$. 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 $l_i(x_0 + x) \le 0$, $i = 1, \dots, k$, $x \in \hat{\Omega}$ is compatible and has 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 $K = \hat{\Omega} = \{x \in \Omega \mid l_1(x) = 0\} \ne \Omega = \mathcal{S}$.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press

Acknowledgements. I obtained Theorem 4.1 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 this theorem. But (as it was mentioned in Introduction) this proof has never been published. The revision of the proof with its partial modification was mainly performed during my stay at École polytechnique, France, in 2007. I would like to express my gratitude to F. Bonnans for our numerous fruitful discussions which initiated this publication. Also I am thankful to anonymous referees for their valuable remarks.

References

[1] J.F. Bonnans and A. Hermant, Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **26** (2009) 561–598.
[2] J.F. Bonnans and N.P. Osmolovskii, Second-order analysis of optimal control problems with control and initial-final state constraints. *J. Convex Anal.* **17** (2010) 885–913.
[3] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimal Control Problems*. Springer, New York (2000).
[4] A.V. Dmitruk, Quadratic conditions for a Pontryagin minimum in an optimal control problem linear in control. I. Deciphering theorem. *Izv. Akad. Nauk SSSR* **50** (1986) 284–312.
[5] A.V. Dmitruk, Quadratic conditions for a Pontryagin minimum in an optimal control problem linear in control. II. Theorem on weakening inequality constraints. *Izv. Akad. Nauk SSSR* **51** (1987) 812–832.
[6] A.Ya. Dubovitski and A.A. Milyutin, Extremum problems in the presence of restrictions. *Zh. Vychislit. Mat. i Mat. Fiz.* **5** (1965) 395–453; English translation in *U.S.S.R. Comput. Math. Math. Phys.* **5** (1965) 1–80.
[7] A.J. Hoffman, On approximate solutions of systems of linear inequalities. *J. Res. Nat'l Bur. Standarts* **49** (1952) 263–265.
[8] E.S. Levitin, A.A. Milyutin and N.P. Osmolovskii, Higher-order local minimum conditions in problems with constraints. *Uspekhi Mat. Nauk* **33** (1978) 85–148; English translation in *Russian Math. Surveys* **33** (1978) 97–168.
[9] K. Malanowski, Stability and sensitivity of solutions to nonlinear optimal control problems. *Appl. Math. Optim.* **32** (1994) 111–141.
[10] K. Malanowski, *Sensitivity analysis for parametric control problems with control–state constraints. Dissertationes Mathematicae* **CCCXCIV**. Polska Akademia Nauk, Instytut Matematyczny, Warszawa (2001) 1–51.
[11] H. Maurer, First and second order sufficient optimality conditions in mathematical programming and optimal control. *Mathematical Programming Study* **14** (1981) 163–177.
[12] H. Maurer and S. Pickenhain, Second order sufficient conditions for optimal control problems with mixed control-state constraints. *J. Optim. Theory Appl.* **86** (1995) 649–667.
[13] A.A. Milyutin, *Maximum Principle in the General Optimal Control Problem*. Fizmatlit, Moscow (2001) [in Russian].
[14] A.A. Milyutin and N.P. Osmolovskii, High-order conditions for a minimum on a set of sequences in the abstract problem with inequality constraints. *Comput. Math. Model.* **4** (1993) 393–400.
[15] A.A. Milyutin and N.P. Osmolovskii, High-order conditions for a minimum on a set of sequences in the abstract problem with inequality and equality constraints. *Comput. Math. Model.* **4** (1993) 401–409.
[16] A.A. Milyutin and N.P. Osmolovskii, High-order conditions with respect to a subsystem of constraints in the abstract minimization problem on a set of sequences. *Comput. Math. Model.* **4** (1993) 410–418.
[17] A.A. Milyutin and N.P. Osmolovskii, *Calculus of Variations and Optimal Control*, *Translations of Mathematical Monographs* **180**. American Mathematical Society, Providence (1998).
[18] N.P. Osmolovskii, On a system of linear inequalities on a convex set. *Usp. Mat. Nauk.* **32** (1977) 223–224 [in Russian].
[19] N.P. Osmolovskii, *Higher-Order Necessary and Sufficient Conditions in Optimal Control.* Parts 1 and 2, Manuscript deposited in VINITI April 1, No. 2190-B and No. 2191-B (1986) [in Russian].
[20] N.P. Osmolovskii, *Theory of higher order conditions in optimal control*. Ph.D. thesis, Moscow (1988) [in Russian].
[21] N.P. Osmolovskii, Quadratic optimality conditions for broken extremals in the general problem of calculus of variations. *J. Math. Sci.* **123** (2004) 3987–4122.
[22] N.P. Osmolovskii, Sufficient quadratic conditions of extremum for discontinuous controls in optimal control problems with mixed constraints. *J. Math. Science* **173** (2011) 1–106.
[23] V. Zeidan, Extended Jacobi sufficiency criterion for optimal control. *SIAM J. Control. Optim.* **22** (1984) 294–301.
[24] V. Zeidan, The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency. *SIAM J. Control Optim.* **32** (1994) 1297–1321.

https://doi.org/10.1051/cocv/2011101 Published online by Cambridge University Press