FIRST- AND SECOND-ORDER NECESSARY CONDITIONS
FOR CONTROL PROBLEMS WITH CONSTRAINTS

ZSOLT PÁLES AND VERA ZEIDAN

ABSTRACT. Second-order necessary conditions are developed for an abstract nonsmooth control problem with mixed state-control equality and inequality constraints as well as a constraint of the form $G(x, u) \in \Gamma$, where Γ is a closed convex set of a Banach space with nonempty interior. The inequality constraints $g(s, x, u) \leq 0$ depend on a parameter s belonging to a compact metric space S. The equality constraints are split into two sets of equations $K(x, u) = 0$ and $H(x, u) = 0$, where the first equation is an abstract control equation, and H is assumed to have a full rank property in u. The objective function is $\max_{t \in T} f(t, x, u)$ where T is a compact metric space, f is upper semicontinuous in t and Lipschitz in (x, u). The results are in terms of a function σ that disappears when the parameter spaces T and S are discrete. We apply these results to control problems governed by ordinary differential equations and having pure state inequality constraints and control state equality and inequality constraints. Thus we obtain a generalization and extension of the existing results on this problem.

1. INTRODUCTION

Second-order necessary conditions for infinite dimensional optimization problems is the main focus of several papers. One ultimate goal is to obtain results for the optimal control setting by applying the theory for the abstract optimization problem. Another route in accomplishing this goal is to directly study the second variation of the optimal control problem with the given constraints (see e.g., [25], [24]). In the literature, one can find second-order conditions for optimization problems with various types of constraints: equality, inequality, or set valued (e.g., [1], [7], [20]). On the other hand, there are second-order necessary conditions for optimal control problems with equality or inequality mixed state-control constraints (see [19], [21], [24]), or for the pure control inequality constraints (see [25]). However, none of these results include the pure-state inequality constraint $g(t, x(t)) \leq 0$, where t-dependence of g is

only upper semicontinuous. In fact, all the known versions would require that, for the active constraints, g_u at the solution be of full rank, which is not the case of a pure state constraint.

The main goal of this paper is to show how this major difficulty can be overcome. First, in Section 4, we derive second-order conditions for an abstract control problem with the following constraints: (1) mixed state-control equations: $K(x, u) = 0$, $H(x, u) = 0$ (where $H_u(\hat{x}, \hat{u})$ is of full rank, $K_x(\hat{x}, \hat{u})$ and $K_u(\hat{x}, \hat{u})$ are Fredholm and compact operators, resp.), (2) constraints of the form $G(x, u) \in \Gamma$ (where Γ is a closed convex set with nonempty interior), and (3) $g(s, x, u) \le 0$, $(s \in S)$ (where S is a compact metric space). The objective function is $F(x, u) = \sup_{t \in T} f(t, x, u)$ where T is a compact metric space. The results generalize that in [20]. The proofs are based on some properties of Fredholm and compact operators and on an earlier work by the authors [22], where a more general problem is considered (see Section 2).

Then, in Section 5, we apply the results of Section 4 to a control problem with ordinary differential equations, mixed equality constraints $H(t, x(t), u(t)) = 0$, pure state and "endpoints" inequalities constraints that depend on a parameter, "endpoints" equality constraints, and constraints of the form $G(t, x(t), u(t)) \in \Gamma$. The objective function is of a general nature:

$$F(x, u) = \max_{1 \le i \le \alpha} \sup_{(t_1, \ldots, t_{\vartheta_i}) \in T_i} f_i(t_1, \ldots, t_{\vartheta_i}, x(t_1), \ldots, x(t_{\vartheta_i})).$$

A certain rank property is imposed on $\widehat{G}_u(t)$ and $\widehat{H}_u(t)$ to insure that the multipliers associated with $G(t, x(t), u(t)) \in \Gamma$ and $H(t, x(t), u(t)) = 0$ are integrable functions. In the last section, we apply these results to an optimal control problem in Mayer's form with pure state inequality constraints $g(t, x(t)) \le 0$, mixed equality and inequality constraints $H(t, x(t), u(t)) = 0$ and $G(t, x(t), u(t)) \le 0$ and with endpoints inequality and equality constraints. The obtained necessary conditions generalize those in [21] and [24] to the case where pure state inequality constraints are present.

2. THE LAGRANGE MULTIPLIER RULE

In this section we recall the multiplier rule established in our previous paper [22]. The general problem to be considered here is the following:

Assume that Z and Y are Banach spaces (over $\mathbf{R}$), $D \subset Z$ is nonempty and open, T and S are compact metric spaces, $f: T \times D \to \mathbf{R}$ and $g: S \times D \to \mathbf{R}$, $H: D \to Y$; furthermore, Q is a closed convex subset of Z with nonempty interior. Denote

$$F(z) = \sup_{t \in T} f(t, z).$$

The problem $(\mathscr{P})$ is to minimize $F(z)$ subject to

$$z \in D: g(s, z) \le 0, \quad (s \in S), \quad z \in Q, \quad H(z) = 0.$$

A point $\hat{z} \in D$ is called an *admissible point* for this problem $(\mathscr{P})$ if $g(s, \hat{z}) \le 0$, $(s \in S)$, $\hat{z} \in Q$, and $H(\hat{z}) = 0$. A point $\hat{z} \in D$ is a *local solution* for $(\mathscr{P})$ if it is an admissible point and there exists a neighbourhood U of $\hat{z}$ such that $F(z) \ge F(\hat{z})$ holds for all admissible points $z \in U$. The multiplier rule below gives first- and second-order necessary conditions on $\hat{z}$ in order that $\hat{z}$ were a local solution for $(\mathscr{P})$.

To formulate the rule, we shall need some regularity assumptions on the data of $(\mathscr{P})$. We have to recall first some concepts on differentiability.

If T is a compact metric space and $D \subset Z$ is an open set then $\mathscr{F}(T, D)$ consists of those functions $f: T \times D \to \mathbf{R}$ that satisfy the following two conditions:

- f is an upper semicontinuous function in the first variable, i.e., $t \mapsto f(t, z)$ is upper semicontinuous for each fixed $z \in D$.
- f is uniformly locally Lipschitzian at $\hat{z}$ in the second variable, i.e., there exists $\varepsilon > 0$ and $K \in \mathbf{R}$ such that $\|f(t, z') - f(t, z'')\| \le K\|z' - z''\|$ if $\|z' - \hat{z}\| < \varepsilon$ and $\|z'' - \hat{z}\| < \varepsilon$.

One can observe that these conditions imply the upper semicontinuity of $(t, z) \mapsto f(t, z)$. For functions $f \in \mathscr{F}(T, D)$, define

$$f_{[T]}^{\circ}(t, \hat{z}; d) := \limsup_{(\tau, z, \varepsilon) \to (t, \hat{z}; 0+)} \frac{f(\tau, z + \varepsilon d) - f(\tau, z)}{\varepsilon},$$

$$f_{[T]}^{\circ\circ}(t, \hat{z}; d) := \limsup_{(\tau, \varepsilon) \to (t, 0+)} 2\frac{f(\tau, \hat{z} + \varepsilon d) - f(\tau, \hat{z}) - \varepsilon f_{[T]}^{\circ}(\tau, \hat{z}; d)}{\varepsilon^2}.$$

Clearly, $f_{[T]}^{\circ}$ is a real valued function and $f_{[T]}^{\circ\circ}$ is an extended real valued function. The notion of $f_{[T]}^{\circ}$ is an extension of Clarke's generalized derivative to the case when the function depends also on a parameter. If f is Fréchet differentiable (in the second variable) and $(t, z) \mapsto f'(t, z)$ is continuous on $T \times D$, then $f_{[T]}^{\circ}(t, \hat{z}; d) = f'(t, \hat{z})(d)$. Moreover, if f is twice Fréchet differentiable with continuous $(t, z) \to f''(t, z)$ on $T \times D$, then $f_{[T]}^{\circ\circ}(t, \hat{z}; d) = f''(t, \hat{z})(d, d)$. The most important properties of $f_{[T]}^{\circ}$ and $f_{[T]}^{\circ\circ}$ are summarized in Lemma 5 of [22].

For fixed t and $\hat{z}$, the functional $d \mapsto f_{[T]}^{\circ}(t, \hat{z}; d)$ is sublinear; therefore, the Hahn-Banach theorem yields the existence of continuous linear functionals $z^* \in Z^*$ such that $z^*(d) \le f_{[T]}^{\circ}(t, \hat{z}; d)$ for all $d \in Z$. The set of these linear functionals will be denoted by $\partial_{[T]} f(t, \hat{z})$. Clearly, this set is convex and weak* compact. One can also see that if $t_n \to t_0$, $z_n^* \to z_0^*$ as $n \to \infty$ and $z_n^* \in \partial_{[T]} f(t_n, \hat{z})$, then $z_0^* \in \partial_{[T]} f(t_0, \hat{z})$, i.e., $t \mapsto \partial_{[T]} f(t, \hat{z})$ is an upper semicontinuous set valued function. This function is called the *Clarke's subgradient* (see [6]).

We introduce the following notation: If $\mathscr{C}$ is a condition for the points of T, then $T_{\mathscr{C}}$ denotes the set of those points of T where $\mathscr{C}$ is satisfied, e.g., $T_{f>0}$ denotes the set $\{t \in T | f(t) > 0\}$, where $f: T \to \mathbf{R}$ is an arbitrary function. If R is a subset of T, then ∂R stands for the boundary of R.

If $a, b: T \to \mathbf{R}$ are arbitrary functions then define $\sigma_{a,b}: T \to [-\infty, 0]$ by

$$\sigma_{a,b}(t) := \begin{cases} \liminf_{\tau \to t} \dfrac{b^2(\tau)}{4a(\tau)}, & \text{if } t \in T_{a=0, b=0} \cap \partial(T_{a<0, b>0}), \\ a(\tau) < 0, \ b(\tau) > 0 \\ 0, & \text{otherwise.} \end{cases}$$

As is stated in the Key Lemma of [22], if a and b are upper semicontinuous functions, then $\sigma_{a,b}$ is a lower semicontinuous function and $T_{\sigma_{a,b}<0}$ is nowhere dense in T. Therefore, if T is discrete, then $\sigma_{a,b}$ turns out to be identically

zero. This function will play a significant role in the description of second-order conditions for inequality constraints.

The above concepts were necessary to describe our regularity assumptions concerning the objective function and the inequality constraint. In what follows we summarize what is necessary to assume for the equality constraints.

The function $H: D \to Y$ is called *strictly Fréchet differentiable at* $\hat{z}$, if there exists a bounded linear operator $H'(\hat{z}): Z \to Y$ such that $\forall \varepsilon > 0$, $\exists \delta(\varepsilon) > 0$ with

$$\|H(z') - H(z'') - H'(\hat{z})(z' - z'')\| \le \varepsilon \|z' - z''\|$$

whenever z' and z'' satisfy $\|z' - \hat{z}\| < \delta(\varepsilon)$ and $\|z'' - \hat{z}\| < \delta(\varepsilon)$.

If H is strictly Fréchet differentiable at $\hat{z}$ and $\overline{d} \in Z$, then we introduce the second-order weak directional derivative of H by the formula

$$H''(\hat{z};\overline{d}) := \left\{ z \in Z \,\middle|\, \liminf_{\varepsilon \to 0+} \left\| z - 2\frac{H(\hat{z} + \varepsilon\overline{d}) - H(\hat{z}) - \varepsilon H'(\hat{z})\overline{d}}{\varepsilon^2} \right\| = 0 \right\}.$$

In other words, using the concept of the upper set-limit of [3],

$$H''(\hat{z};\overline{d}) = \operatorname*{Limsup}_{\varepsilon \to 0+} \left\{ 2\frac{H(\hat{z} + \varepsilon\overline{d}) - H(\hat{z}) - \varepsilon H'(\hat{z})\overline{d}}{\varepsilon^2} \right\}.$$

This set is possibly empty. If it is not empty, then we say that H is *twice weakly directionally differentiable at* $\hat{z}$ *in the direction* $\overline{d}$. If the above relations hold with "limsup" and "Liminf", respectively, then we speak about strong directional differentiability. In that case $H''(\hat{z};\overline{d})$ clearly consists of one point. When H is $\mathscr{C}^2$ near $\hat{z}$ then one can see that $H''(\hat{z};\overline{d}) = \{H''(\hat{z})(\overline{d},\overline{d})\}$.

If $Q \subset Z$ is a convex set, then we shall use the following notations.

The symbol $\delta^*(\cdot\,;Q)$ denotes the support function of Q defined by

$$\delta^*(z^*;Q) := \sup\{z^*(z): z \in Q\}, \quad z^* \in Z^*.$$

The set $Q^\circ(\hat{z},\overline{d})$, when $\hat{z} \in Q$, $\overline{d} \in Z$, is defined by

$$Q^\circ(\hat{z},\overline{d}) := \bigcup_{\overline{\varepsilon} > 0} \bigcap_{\substack{\varepsilon < \overline{\varepsilon} \\ \|w\| < \overline{\varepsilon}}} \left[\frac{1}{\varepsilon^2}(Q - \hat{z} - \varepsilon\overline{d}) + w \right].$$

This set turns out to be important in the description of the second-order necessary condition for problem $(\mathscr{P})$. It was introduced in [22] and proved that its closure is the set investigated and invented by Kawasaki [15], [16] (cf. Ioffe [12], [13], Cominetti [7]). Its nonemptyness is an important fact; one can show (see [22]), that it is necessary that $\overline{d} \in \overline{\operatorname{cone}}(Q - \hat{z})$, where "$\overline{\operatorname{cone}}$" and "cone" stand for the closed conical and conical hull in the sequel.

Now we are able to formulate our regularity assumptions on the above optimum problem $(\mathscr{P})$.

A point $\hat{z} \in D$ is called *regular* for the objective function, inequality constraint, and equality constraint of problem $(\mathscr{P})$, respectively, if the following conditions are met:

- $f \in \mathscr{F}(T,\widehat{D})$ for some neighbourhood $\widehat{D} \subset D$ of $\hat{z}$;
- $g \in \mathscr{F}(S,\widehat{D})$ for some neighbourhood $\widehat{D} \subset D$ of $\hat{z}$;
- H is strictly Fréchet differentiable at $\hat{z}$ and the range of the linear operator $H'(\hat{z})$ is a closed subspace of Y.

If all of these three conditions are satisfied then $\hat{z}$ is called a *regular point* for $(\mathcal{P})$.

Let $\hat{z}$ be a regular admissible point for the problem $(\mathcal{P})$. For arbitrary directions $\overline{d}$ denote

$$\sigma_{[f]}(t, \hat{z}; \overline{d}) := \sigma_{a_f, b_f}(t), \quad \text{with } a_f(t) := f(t, \hat{z}) - F(\hat{z}), \ b_f(t) := f^{\circ}_{[T]}(t, \hat{z}; \overline{d})$$

and

$$\sigma_{[g]}(s, \hat{z}; \overline{d}) := \sigma_{a_g, b_g}(s), \quad \text{with } a_g(s) := g(s, \hat{z}), \ b_g(s) := g^{\circ}_{[S]}(s, \hat{z}; \overline{d}).$$

A direction $\overline{d}$ is called *regular* for the objective function, inequality constraint, and equality constraint of our problem $(\mathcal{P})$ at $\hat{z}$ if, respectively,

- $f^{\circ\circ}_{[T]}(t, \hat{z}; \overline{d})$ and $\sigma_{[f]}(t, \hat{z}; \overline{d})$ are finite on T,
- $g^{\circ\circ}_{[S]}(s, \hat{z}; \overline{d})$ and $\sigma_{[g]}(s, \hat{z}; \overline{d})$ are finite on S,
- H is twice weakly directionally differentiable at $\hat{z}$ in the direction $\overline{d}$, that is $H''(\hat{z}; \overline{d})$ is nonempty,
- $Q^{\circ}(\hat{z}, \overline{d}) \neq \varnothing$.

A direction $\overline{d}$ is *regular* at $\hat{z}$ for $(\mathcal{P})$ if all of the above conditions are satisfied.

A direction $\overline{d}$ is called *critical* for the objective function, inequality constraint, equality constraint, and the convex set Q at $\hat{z}$, if, respectively,

- $f^{\circ}_{[T]}(t, \hat{z}; \overline{d}) \leq 0$ whenever $f(t, \hat{z}) = F(\hat{z})$;
- $G^{\circ}_{[S]}(s, \hat{z}; \overline{d}) \leq 0$ whenever $g(s, \hat{z}) = 0$;
- $H'(\hat{z}; \overline{d}) = 0$;
- $\overline{d} \in \overline{\text{cone}}(Q - \hat{z})$.

The direction $\overline{d}$ is *critical* for $(\mathcal{P})$ at $\hat{z}$ if it is critical for all of the constraints and the objective function.

One can check that 0 is always a regular and critical direction at $\hat{z}$ for $(\mathcal{P})$.

Denote by $\mathcal{M}(T)$ and $\mathcal{M}(S)$ the set of bounded nonnegative Borel regular measures on T and S, respectively. The Lagrange multiplier rule is stated in the following.

Theorem 1 ([22, Theorem 6]). *Let $\hat{z}$ be a regular solution of the above problem $(\mathcal{P})$. Then for all regular critical directions $\overline{d}$ and convex set $K \subset H''(\hat{z}; \overline{d})$ there exist Lagrange multipliers $\mu \in \mathcal{M}(T)$, $\nu \in \mathcal{M}(S)$, $y^* \in Y^*$, and $z^* \in Z^*$ such that at least one of ν, μ, y^* is different from zero and the following relations hold:*

(1)
$$\text{supp}\,\mu \subset T_{f(t, \hat{z}) = F(\hat{z})}, \quad \text{supp}\,\nu \subset S_{g(s, \hat{z}) = 0},$$
$$z^*(z) \leq 0 \quad \text{for } z \in Q - \hat{z},$$

(2) $$\int_T f^{\circ}_{[T]}(t, \hat{z}; z)\,d\mu(t) + \int_S g^{\circ}_{[S]}(s, \hat{z}; z)\,d\nu(s) + y^*(H'(\hat{z})z) + z^*(z) \geq 0$$

for all $z \in Z$ and

(3) $$\int_T (f^{\circ\circ}_{[T]} - 2\sigma_{[f]})(t, \hat{z}; \overline{d})\,d\mu(t) + \int_S (g^{\circ\circ}_{[S]} - 2\sigma_{[g]})(s, \hat{z}; \overline{d})\,d\nu(s)$$
$$- \delta^*(-y^*; K) - 2\delta^*(z^*; Q^{\circ}(\hat{z}, \overline{d})) \geq 0.$$

Remarks. 1. If one takes the direction $\overline{d} = 0$ and $K = \{0\}$, then the above theorem specializes to a first-order rule, in which the functions $\sigma_{[f]}$ and $\sigma_{[g]}$ do not play any role.

As it was proved in [22],

$$Q^\circ(\hat{z}, \overline{d}) \subset \text{cone}(\text{cone}(Q - \hat{z}) - \overline{d}).$$

The right-hand side being a cone, the supporting function of the left hand side is always nonpositive (if it is bounded above). When the critical direction $\overline{d}$ belongs to $\text{cone}(Q - \hat{z})$, then the above sets are equal, whence $\delta^*(z^*; Q^\circ(\hat{z}, \overline{d})) = 0$ follows.

If the sets T and S are discrete, then the functions $\sigma_{[f]}$ and $\sigma_{[g]}$ vanish, respectively. (This follows from the properties of the function σ listed earlier.)

2. Substituting $z = \overline{d}$ into (2) and applying the criticality of $\overline{d}$ and (1) we obtain that $z^*(\overline{d}) = 0$ and

$$\text{supp}\,\mu \subset \{t \in T | f(t, \hat{z}) = F(\hat{z}), f_{[T]}^\circ(t, \hat{z}; \overline{d}) = 0\},$$

$$\text{supp}\,\nu \subset \{s \in S | g(s, \hat{z}) = 0, g_{[S]}^\circ(s, \hat{z}; \overline{d}) = 0\}.$$

3. Using the notion of Clarke's subgradient, the inequality (2) is equivalent to the following: *There exist Borel measurable selections* $\alpha: T \to Z^*$ *and* $\beta: S \to Z^*$ *of* $\partial_{[T]}f(\cdot, \hat{z})$ *and* $\partial_{[S]}g(\cdot, \hat{z})$, *respectively, such that*

$$\int_T \alpha(t)(z)\,d\mu(t) + \int_S \beta(s)(z)\,d\nu(s) + y^*(H'(\hat{z})z) + z^*(z) \geq 0$$

for all $z \in Z$. The proof of this statement is a consequence of the Strassen disintegration theorem (see, e.g., [10]):

Lemma 1. *Let* (Ω, λ) *be a measure space and let* X *be a subcone of* Z. *Let* $p: \Omega \times Z \to \mathbf{R}$ *be a function with the following properties*: (i) $\omega \mapsto p(\omega, z)$ *is integrable for each* $z \in Z$; (ii) $z \mapsto p(\omega, z)$ *is sublinear for each* $\omega \in \Omega$. *Furthermore, assume that*

$$\int_\Omega p(\omega, x)\,d\lambda(\omega) \geq q(x)$$

holds for $x \in X$, *where* $-q: Z \to \mathbf{R}$ *is a sublinear function. Then there exists a function* $P: \Omega \times Z \to \mathbf{R}$ *such that* $\omega \mapsto P(\omega, z)$ *is integrable for each* $z \in Z$, $z \mapsto P(\omega, z)$ *is linear for each* $\omega \in \Omega$, $P(\omega, z) \leq p(\omega, z)$ *for all* $(\omega, z) \in \Omega \times Z$ *and*

$$\int_\Omega P(\omega, x)\,d\lambda(\omega) \geq q(x)$$

for $x \in X$.

Roughly speaking, this disintegration theorem states that integral of sublinear terms in inequalities can be decreased to linear ones. This idea will also be applied later, when we deal with control problems concerning ordinary differential equations.

Using this result twice, taking X as the conical hull of $Q - \hat{z}$, $(\Omega, \lambda) :=$ (T, μ) and $(\Omega, \lambda) := (S, \nu)$; furthermore, $p(t, z) := f_{[T]}^\circ(t, \hat{z}; z)$ and $p(t, z)$

$:= g^{\circ}_{[S]}(s,\hat{z};z)$, respectively, the above rewritten form of (2) can be obtained. This form of (2) often turns out to be an equation.

3. FREDHOLM AND COMPACT OPERATORS

In this section we briefly summarize what we shall need about bounded linear operators, especially Fredholm and compact ones. The regularity assumption on equality constraints requires the closed range property of a Fréchet derivative. In the applications this derivative is expressed in terms of Fredholm, compact, and other linear operators that come from the integral equation form of the control equation. The main goal of this section is a result on the closed range property of those operators.

Let X and Y be Banach spaces. A bounded linear operator $K\colon X \to Y$ is called *compact* if it maps bounded subsets of X into relatively compact subsets of Y. Properties of compact operators are summarized in the following

Lemma 2. *The family of compact operators* $K\colon X \to Y$ *forms a Banach subspace of* $\mathscr{B}(X,Y)$, *the space of all bounded linear operators from* X *to* Y. *The composition of a compact and a bounded linear operator is again a compact linear operator.*

Proof. See [26]. □

It turns out that the Volterra and Fredholm integral operators are always compact operators.

A bounded linear operator $F\colon X \to Y$ is called a *Fredholm operator* if the dimensions of

$$\operatorname{Ker} F := \{x \in X \,|\, F(x) = 0\} \quad \text{and} \quad \operatorname{Coker} F := Y/\operatorname{Im} F$$

are finite. (Here $\operatorname{Im} F$ denotes the range of F.) Then

$$\operatorname{ind} F := \dim \operatorname{Ker} F - \dim \operatorname{Coker} F$$

is called the index of F. (We note that Fredholm integral operators are not Fredholm operators.) We shall need the following properties of Fredholm operators:

Lemma 3. *Let* $F\colon X \to Y$ *be a Fredholm operator. Then the range of* F *is a closed subspace of* Y. *Furthermore, if* $K\colon X \to Y$ *is a compact linear operator, then* $F + K$ *is also a Fredholm operator and*

$$\operatorname{ind}(F + K) = \operatorname{ind} F.$$

Proof. The proof of this result and other more general statements can be found in the standard textbooks, e.g., in [26, Vol. I, pp. 365–372], or in [23, p. 114]. □

As a consequence of this result we have that an invertible operator perturbed with a compact one is always is Fredholm operator with index zero. We shall see later that operators of this kind come from the control equation, when written in integral equation form.

The following result gives a condition on the closed range of *glued* operators.

Lemma 4. *Let* X, Y, W *be arbitrary Banach spaces and* $A\colon X \to W$ *and* $B\colon X \to Y$ *be bounded linear operators such that* $\operatorname{Im} A$ *and* $B(\operatorname{Ker} A)$ *are closed*

subspaces. Then the operator $L: X \to W \times Y$ *defined by* $Lx := (Ax, Bx)$, $x \in X$, *admits a closed range.*

Proof. See [1, Lemma 2.1.6]. □

The next lemma summarizes the properties of glued operators, when B is of finite rank, i.e., when the range of B is of finite dimension. We shall need this statement in the sequel.

Lemma 5. *Let* X, W *be arbitrary Banach spaces and* $F: X \to W$ *be a Fredholm operator,* $K: X \to W$ *be a compact operator and* $B: X \to \mathbf{R}^m$ *be a bounded linear operator. Then the operators* $X \to W \times \mathbf{R}^m$ *defined by*

$$x \mapsto (Fx, Bx), \qquad x \mapsto (Kx, Bx)$$

are also Fredholm and compact operators, respectively.

Proof. The proof is elementary, so we can omit it. □

Now we are able to formulate the main result of this section.

Theorem 2. *Let* $X, Y, U,$ *and* W *be arbitrary Banach spaces and let*

$$A: X \to W, \ S: U \to W, \ F: X \to Y, \quad and \quad K: U \to Y$$

be given bounded linear operators. Assume that the following conditions are satisfied: (1) S *is surjective and there exists a linear bounded right inverse* $R: W \to U$ *of* S; (2) F *is a Fredholm operator;* (3) K *is a compact operator. Then the bounded linear operator* $L: X \times U \to Y \times W$ *defined by*

$$L(x, u) := \begin{pmatrix} A & S \\ F & K \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}$$

admits a closed range.

Proof. Let $\widehat{X} := X \times U$ and $\widehat{A}: \widehat{X} \to W$ and $\widehat{B}: \widehat{X} \to Y$ defined by $\widehat{A}(x, u) := Ax + Su$ and $\widehat{B}(x, u) := Fx + Su$. Then, based on Lemma 4, it is enough to show that the range of $\widehat{A}$ and $\widehat{B}(\text{Ker}\,\widehat{A})$ are closed subspaces.

Since S is surjective, therefore $\text{Im}\,\widehat{A} = W$, which is obviously closed.

Now observe that the kernel of $\widehat{A}$ can be expressed in the form

$$\text{Ker}\,\widehat{A} = \{(x, -R(Ax))|x \in X\} + (\{0\} \times \text{Ker}\,S).$$

Therefore,

$$\widehat{B}(\text{Ker}\,\widehat{A}) = \{Fx - K(R(Ax))|x \in X\} + K(\text{Ker}\,S)$$
$$= \text{Im}(F - KRA) + K(\text{Ker}\,S).$$

Using Lemma 2 and Lemma 3, one can see that $\overline{F} := F - KRA$ is a Fredholm operator, therefore its image is closed and is of finite codimension. Thus the sum $\text{Im}\,\overline{F} + X_0$ is always closed, where X_0 is an arbitrary subspace of X.

The proof is therefore complete. □

4. ABSTRACT CONTROL PROBLEMS WITH CONSTRAINTS

In this section we are going to deal with abstract control problems of the following form:

Assume that X, U, Y, V, and W are Banach spaces (over **R**), $D \subset X \times U$ is nonempty and open, T and S are compact metric spaces, $f: T \times D \to \mathbf{R}$ and $g: S \times D \to \mathbf{R}$, $G: D \to V$, $H: D \to W$, $K: D \to Y$; furthermore, $\Gamma \subset V$ is a closed convex set with $\Gamma^\circ \neq \varnothing$. Denote

$$F(x, u) = \sup_{t \in T} f(t, x, u).$$

The problem $(\mathscr{P})^*$ is to minimize $F(x, u)$ in $(x, u) \in D$ subject to

- (i) $g(s, x, u) \leq 0$, $(s \in S)$ (real-valued inequalities),
- (ii) $G(x, u) \in \Gamma$ (Banach space valued mixed state-control inequality and control set constraint),
- (iii) $H(x, u) = 0$ (Banach space valued mixed state-control equality),
- (iv) $K(x, u) = 0$ (control system).

The admissibility and optimality of a pair $(x, u) \in D$ is defined similarly to that of problem $(\mathscr{P})$.

The second constraint $G(x, u) \in \Gamma$ is able to handle Banach space valued inequalities and control set constraint as well. For instance, if Γ is a closed convex cone with nonempty interior, then introducing the ordering $\leq_\Gamma$ in V by $x \leq_\Gamma y \Leftrightarrow x - y \in \Gamma$, one can see that our second constraint can be rewritten as $G(x, u) \leq_\Gamma 0$. On the other hand the constraints $u \in Q$ or $x \in Q$, where Q is a convex set with nonempty interior, are obviously a particular case of (ii). In this case the problem $(\mathscr{P})^*$ specializes to the mixed problem dealt with in [14, Section 1.1.3, p. 70]. However, both the regularity assumptions and the results are of different nature in our case.

At this stage one cannot make any difference between the mixed state-control equality and control system constraints. However, the difference becomes clear when evoking the regularity conditions stated below.

A pair $(\hat{x}, \hat{u}) \in D$ is called *regular* for the problem $(\mathscr{P})^*$, if the following conditions are satisfied:

- $f \in \mathscr{F}(T, \hat{D})$ for some neighbourhood $\hat{D} \subset D$ of $(\hat{x}, \hat{u})$;
- $g \in \mathscr{F}(S, \hat{D})$ for some neighbourhood $\hat{D} \subset D$ of $(\hat{x}, \hat{u})$;
- G is strictly Fréchet differentiable at $(\hat{x}, \hat{u})$;
- H is strictly Fréchet differentiable at $(\hat{x}, \hat{u})$ and the partial Fréchet derivative $H_u(\hat{x}, \hat{u}): U \to W$ has the *full rank property*, that is it has a bounded right inverse;
- K is strictly Fréchet differentiable at $(\hat{x}, \hat{u})$ and the equation is an *abstract control system* at $(\hat{x}, \hat{u})$, i.e., the partial derivative $K_x(\hat{x}, \hat{u})$ is a Fredholm operator and $K_u(\hat{x}, \hat{u})$ is compact.

We note that when K fulfills the above assumption at each point of D, then the equation $K = 0$ will be called a *(global) control system*. It is worth observing that if K is continuously Fréchet differentiable on D, K_x is a Fredholm operator and D is a connected set, then $\operatorname{ind} K_x$ is constant on D and hence, the index of a control system could be defined.

We indicate by $\widehat{\Phi}$ the evaluation of the function Φ at $(\hat{x}, \hat{u})$.

Let $(\hat{x}, \hat{u})$ be a regular admissible pair for the problem $(\mathscr{P})^*$. For directions (ξ, η) we can define $\sigma_{[f]}(t, \hat{x}, \hat{u}; \xi, \eta) =: \hat{\sigma}_{[f]}(t; \xi, \eta)$ and $\sigma_{[g]}(s, \hat{x}, \hat{u}; \xi, \eta)$ $=: \hat{\sigma}_{[g]}(s; \xi, \eta)$ as was defined for the problem $(\mathscr{P})$ in the first section. A

direction (ξ, η) is called *regular* for our problem $(\mathscr{P})^*$ at $(\hat{x}, \hat{u})$ if

- $\hat{f}^{\circ\circ}_{[T]}(t; \xi, \eta)$ and $\hat{\sigma}_{[f]}(t; \xi, \eta)$ are finite on T,
- $\hat{g}^{\circ\circ}_{[S]}(s; \xi, \eta)$ and $\hat{\sigma}_{[g]}(s; \xi, \eta)$ are finite on S,
- G, H, and K are twice strongly directionally differentiable at $(\hat{x}, \hat{u})$ in the direction (ξ, η),
- $\Gamma^{\circ}(\widehat{G}, \widehat{G}_x\xi + \widehat{G}_u\eta) \neq \varnothing$.

The third assumption here always guarantees that the second-order directional derivative of G, H, and K in the direction (ξ, η) is a singleton, i.e., it is uniquely determined. These derivatives are denoted by $\widehat{G}''(\xi, \eta)$, $\widehat{H}''(\xi, \eta)$, and $\widehat{K}''(\xi, \eta)$, respectively.

A direction (ξ, η) is called *critical* for $(\mathscr{P})^*$ at $(\hat{x}, \hat{u})$ if

$$\hat{f}^{\circ}_{[T]}(t; \xi, \eta) \leq 0 \quad \text{whenever } \hat{f}(t) = \widehat{F},$$

$$\hat{g}^{\circ}_{[S]}(s; \xi, \eta) \leq 0 \quad \text{whenever } \hat{g}(s) = 0,$$

and

$$\widehat{G}_x\xi + \widehat{G}_u\eta \in \overline{\text{cone}}(\Gamma - \widehat{G}), \quad \widehat{H}_x\xi + \widehat{H}_u\eta = 0, \quad \widehat{K}_x\xi + \widehat{K}_u\eta = 0.$$

One can check that $(\xi, \eta) = (0, 0)$ is always a regular and critical direction at $(\hat{x}, \hat{u})$ for $(\mathscr{P})^*$.

Theorem 3. *Let $(\hat{x}, \hat{u})$ be a regular solution for the problem $(\mathscr{P})^*$. Then for every regular critical direction (ξ, η) there exist Lagrange multipliers $\mu \in \mathscr{M}(T)$, $\nu \in \mathscr{M}(S)$, $v^* \in V^*$, $w^* \in W^*$, and $y^* \in Y^*$ such that at least one of them is different from zero and the following relations hold:*

(4) $\text{supp}\, \mu \subset \{t \in T | \hat{f}(t) = \widehat{F}\}, \quad \text{supp}\, \nu \subset \{s \in S | \hat{g}(s) = 0\},$

(5) $v^*(v) \leq 0 \quad \text{for all } v \in \Gamma - \widehat{G} \quad \text{and} \quad v^*(\widehat{G}_x\xi + \widehat{G}_u\eta) = 0,$

(6) $\displaystyle\int_T \hat{f}^{\circ}_{[T]}(t; x, 0)\, d\mu(t) + \int_S \hat{g}^{\circ}_{[S]}(s; x, 0)\, d\nu(s)$

$\displaystyle\qquad\qquad + v^*(\widehat{G}_x x) + w^*(\widehat{H}_x x) + y^*(\widehat{K}_x x) \geq 0 \quad \text{for } x \in X,$

(7) $\displaystyle\int_T \hat{f}^{\circ}_{[T]}(t; 0, u)\, d\mu(t) + \int_S \hat{g}^{\circ}_{[S]}(s; 0, u)\, d\nu(s)$

$\displaystyle\qquad\qquad + v^*(\widehat{G}_u u) + w^*(\widehat{H}_u u) + y^*(\widehat{K}_u u) \geq 0 \quad \text{for } u \in U,$

and

(8)
$$\int_T (\hat{f}^{\circ\circ}_{[T]} - 2\hat{\sigma}_{[f]})(t; \xi, \eta)\, d\mu(t) + \int_S (\hat{g}^{\circ\circ}_{[S]} - 2\hat{\sigma}_{[g]})(s; \xi, \eta)\, d\nu(s)$$

$$+ v^*(\widehat{G}''(\xi, \eta)) - 2\delta^*(v^*; \Gamma^{\circ}(\widehat{G}, \widehat{G}_x\xi + \widehat{G}_u\eta))$$

$$+ w^*(\widehat{H}''(\xi, \eta)) + y^*(\widehat{K}''(\xi, \eta)) \geq 0.$$

Remarks. 1. If the critical pair (ξ, η) satisfies the stronger relation

$$\widehat{G}_x\xi + \widehat{G}_u\eta \in \text{cone}(\Gamma - \widehat{G}) = \{\lambda(\gamma - \widehat{G}): \lambda \geq 0, \gamma \in \Gamma\}$$

(instead of the left-hand side being in $\overline{\text{cone}}(\Gamma - \widehat{G})$), then the term in (8) containing δ^* vanishes. (Cf. Remark 1 to Theorem 1.) Similarly, if T or S are

discrete sets, then $\hat{\sigma}_{[f]}$ and $\hat{\sigma}_{[g]}$ vanish, respectively. The choice $(\xi, \eta) = (0, 0)$ leads to a first-order rule, where these functions do not play any role.

2. The inequality stated in (5) can be extended to a larger set instead of $\Gamma - \widehat{G}$. Introduce first the following notations:

(i) If $Q \subset Z$ is a convex set, $\hat{z} \in Q$, then let $Q_{\hat{z}} := \overline{\text{cone}}(Q - \hat{z})$.

(ii) If $K \subset Z$ is a convex cone, then denote by $\Lambda(K)$ the maximal linear subspace contained in K, that is $\Lambda(K) = K \cap (-K)$.

Then it follows from (5) that

$$v^*(v) \leq 0 \quad \text{for } v \in \Gamma_{\widehat{G}}$$

and

(9) $$v^*(v) = 0 \quad \text{for } v \in \Lambda(\Gamma_{\widehat{G}}),$$

since v^* is continuous and linear. This equation together with (7) plays an important role in the formulation of the strong full rank property, which is explained after the proof of the theorem.

3. Consider the case treated in [20], that is when T is a singleton, f is twice Fréchet differentiable, Γ is a cone, and $g \equiv H \equiv K \equiv 0$. In this case Theorem 3 generalizes Theorems 3.1, 3.2, and 3.3 in [20] to the case where relation (2.3) of that paper is not satisfied.

4. Consider the following problem:

$$(\overline{\mathscr{P}}): \text{Minimize } F(x, u) := \max_{1 \leq i \leq \alpha} \sup_{t \in T_i} f_i(t, x, u)$$

subject to

$$g_1(s, x, u) \leq 0, (s \in S_1), \ldots, g_\beta(s, x, u) \leq 0, (s \in S_\beta),$$
$$G_1(x, u) \in \Gamma_1, \ldots, G_\gamma(x, u) \in \Gamma_\gamma,$$

and

$$H(x, u) = 0, \quad K(x, u) = 0, \quad k_1(x, u) = 0, \ldots, k_\kappa(x, u) = 0,$$

where $T_1, \ldots, T_\alpha, S_1, \ldots, S_\beta$ are compact metric spaces, $X, U, V_1, \ldots, V_\gamma$, W, Y are Banach spaces, $\Gamma_1 \subset V_1, \ldots, \Gamma_\gamma \subset V_\gamma$ are closed convex sets with nonempty interior, $D \subset X \times U$ is an open set and

$$f_1: T_1 \times D \to \mathbf{R}, \ldots, f_\alpha: T_\alpha \times D \to \mathbf{R},$$
$$g_1: S_1 \times D \to \mathbf{R}, \ldots, g_\beta: S_\beta \times D \to \mathbf{R},$$
$$G_1: D \to V_1, \ldots, G_\gamma: D \to V_\gamma,$$
$$H: D \to W, \quad K: D \to Y, \quad k_1: D \to \mathbf{R}, \ldots, k_\kappa: D \to \mathbf{R}.$$

The notions of admissibility, optimality, and critical directions are similar to those defined for problem $(\mathscr{P})^*$. First- and second-order necessary conditions for this problem can be derived by applying Theorem 3 using the following data.

$$T := \bigcup_{i=1}^{\alpha} \{i\} \times T_i, \qquad S := \bigcup_{i=1}^{\beta} \{i\} \times S_i,$$
$$f((i, t), x, u) := f_i(t, x, u), \quad (i, t) \in T, \quad (x, u) \in D,$$
$$g((i, t), x, u) := g_i(t, x, u), \quad (i, s) \in S, \quad (x, u) \in D,$$
$$G(x, u) := (G_1(x, u), \ldots, G_\gamma(x, u)), \quad \Gamma := \Gamma_1 \times \cdots \times \Gamma_\gamma,$$

and

$$\overline{K}(x, u) := (K(x, u), k_1(x, u), \ldots, k_\kappa(x, u)).$$

Note that, using Lemma 5, $\overline{K}(x, u) = 0$ is also an abstract control equation. In the special case when the sets $T_1, \ldots, T_\alpha$ and $S_1, \ldots, S_\beta$ are singletons, the function σ disappears from the second-order necessary condition.

Proof of Theorem 3. Introduce the following notations:

$$Z := X \times U \times V, \quad \overline{Y} := V \times W \times Y, \quad \overline{D} := D \times V, \quad Q := X \times U \times \Gamma.$$

Furthermore, let

$$\overline{f}(t, z) := f(t, x, u), \quad \overline{g}(s, z) := g(s, x, u),$$
$$\overline{H}(z) := (G(x, u) - v, H(x, u), K(x, u))$$

for $z = (x, u, v) \in \overline{D}$ and $t \in T, s \in S$. Then our problem $(\mathscr{P})^*$ is equivalent to the optimization problem $(\overline{\mathscr{P}})$, which is obtained from $(\mathscr{P})$ if one replaces Y, D, f, g, and H by $\overline{Y}, \overline{D}, \overline{f}, \overline{g}$, and $\overline{H}$, respectively. Clearly, $(\hat{x}, \hat{u})$ is an optimal solution of $(\mathscr{P})^*$ if and only if $\hat{z} := (\hat{x}, \hat{u}, \hat{v}) := (\hat{x}, \hat{u}, G(\hat{x}, \hat{u}))$ solves $(\overline{\mathscr{P}})$. We still have to check that if $(\hat{x}, \hat{u})$ is a regular solution for $(\mathscr{P})^*$, then $\hat{z}$ is also a regular solution for $(\overline{\mathscr{P}})$. We are going to show only the closed range property of $\overline{H}'(\hat{z})$, the other regularity assumptions concerning differentiability are trivially satisfied.

We have

$$(10) \qquad \widehat{\overline{H}'}(\xi, \eta, \zeta) = \overline{H}'(\hat{z})(\xi, \eta, \zeta) = \begin{pmatrix} \widehat{G}_x & \widehat{G}_u & -I \\ \widehat{H}_x & \widehat{H}_u & 0 \\ \widehat{K}_x & \widehat{K}_u & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}.$$

If we define

$$\mathbf{A}\xi := \begin{pmatrix} \widehat{G}_x\xi \\ \widehat{H}_x\xi \end{pmatrix}, \quad \mathbf{S}(\eta, \zeta) := \begin{pmatrix} \widehat{G}_u & -I \\ \widehat{H}_u & 0 \end{pmatrix} \begin{pmatrix} \eta \\ \zeta \end{pmatrix}, \quad \mathbf{F}\xi := \widehat{K}_x\xi, \quad \mathbf{K}(\eta, \zeta) := \widehat{K}_u\eta,$$

then, by our regularity assumptions, $\mathbf{F}$ is a Fredholm operator and $\mathbf{K}$ is compact. To obtain that the range of $\overline{H}'(\hat{z})$ is closed, we have only to show that $\mathbf{S}: U \times V \to V \times W$ has a bounded right inverse, then Theorem 2 of the previous section yields the statement. (One has to replace A, S, F, and K there by $\mathbf{A}, \mathbf{S}, \mathbf{F}$, and $\mathbf{K}$, respectively.)

The operator $\widehat{H}_u$ has a bounded right inverse $P: W \to U$ by its full rank property. Define $\mathscr{R}: V \times W \to U \times V$ by

$$\mathscr{R}(v, w) := \begin{pmatrix} 0 & P \\ -I & \widehat{G}_u P \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}.$$

(Here and above I denotes the identical mapping on V.) Then a simple calculation shows that $\mathscr{R}$ is a right inverse of $\mathbf{S}$, thus $\mathbf{S}$ is obviously surjective and hence $\operatorname{Im}\overline{H}'(\hat{z})$ is closed.

If (ξ, η) is a regular critical direction for $(\mathscr{P})^*$, then the direction $\overline{d} := (\xi, \eta, \zeta) := (\xi, \eta, \widehat{G}_x\xi + \widehat{G}_u\eta)$ is obviously regular and critical with respect to $(\overline{\mathscr{P}})$.

Now we are able to apply Theorem 1 for the problem $(\overline{\mathscr{P}})$. Then there exist Lagrange multipliers $\mu \in \mathscr{M}(T)$, $\nu \in \mathscr{M}(S)$, $\overline{y}^* \in \overline{Y}^*$, and $z^* \in Z^*$ not all zero such that the relations in (4),

$$(11) \qquad z^*(x, u, v) \le 0 \quad \text{for } (x, u, v) \in Q - (\hat{x}, \hat{u}, \hat{v}),$$

$$(12) \qquad \begin{aligned} \int_T \hat{f}^\circ_{[T]}(t; x, u)\, d\mu(t) + \int_S \hat{g}^\circ_{[S]}(s; x, u)\, d\nu(s) \\ + \overline{y}^*(\widehat{\overline{H}}'(x, u, v)) + z^*(x, u, v) \ge 0 \end{aligned}$$

for all x, u, v, and

$$(13) \qquad \begin{aligned} \int_T (\hat{f}^{\circ\circ}_{[T]} - 2\hat{\sigma}_{[f]})(t; \xi, \eta)\, d\mu(t) + \int_S (\hat{g}^{\circ\circ}_{[S]} - 2\hat{\sigma}_{[g]})(s; \xi, \eta)\, d\nu(s) \\ + \overline{y}^*(\widehat{\overline{H}}''(\xi, \eta, \zeta)) - 2\delta^*(z^*; Q^\circ((\hat{x}, \hat{u}, \hat{v}), (\xi, \eta, \zeta)) \ge 0. \end{aligned}$$

are satisfied.

Define $v^* \in V^*$, $w^* \in W^*$, and $y^* \in Y^*$ by

$$v^*(v) := \overline{y}^*(v, 0, 0), \quad w^*(w) := \overline{y}^*(0, w, 0), \quad y^*(y) := \overline{y}^*(0, 0, y).$$

Using (10), the term $\overline{y}^*(\widehat{\overline{H}}'(x, u, v))$ in (12) can be rewritten as

$$\overline{y}^*(\widehat{\overline{H}}'(x, u, v)) = v^*(\widehat{G}_x x + \widehat{G}_u u - v) + w^*(\widehat{H}_x x + \widehat{H}_u u) + y^*(\widehat{K}_x x + \widehat{K}_u u).$$

Putting $x = 0$, $u = 0$ in (12), we get

$$z^*(0, 0, v) = v^*(v) \quad \text{for } v \in V.$$

On the other hand, (11) is equivalent to

$$z^*(x, u, v) \le 0 \quad \text{for } x \in X, u \in U, v \in \Gamma - \hat{v}.$$

Therefore

$$z^*(x, u, v) = z^*(0, 0, v) = v^*(v) \quad \text{for } x \in X, u \in U, v \in V.$$

Using this representation of z^*, (11) reduces to the first inequality of (5), and the substitutions $u = 0$, $v = 0$ and $x = 0$, $v = 0$ in (12) yield (6) and (7), respectively. As we have noted in Remark 2 to Theorem 1, the necessary conditions of Theorem 1 imply $z^*(\overline{d}) = 0$ there. Applying this relation to $\overline{d} = (\xi, \eta, \zeta)$ we get the second equality of (5).

To complete the proof of the theorem, observe that

$$\widehat{\overline{H}}''(\xi, \eta, \zeta) = \overline{H}''(\hat{z})(\xi, \eta, \zeta) = (\widehat{G}''(\xi, \eta), \widehat{H}''(\xi, \eta), \widehat{K}''(\xi, \eta)).$$

and

$$Q^\circ((\hat{x}, \hat{u}, \hat{v}), (\xi, \eta, \zeta)) = X \times U \times \Gamma^\circ(\hat{G}, \hat{G}_x\xi + \hat{G}_u\eta).$$

Then one can see that (13) is equivalent to inequality (8) of the theorem. $\square$

Remark. Having proved the theorem, we are able to formulate the *strong full rank property*: The functions G and H of the above problem $(\mathscr{P})^*$ are said to have this property if the operator $\mathbf{S}: U \times V \to V \times W$ (defined in the proof of the above theorem) has a bounded right inverse $\mathscr{R}$ with the property $\operatorname{Im}\mathscr{R} \subset U \times \Lambda(\Gamma_{\hat{G}})$. It turns out that if (7) is satisfied with equality (for

instance, when f and g are differentiable with respect to the u-variable), then (7) and the above relation (9) completely determine the linear functionals v^* and w^* as functions of the other multipliers. This fact is especially important when V and W are spaces of bounded measurable functions. In that case, the dual spaces are too large and there is no nice representation of their linear functionals without additional conditions. However, when the above strong full rank property is imposed (see Section 6), then it will always turn out that the multipliers v^* and w^* can be expressed in terms of integrable functions.

5. Control problems for ordinary differential equations with constraints

The aim of this section is to apply Theorem 3 to control problems governed by ordinary differential equations. The problem $(\widehat{\mathscr{P}})$ we deal with is the following:
Minimize

$$F(x, u) = \max_{1 \le i \le \alpha} \sup_{(t_1, \ldots, t_{\vartheta_i}) \in T_i} f_i(t_1, \ldots, t_{\vartheta_i}, x(t_1), \ldots, x(t_{\vartheta_i}))$$

subject to

  (i) $g_i(s_1, \ldots, s_{\pi_i}, x(s_1), \ldots, x(s_{\pi_i})) \le 0$ for $(s_1, \ldots, s_{\pi_i}) \in S_i$, $i = 1,$ $\ldots, \beta$ (pure state inequalities—"endpoint" inequalities);
  (ii) $G(t, x(t), u(t)) \in \Gamma$ for a.e. $t \in [a, b]$ (mixed state-control inequalities—control set constraint);
  (iii) $H(t, x(t), u(t)) = 0$ for a.e. $t \in [a, b]$ (mixed state-control equalities);
  (iv) $x'(t) = K(t, x(t), u(t))$ for a.e. $t \in [a, b]$ (control equation);
  (v) $k_i(x(r_{i1}), \ldots, x(r_{i\varrho_i})) = 0$, $i = 1, \ldots, \kappa$ ("endpoint" equalities).

Here we have the following assumptions on the ranges of the occuring functions:

$$f_1, \ldots, f_\alpha, g_1, \ldots, g_\beta, k_1, \ldots, k_\kappa \in \mathbf{R}, \quad G \in \mathbf{R}^\gamma, \quad H \in \mathbf{R}^\delta, \quad K \in \mathbf{R}^n;$$

furthermore, $T_i \subset [a, b]^{\vartheta_i}$, $(i = 1, \ldots, \alpha)$, $S_i \subset [a, b]^{\pi_i}$, $(i = 1, \ldots, \beta)$ are compact subsets, Γ is a closed convex set with nonempty interior in $\mathbf{R}^\gamma$ and $r_i := (r_{i1}, \ldots, r_{i\varrho_i}) \in [a, b]^{\varrho_i}$ for $(i = 1, \ldots, \kappa)$ are given fixed points.

A pair (x, u) is an *arc* if $x: [a, b] \to \mathbf{R}^n$ is an absolutely continuous function and $u: [a, b] \to R^m$ is a bounded measurable function. An arc (x, u) is *admissible* for the problem $(\widehat{\mathscr{P}})$ if the conditions listed in (i)–(v) above are satisfied.

An arc $(\hat{x}, \hat{u})$ is called a *(local) solution* for $(\widehat{\mathscr{P}})$ if it is admissible and there exists $\varepsilon > 0$ such that $F(\hat{x}, \hat{u}) \le F(x, u)$ whenever (x, u) is admissible and

(14) $|x(t) - \hat{x}(t)| < \varepsilon, \qquad |u(t) - \hat{u}(t)| < \varepsilon$

hold for a.e. $t \in [a, b]$.

For practical reasons, we introduce the following notations. If T is a subset of $[a, b]^d$ and $t = (t_1, \ldots, t_d) \in T$, then $x(t)$ means the d-tuple $(x(t_1), \ldots, x(t_d))$, which can also be interpreted as an element of $\mathbf{R}^{n \times d}$, the space of the $n \times d$-matrices. Thus the objective function can be rewritten as

$$F(x, u) = \max_{1 \le i \le \alpha} \sup_{t \in T_i} f_i(t, x(t)).$$

The dual space of $\mathbf{R}^{n \times d}$ will be identified with $\mathbf{R}^{d \times n}$, the duality pairing is then defined by

$$\langle \alpha, x \rangle = \operatorname{tr} \alpha \cdot x \quad \text{for } x \in \mathbf{R}^{n \times d}, \ \alpha \in \mathbf{R}^{d \times n}.$$

(Here tr stands for the trace of a square matrix, i.e., for the sum of the diagonal elements.)

The constraint (i) reduces to $g_i(s, x(s)) \leq 0$ for $s \in S_i$, $(i = 1, \dots, \beta)$. The endpoint equality constraints are now of the form $k_i(x(r_i)) = 0$, $(i = 1, \dots, \kappa)$. If T is a subset of $[a, b]^d$ and $\hat{w}: [a, b] \to \mathbf{R}^\omega$ is an arbitrary function, then the ε-*tube* on T around $\hat{w}$ is the set

$$\mathscr{T}_\varepsilon(\hat{w}; T) = \{(t, w) \in T \times \mathbf{R}^{\omega \times d} : |w - \hat{w}(t)| < \varepsilon \text{ for } t \in T\}.$$

When $T = \{t\}$ is a singleton, then $\{w: (t, w) \in \mathscr{T}_\varepsilon(\hat{w}; \{t\})\}$ will be denoted by $\mathscr{T}_\varepsilon(\hat{w}(t))$.

Denote by $\mathscr{L}$ the class of Lebesgue measurable subsets in $[a, b]$, and by $\mathscr{B}$ the class of Borel measurable subsets in a metric space.

An arc $(\hat{x}, \hat{u})$ is *regular* for $(\widehat{\mathscr{P}})$ if there exists an $\varepsilon > 0$ such that the following conditions are satisfied:

- f_i is defined in $\mathscr{T}_\varepsilon(\hat{x}; T_i)$, the mapping $t \to f_i(t, x)$ is upper semi-continuous on $\{t \in T_i | x \in \mathscr{T}_\varepsilon(\hat{x}(t))\}$ for fixed $x \in \mathbf{R}^{n \times \vartheta_i}$; furthermore, there exists a constant C such that $|f_i(t, x') - f_i(t, x'')| \leq C|x' - x''|$, whenever $(t, x'), (t, x'') \in \mathscr{T}_\varepsilon(\hat{x}; T_i)$, $(i = 1, \dots, \alpha)$;
- g_i is defined in $\mathscr{T}_\varepsilon(\hat{x}; S_i)$, the mapping $s \to g_i(s, x)$ is upper semi-continuous on $\{s \in S_i | x \in \mathscr{T}_\varepsilon(\hat{x}(s))\}$ for fixed $x \in \mathbf{R}^{n \times \pi_i}$; furthermore, there exists a constant C such that $|g_i(s, x') - g_i(s, x'')| \leq C|x' - x''|$, whenever $(s, x'), (s, x'') \in \mathscr{T}_\varepsilon(\hat{x}; S_i)$, $(i = 1, \dots, \beta)$;
- G and H are $\mathscr{L} \times \mathscr{B} \times \mathscr{B}$-measurable functions defined in $\mathscr{T}_\varepsilon(\hat{x}, \hat{u}; [a, b])$; for a.e. $t \in [a, b]$, the map $(x, u) \mapsto L(t, x, u) := (G(t, x, u), H(t, x, u))$ is strictly Fréchet differentiable at $(\hat{x}(t), \hat{u}(t))$; moreover, $\forall \ \bar{\varepsilon} > 0 \ \exists \delta > 0$ such that, for a.e. t,

$$|L(t, z_1) - L(t, z_2) - \hat{L}'(t)(z_1 - z_2)| \leq \bar{\varepsilon}|z_1 - z_2|$$

if $|z_i - (\hat{x}(t), \hat{u}(t))| \leq \delta$, $(i = 1, 2)$. Furthermore, it is also assumed that the functions $\hat{G}, \hat{G}_x, \hat{G}_u, \hat{H}, \hat{H}_x$, and $\hat{H}_u$ are bounded;
- Let Γ be the set of those Lebesgue measurable functions $v \in \mathbf{L}_y^{(\infty)}[a, b]$ that satisfy $v(t) \in \Gamma$ for a.e. $t \in [a, b]$, and $\Gamma_{\hat{G}}$ be the closed cone $\overline{\operatorname{cone}}(\Gamma - \hat{G})$. Furthermore, let $\Lambda(\Gamma_{\hat{G}}) := \Gamma_{\hat{G}} \cap (-\Gamma_{\hat{G}})$. We assume that G and H have the following property: There exist bounded Lebesgue measurable functions $v_1, \dots, v_q \in \Lambda(\Gamma_{\hat{G}})$ such that

$$\det[J(t) \cdot J^T(t)] \geq \tau > 0$$

for a.e. $t \in [a, b]$, where the $(\gamma + \delta) \times (m + q)$-matrix valued function J is defined by

$$J(t) := \begin{pmatrix} \hat{G}_u(t) & v_1(t) & \cdots & v_q(t) \\ \hat{H}_u(t) & 0 & \cdots & 0 \end{pmatrix}.$$

(The notation $(\ )^T$ stands for transposition.)

- K is a $\mathscr{L} \times \mathscr{B} \times \mathscr{B}$-measurable function defined in $\mathscr{T}_\varepsilon(\hat{x}, \hat{u}; [a, b])$; for a.e. $t \in [a, b]$, the map $(x, u) \mapsto K(t, x, u)$ is strictly Fréchet differentiable at $(\hat{x}(t), \hat{u}(t))$. Moreover, $\forall \bar{\varepsilon} > 0 \, \exists \delta > 0$ and $\exists \Psi \colon [a, b] \to [0, \infty]$ with $\int_a^b \Psi(t)\,dt \le \bar{\varepsilon}$ such that, for a.e. t,

$$|K(t, z_1) - K(z, z_2) - \widehat{K}'(t)(z_1 - z_2)| \le \Psi(t)|z_1 - z_2|$$

  if $|z_i - (\hat{x}(t), \hat{u}(t))| \le \delta$, $(i = 1, 2)$. Furthermore, it is also assumed that $\widehat{K}$, $\widehat{K}_x$, and $\widehat{K}_u$ are integrable functions;
- k_i is defined in $\mathscr{T}_\varepsilon(\hat{x}(r_i))$; furthermore, it is strictly Fréchet differentiable at $\hat{x}(r_i)$ for $(i = 1, \ldots, \kappa)$.

Having the Lipschitz property of f_i, g_i, the Clarke's generalized derivatives $\hat{f}_{i[T_i]}^\circ$ and $\hat{g}_{i[S_i]}^\circ$ also exist. The functional $x \mapsto \hat{f}_{i[T_i]}^\circ(t; x)$ is sublinear on $\mathbf{R}^{n \times \vartheta_i}$; Therefore, by the Hahn-Banach theorem, the set of matrices $\alpha_i \in \mathbf{R}^{\vartheta_i \times n}$ such that

$$\langle \alpha_i, x \rangle = \mathrm{tr}(\alpha_i \cdot x) \le \hat{f}_{i[T_i]}^\circ(t; x) \quad \text{for } x \in \mathbf{R}^{n \times \vartheta_i}$$

is nonempty compact and convex. We shall denote this set by $\partial_{[T_i]}\hat{f}_i(t)$. One can see that this set can also be obtained in the following way:

$$\partial_{[T_i]}\hat{f}_i(t) = \mathrm{co}\left\{ \lim_{j \to \infty} \left(\frac{\partial f_i}{\partial x_{\lambda, \mu}}(t^j, x^j) \right)_{\substack{1 \le \lambda \le \vartheta_i \\ 1 \le \mu \le n}} \in \mathbf{R}^{\vartheta_i \times n} \colon t^j \to t \text{ and } x^j \to \hat{x}(t) \right\},$$

where the limit is taken for those sequences for which the derivatives exist (cf. [6]).

In a similar way, $\partial_{[S_i]}\hat{g}_i(s)$ can also be defined.

Let $(\hat{x}, \hat{u})$ be a regular admissible arc for problem $(\widehat{\mathscr{P}})$. Then we shall also need the notion of critical directions. As arc (ξ, η) is called *critical* for $(\widehat{\mathscr{P}})$ at $(\hat{x}, \hat{u})$ if

- $\hat{f}_{i[T_i]}^\circ(t; \xi(t)) \le 0$, whenever $t \in T_i$, $\hat{f}_i(t) = \widehat{F}$, $(i = 1, \ldots, \alpha)$;
- $\hat{g}_{i[S_i]}^\circ(s; \xi(s)) \le 0$, whenever $s \in S_i$, $\hat{g}_i(s) = 0$, $(i = 1, \ldots, \beta)$;
- $\widehat{G}_x \xi + \widehat{G}_u \eta \in \Gamma_{\widehat{G}}$;
- $\widehat{H}_x(t)\xi(t) + \widehat{H}_u(t)\eta(t) = 0$ for a.e. $t \in [a, b]$;
- $\xi'(t) = \widehat{K}_x(t)\xi(t) + \widehat{K}_u(t)\eta(t)$ for a.e. $t \in [a, b]$;
- $\hat{k}_i'\xi(r_i) = 0$ for $i = 1, \ldots, \kappa$.

An arc (ξ, η) is called a *regular direction* for $(\widehat{\mathscr{P}})$ at $(\hat{x}, \hat{u})$ if

- $\hat{f}_{i[T_i]}^{\circ\circ}(t; \xi(t))$ and $\hat{\sigma}_{[f_i]}(t; \xi(t))$ are finite for $t \in T_i$, $(i = 1, \ldots, \alpha)$;
- $\hat{g}_{i[S_i]}^{\circ\circ}(s; \xi(s))$ and $\hat{\sigma}_{[g_i]}(s; \xi(s))$ are finite for $s \in S_i$, $(i = 1, \ldots, \beta)$;
- $\Gamma^\circ(\widehat{G}, \widehat{G}_x \xi + \widehat{G}_u \eta) \ne \varnothing$;
- The mappings $(x, u) \mapsto L(t, x, u) := (G(t, x, u), H(t, x, u))$ and $(x, u) \mapsto K(t, x, u)$ are twice strongly directionally differentiable at $(\hat{x}(t), \hat{u}(t))$ in the direction $(\xi(t), \eta(t))$ for a.e. $t \in [a, b]$. Furthermore, the function

$$t \mapsto \frac{L(t, \hat{x}(t) + \varepsilon\xi(t), \hat{u}(t) + \varepsilon\eta(t)) - \widehat{L}(t) - \varepsilon\widehat{L}'(t)(\xi(t), \eta(t))}{\varepsilon^2}$$

tends to $\widehat{L}''(t; \xi(t), \eta(t))$ in L^∞-norm as $\varepsilon \to 0$. Analogously, the function

$$t \mapsto \frac{K(t, \hat{x}(t) + \varepsilon\xi(t), \hat{u}(t) + \varepsilon\eta(t)) - \widehat{K}(t) - \varepsilon\widehat{K}'(t)(\xi(t), \eta(t))}{\varepsilon^2}$$

goes to $\widehat{K}''(t; \xi(t), \eta(t))$ in L^1-norm. (It follows from these uniform convergences that $t \mapsto (\widehat{G}''(t; \xi(t), \eta(t)), \widehat{H}''(t; \xi(t), \eta(t)))$ is bounded and $t \mapsto \widehat{K}''(t; \xi(t), \eta(t))$ is integrable.)

- k_i is twice strongly directionally differentiable at $\hat{x}(r_i)$ in the direction $\xi(r_i)$ for $i = 1, \ldots, \kappa$.

In order to formulate the main result of this section we have to introduce the following notation: If $\Omega \subset [a, b]$ then the extended characteristic function $\chi_\Omega: [a, b]^d \to \mathbf{R}^{1 \times d}$ is defined by

$$\chi_\Omega(t_1, \ldots, t_d) := (\chi_\Omega(t_1), \ldots, \chi_\Omega(t_d)),$$

where χ_Ω denotes the characteristic function of the set Ω.

Define the Hamiltonian of the problem $(\widehat{\mathscr{P}})$ by

$$\mathscr{H}(t, x, u, p, \varphi, \psi) := p^T K(t, x, u) + \varphi^T G(t, x, u) + \psi^T H(t, x, u).$$

Now we are able to formulate

Theorem 4. *Let the arc $(\hat{x}, \hat{u})$ be a regular solution of the problem $(\widehat{\mathscr{P}})$. Then, for every regular critical arc (ξ, η), there exist Borel regular measures*

$$\mu_1 \in \mathscr{M}(T_1), \ldots, \mu_\alpha \in \mathscr{M}(T_\alpha), \qquad \nu_1 \in \mathscr{M}(S_1), \ldots, \nu_\beta \in \mathscr{M}(S_\beta),$$

two integrable functions $\varphi: [a, b] \to \mathbf{R}^\gamma$ and $\psi: [a, b] \to \mathbf{R}^\delta$, a function $p: [a, b] \to \mathbf{R}^n$ of bounded variation which is right continuous on $[a, b[$ and $p(b) = 0$, and $\lambda \in \mathbf{R}^\kappa$ such that at least one of them is different from zero; furthermore, there exist Borel measurable functions $\alpha_i: T_i \to \mathbf{R}^{\vartheta_i \times n}$, $\beta_i: S_i \to \mathbf{R}^{\pi_i \times n}$ with the selection property

$$\alpha_i(t) \in \partial_{[T_i]} \hat{f}_i(t), \quad \text{for } t \in T_i \quad \text{and} \quad \beta_i(s) \in \partial_{[S_i]} \hat{g}_i(s), \quad \text{for } s \in S_i$$

such that the following relations hold:

$$(15) \quad \begin{aligned} &\operatorname{supp} \mu_1 \subset \{t \in T_1 | \hat{f}_1(t) = \widehat{F}\}, \ldots, \operatorname{supp} \mu_\alpha \subset \{t \in T_\alpha | \hat{f}_\alpha(t) = \widehat{F}\}, \\ &\operatorname{supp} \nu_1 \subset \{s \in S_1 | \hat{g}_1(s) = 0\}, \ldots, \operatorname{supp} \nu_\beta \subset \{s \in S_\beta | \hat{g}_\beta(s) = 0\}, \end{aligned}$$

$$(16) \quad \begin{aligned} &\sup_{v \in \Gamma} \varphi^T(t)v \le \varphi^T(t)\widehat{G}(t) \quad \text{a.e. } t \in [a, b] \quad \text{and} \\ &\varphi^T(\widehat{G}_x \xi + \widehat{G}_u \eta) = 0 \quad \text{a.e.}, \end{aligned}$$

$$(17) \quad \begin{aligned} p^T(a) = &-\sum_{i=1}^{\alpha} \int_{T_i} \chi_{\{a\}}(t)\alpha_i(t) \, d\mu_i(t) - \sum_{i=1}^{\beta} \int_{S_i} \chi_{\{a\}}(s)\beta_i(s) \, d\nu_i(s) \\ &-\sum_{i=1}^{\kappa} \lambda_i \sum_{j=1}^{\varrho_i} \chi_{\{a\}}(r_{ij})\hat{k}_{ix_j}, \end{aligned}$$

and

$$(18) \quad p^T(\tau) = \sum_{i=1}^{\alpha} \int_{T_i} \chi_{]\tau,b]}(t)\alpha_i(t)\,d\mu_i(t) + \sum_{i=1}^{\beta} \int_{S_i} \chi_{]\tau,b]}(s)\beta_i(s)\,d\nu_i(s)$$

$$+ \int_{\tau}^{b} \widehat{\mathcal{H}_x}(t,p(t),\varphi(t),\psi(t))\,dt + \sum_{i=1}^{\kappa} \lambda_i \sum_{j=1}^{\varrho_i} \chi_{]\tau,b]}(r_{ij})\hat{k}_{ix_j}$$

for all $a \le \tau \le b$,

$$(19) \qquad\qquad \widehat{\mathcal{H}_u}(t,p(t),\varphi(t),\psi(t)) = 0$$

for a.e. $t \in [a,b]$, and

(20)

$$\sum_{i=1}^{\alpha} \int_{T_i} (\hat{f}^{\circ\circ}_{i[T_i]} - 2\hat{\sigma}_{[f_i]})(t;\xi(t))\,d\mu_i(t) + \sum_{i=1}^{\beta} \int_{S_i} (\hat{g}^{\circ\circ}_{i[S_i]} - 2\hat{\sigma}_{[g_i]})(s;\xi(s))\,d\nu_i(s)$$

$$+ \int_a^b (\widehat{\mathcal{H}''}(t,p(t),\varphi(t),\psi(t);\xi(t),\eta(t)))\,dt + \sum_{i=1}^{\kappa} \lambda_i \hat{k}_i''(\xi(r_i))$$

$$\ge 2\sup\left\{ \int_a^b \varphi^T(t)v(t)\,dt \,\Big|\, v \in \Gamma^{\circ}(\widehat{G},\widehat{G}_x\xi + \widehat{G}_u\eta) \right\}.$$

Remarks. 1. If $T_i = \{(a,b)\} \subset [a,b]^2$ in the objective function, then $\sup_{t \in T_i} f_i(t,x(t))$ becomes the "endpoint term" $f_i(a,b,x(a),x(b))$. If $T_i = [a,b]$, then we arrive at a minimax problem in the ith term of the objective function. If $T_i = [a,b]^2$ and $f_i(t_1,t_2,x(t_1),x(t_2)) = x(t_1) - x(t_2)$, then the ith term of the objective function becomes the diameter of the trajectory x.

2. Similar observations are valid for the inequality constraints $g_i(s,x(s)) \le 0$. These inequalities may take the form $g_i(a,b,x(a),x(b)) \le 0$, $g_i(s,x(s)) \le 0$, $(s \in [a,b])$ and $\operatorname{diam}(x) \le c$ as S_i and g_i are chosen properly.

3. When $G(t,x,u) = u$, we arrive at the case of a control set constraint: $u(t) \in \Gamma$ for a.e. $t \in [a,b]$.

4. Taking the limit $\tau \to b$ in (18), $p(b-)$ can be obtained:

$$p^T(b-) = \sum_{i=1}^{\alpha} \int_{T_i} \chi_{\{b\}}(t)\alpha_i(t)\,d\mu_i(t) + \sum_{i=1}^{\beta} \int_{S_i} \chi_{\{b\}}(s)\beta_i(s)\,d\nu_i(s)$$

$$+ \sum_{i=1}^{\kappa} \lambda_i \sum_{j=1}^{\varrho_i} \chi_{\{b\}}(r_{ij})\hat{k}_{ix_j}.$$

5. In the applications of Theorem 4 the determination of the sets $\Gamma_{\widehat{G}}$ and $\Gamma^{\circ}(\widehat{G},\widehat{G}_x\xi + \widehat{G}_u\eta)$ is a crucial point. There is no general method to solve this problem. In the next section we shall deal with the case $\Gamma =]-\infty,0]^{\gamma}$ and describe these two sets exactly.

The right-hand side of (20) is always nonpositive for all

$$v \in \Gamma^{\circ}(\widehat{G},\widehat{G}_x\xi + \widehat{G}_u\eta);$$

therefore, the optimal case is when it is zero. This happens when $v = 0 \in \Gamma^{\circ}(\widehat{G},\widehat{G}_x\xi + \widehat{G}_u\eta)$. However, in most of the cases this is not so; therefore, the

above term in (20) is essential. In the case $\Gamma =]-\infty, 0]^\gamma$ we shall point out that the optimal value 0 is attained.

6. As we shall see in the proof, the multipliers φ and ψ are uniquely determined by p, and as a consequence of this, they are essentially bounded measurable functions whenever $\widehat{K}_u$ is bounded.

Proof of Theorem 4. We introduce the following Banach spaces:

$$X := C_n[a, b] = \{x : [a, b] \to \mathbf{R}^n | x \text{ is continuous}\},$$
$$U := \mathbf{L}_m^{(\infty)}[a, b] = \{u : [a, b] \to \mathbf{R}^m | u \text{ is bounded and measurable}\},$$
$$V := \mathbf{L}_\gamma^{(\infty)}[a, b] = \{v : [a, b] \to \mathbf{R}^\gamma | v \text{ is bounded and measurable}\},$$
$$W := \mathbf{L}_\delta^{(\infty)}[a, b] = \{w : [a, b] \to \mathbf{R}^\delta | w \text{ is bounded and measurable}\},$$
$$Y := C_n[a, b]_0 = \{y : [a, b] \to \mathbf{R}^n | y \text{ is continuous and } y(a) = 0\}.$$

Let D be the set of those pairs $(x, u) \in X \times U$ whose graph is essentially contained in the tube $\mathscr{T}_{\varepsilon'}(\hat{x}, \hat{u}; [a, b])$ for some $\varepsilon' < \varepsilon$. Then D is an open set in $X \times U$. Define the mappings $\mathbf{f}_i : T_i \times D \to \mathbf{R}$, $\mathbf{g}_i : S_i \times D \to \mathbf{R}$, $\mathbf{G} : D \to V$, $\mathbf{H} : D \to W$, $\mathbf{K} : D \to Y$, and $\mathbf{k}_i : D \to \mathbf{R}$ by

$$\mathbf{f}_i(t, x, u) := f_i(t, x(t)) \quad \text{for } i = 1, \ldots, \alpha,$$
$$\mathbf{g}_i(s, x, u) := g_i(s, x(s)) \quad \text{for } i = 1, \ldots, \beta,$$
$$\mathbf{G}(x, u)(t) := G(t, x(t), u(t)),$$
$$\mathbf{H}(x, u)(t) := H(t, x(t), u(t)),$$
$$\mathbf{K}(x, u)(t) := \int_a^t K(\tau, x(\tau), u(\tau)) \, d\tau - x(t) + x(a),$$
$$\mathbf{k}_i(x, u) := k_i(x(r_i)) \quad \text{for } i = 1, \ldots, \kappa.$$

Then Γ is a closed convex subset of V with nonempty interior. One can observe that if (x, u) is an arc and $x'(t) = K(t, x(t), u(t))$ holds for a.e. t, then $\int_a^t K(\tau, x(\tau), u(\tau)) \, d\tau - x(t) + x(a) = 0$ is also valid for all t. Conversely, if $(x, u) \in D$ and the integral equation holds, then (x, u) is an arc (i.e., x is not only continuous, but it is absolutely continuous) and the differential equation is satisfied as well. Therefore, our problem $(\widehat{\mathscr{P}})$ is equivalent to $(\widetilde{\mathscr{P}})$ if $f_i, g_i, G_1, H, K, k, \Gamma_1$, and γ there are replaced by $\mathbf{f}_i, \mathbf{g}_i, \mathbf{G}, \mathbf{H}, \mathbf{K}, \mathbf{k}, \Gamma$ and 1, respectively.

To obtain the regularity of $(\hat{x}, \hat{u})$, we only need that $\mathbf{G}, \mathbf{H}, \mathbf{K}$, and $\mathbf{k}_i$ be strictly Fréchet differentiable, $\mathbf{K}_x$ and $\mathbf{K}_u$ be Fredholm and compact operators, respectively, and $\mathbf{H}_u$ obey the full rank property.

The strict Fréchet differentiability of $\mathbf{G}, \mathbf{H}, \mathbf{K}$ is a result of the regularity assumptions concerning G, H, K. Therefore, we have the following relations:

$$\widehat{\mathbf{G}}'(x, u)(t) = \widehat{G}_x(t)x(t) + \widehat{G}_u(t)u(t), \quad \widehat{\mathbf{H}}'(x, u)(t) = \widehat{H}_x(t)x(t) + \widehat{H}_u(t)u(t)$$

and

$$\widehat{\mathbf{K}}'(x, u)(t) = \int_a^t [\widehat{K}_x(\tau)x(\tau) + \widehat{K}_u(\tau)u(\tau)] \, d\tau - x(t) + x(a).$$

Since $\mathbf{K}_u$ is a Volterra integral operator, then it is compact (see [26]). On the other hand, operator $\mathbf{K}_x$ is the sum of a compact (Volterra integral) operator

and the operator $F: X \to Y$ defined by $Fx(t) := -x(t) + x(a)$, which is clearly a Fredholm operator. Therefore, by Lemma 3, $\mathbf{K}_x$ is also a Fredholm.

Analogously, we have

$$\hat{\mathbf{k}}_i'(x, u) = \hat{k}_i' x(r_i) \quad \text{for } i = 1, \dots, \kappa.$$

To prove the full rank property of $\mathbf{H}_u$, we construct a continuous right inverse. As we have assumed, the determinant of the matrix function $J \cdot J^T$ is bounded below by a positive constant. Therefore, for almost all $t \in [a, b]$, there exists the inverse of $J(t) \cdot J^T(t)$ and the components of the inverse matrix are essentially bounded measurable functions. Let $R(t) := J^T(t)[J(t) \cdot J^T(t)]^{-1}$. Then, obviously, $R(t)$ is a right inverse of $J(t)$ for a.e. $t \in [a, b]$; furthermore, the components of R are essentially bounded measurable functions, as well. Let us split $R(t)$ in the form

$$R(t) = \begin{pmatrix} A(t) & B(t) \\ C(t) & D(t) \end{pmatrix}.$$

where A, B, C, and D are $m \times \gamma$, $m \times \delta$, $q \times \gamma$, and $q \times \delta$ matrix valued functions on $[a, b]$, respectively. Now $B(t)$ is a right inverse of $\hat{H}_u(t)$ for almost all t. Define the operator $\mathbf{B}: W \to U$ by $\mathbf{B}w(t) := B(t)w(t)$, $(w \in W)$. Then, since the components of B are bounded measurable functions, $\mathbf{B}$ is a bounded linear operator from W to U; moreover, it is a right inverse of $\mathbf{H}_u$.

If (ξ, η) is a regular critical arc for $(\mathscr{P})$, then it is also a regular and critical direction for the modified problem $(\widetilde{\mathscr{P}})$. After a simple computation, one yields

$$\hat{\mathbf{G}}''(\xi, \eta)(t) = \hat{G}''(t; \xi(t), \eta(t)), \qquad \hat{\mathbf{H}}''(\xi, \eta)(t) = \hat{H}''(t; \xi(t), \eta(t)),$$

$$\hat{\mathbf{K}}''(\xi, \eta)(t) = \int_a^t \hat{K}''(\tau; \xi(\tau), \eta(\tau)) \, d\tau,$$

and

$$\hat{\mathbf{k}}_i''(\xi, \eta) = \hat{k}_i''(\xi(r_i)) \quad \text{for } i = 1, \dots, \kappa.$$

Thus, by Remark 5 following Theorem 3, we can find Lagrange multipliers

$$\mu_1 \in \mathscr{M}(T_1), \dots, \mu_\alpha \in \mathscr{M}(T_\alpha), \qquad \nu_1 \in \mathscr{M}(S_1), \dots, \nu_\beta \in \mathscr{M}(S_\beta),$$

and

$$v^* \in V^*, \quad w^* \in W^*, \, y^* \in Y^*, \quad (\lambda_1, \dots, \lambda_\kappa) \in \mathbf{R}^\kappa$$

such that at least one of them is different from zero, (16) and (17) are satisfied and

$$(21) \qquad v^*(v) \leq 0, \; (v \in \Gamma - \hat{\mathbf{G}}) \quad \text{and} \quad v^*(\hat{\mathbf{G}}_x \xi + \hat{\mathbf{G}}_u \eta) = 0 \quad \text{a.e.},$$

$$(22) \quad \sum_{i=1}^\alpha \int_{T_i} \hat{f}_{i[T_i]}^\circ(t; x(t)) \, d\mu_i(t) + \sum_{i=1}^\beta \int_{S_i} \hat{g}_{i[S_i]}^\circ(s; x(s)) \, d\nu_i(s)$$

$$+ v^*(\hat{\mathbf{G}}_x x) + w^*(\hat{\mathbf{H}}_x x) + y^*(\hat{\mathbf{K}}_x x) + \sum_{i=1}^\kappa \lambda_i \hat{k}_{ix} x(r_i) \geq 0 \quad \text{for } x \in X,$$

$$(23) \qquad v^*(\hat{\mathbf{G}}_u u) + w^*(\hat{\mathbf{H}}_u u) + y^*(\hat{\mathbf{K}}_u u) \geq 0 \quad \text{for } u \in U$$

and

$$\sum_{i=1}^{\alpha} \int_{T_i} (\hat{f}^{\circ\circ}_{i[T_i]} - 2\hat{\sigma}_{[f_i]})(t\,;\xi)\, d\mu_i(t)$$

(24)
$$+ \sum_{i=1}^{\beta} \int_{S_i} (\hat{g}^{\circ\circ}_{i[S_i]} - 2\hat{\sigma}_{[g_i]})(s\,;\xi)\, d\nu_i(s) + v^*(\widehat{\mathbf{G}}''(\xi,\eta))$$

$$- 2\delta^*(v^*\,;\mathbf{\Gamma}^{\circ}(\widehat{G}, \widehat{G}_x\xi + \widehat{G}_u\eta)) + w^*(\widehat{\mathbf{H}}''(\xi,\eta))$$

$$+ y^*(\widehat{\mathbf{K}}''(\xi,\eta)) + \sum_{i=1}^{\kappa} \lambda_i \hat{k}''_i(\xi(r_i)) \geq 0.$$

By the Riesz representation theorem on linear functionals, y^* can be obtained with the help of a bounded signed $\mathbf{R}^n$-valued Borel measure ζ on $[a,b]$ in the following form:

$$y^*(y) = \int_a^b y^T(t)\, d\zeta(t)\,;$$

furthermore, since the elements of Y vanish at a, hence $\zeta(\{a\}) = 0$. Define $p\colon [a,b] \to \mathbf{R}^n$ by

$$p(t) := \zeta(]t,b]).$$

The property $\zeta(\{a\}) = 0$ yields that p is continuous at a. Since $]t,b] = \bigcup_{i=1}^{\infty}]t_i,b]$ if t_i is a decreasing sequence with $\lim_{i\to\infty} t_i = t$, hence p is right continuous at each $t \in \,]a,b[$. Clearly $p(b) = \zeta(\varnothing) = 0$. The boundedness of ζ implies that p is of bounded variation. Furthermore, for integrable functions $L\colon [a,b] \to \mathbf{R}^n$, we have the following identity:

$$\int_a^b \left[\int_a^t L^T(\tau)\, d\tau\right] d\zeta(t) = \int_a^b p^T(t) L(t)\, dt.$$

(This can be checked by using the Fubini theorem.) Therefore, we get

(25) $$y^*(\widehat{\mathbf{K}}_x x) = \int_a^b p^T(t)\widehat{K}_x(t)x(t)\, dt - \int_a^b x^T(t)\, d\zeta(t) + p^T(a)x(a),$$

(26) $$y^*(\widehat{\mathbf{K}}_u u) = \int_a^b p^T(t)\widehat{K}_u(t)u(t)\, dt,$$

(27) $$y^*(\widehat{\mathbf{K}}''(\xi,\eta)) = \int_a^b p^T(t)\widehat{K}''(t\,;\xi(t),\eta(t))\, dt.$$

Our next aim is to prove that, as a consequence of the strong full rank property, v^* and w^* can be expressed in terms of integrable functions.

We show first that $\Lambda(\Gamma_{\widehat{G}}) \subset V$ is a modulus over the ring of bounded real valued measurable functions on $[a,b]$, i.e., for $v \in \Lambda(\Gamma_{\widehat{G}})$ and for every bounded measurable function $c\colon [a,b] \to \mathbf{R}$, the function $(cv)(t) := c(t)v(t)$ is again in $\Lambda(\Gamma_{\widehat{G}})$.

Let $c \in \mathbf{L}^{(\infty)}[a,b]$ and $v \in \Lambda(\Gamma_{\widehat{G}})$. The function c can be decomposed as $c = c_+ - c_-$, where c_+ and c_- are nonnegative bounded measurable functions. Since $v \in \Gamma_{\widehat{G}}$, hence there exist sequences $\tau_i \geq 0$ and $v_i \in \Gamma$ such that

$\lim_{i\to\infty} \tau_i(v_i(t) - \widehat{G}(t)) = v(t)$ for a.e. $t \in [a, b]$. Define

$$\overline{\tau}_i := \|c_+\|\tau_i \quad \text{and} \quad \overline{v}_i(t) := \frac{c_+(t)}{\|c_+\|}v_i(t) + \left(1 - \frac{c_+(t)}{\|c_+\|}\right)\widehat{G}(t).$$

Then, since Γ is convex, we have $\overline{v}_i \in \Gamma$ for all $i \in \mathbf{N}$. On the other hand

$$\lim_{i\to\infty} \overline{\tau}_i(\overline{v}_i(t) - \widehat{G}(t)) = c_+(t)\lim_{i\to\infty} \tau_i(v_i(t) - \widehat{G}(t)) = c_+(t)v(t)$$

for a.e. $t \in [a, b]$. Therefore, $c_+v \in \Gamma_{\widehat{G}}$. A similar argument shows that $c_-(-v) \in \Gamma_{\widehat{G}}$ also holds (since $-v \in \Gamma_{\widehat{G}}$). Thus we obtain $cv = c_+v - c_-v \in \Gamma_{\widehat{G}}$. Repeating the above argument with $-v$ instead of v, we get that $cv \in -\Gamma_{\widehat{G}}$, and hence $cv \in \Lambda(\Gamma_{\widehat{G}})$. This completes the proof of the modulus property of $\Lambda(\Gamma_{\widehat{G}})$.

As a consequence of (21), we have $v^*(v) = 0$ for all $v \in \Lambda(\Gamma_{\widehat{G}})$. On the other hand, the functions $v_1, \ldots, v_q$ are in $\Lambda(\Gamma_{\widehat{G}})$ (by the strong full rank property imposed on G and H). Thus, for all bounded measurable functions $c_1, \ldots, c_q: [a, b] \to \mathbf{R}$, we obtain

$$(28) \qquad\qquad v^*(c_1v_1 + \cdots + c_qv_q) = 0.$$

In (23) we may write equality instead of inequality, since the left-hand side is a linear functional of u. Adding (23) to (28) and using the description of y^*, we get

$$v^*(\widehat{G}_u u + c_1v_1 + \cdots + c_qv_q) + w^*(\widehat{H}_u u) = -\int_a^b p^T(t)\widehat{K}_u(t)u(t)\,dt$$

for $u \in U$ and $c_1, \ldots, c_q \in \mathbf{L}^{(\infty)}[a, b]$.

Using the right inverse function R of J, we make the following substitution:

$$\begin{pmatrix} u \\ c_1 \\ \vdots \\ c_q \end{pmatrix} := R \cdot \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \cdot \begin{pmatrix} v \\ w \end{pmatrix},$$

where $v \in V$ and $w \in W$ are arbitrary functions. Then we get

$$v^*(v) + w^*(w) = -\int_a^b p^T(t)\widehat{K}_u(t)[A(t)v(t) + B(t)w(t)]\,dt.$$

for $v \in V$ and $w \in W$. Define now $\varphi: [a, b] \to \mathbf{R}^\gamma$ and $\psi: [a, b] \to \mathbf{R}^\delta$ by

$$\varphi^T(t) := -p^T(t)\widehat{K}_u(t)A(t) \quad \text{and} \quad \psi^T(t) := -p^T(t)\widehat{K}_u(t)B(t).$$

Then these functions are clearly integrable and we have the following representations

$$(29) \qquad v^*(v) = \int_a^b \varphi^T(t)v(t)\,dt \quad \text{and} \quad w^*(w) = \int_a^b \psi^T(t)w(t)\,dt.$$

Note that when $\widehat{K}_u$ is essentially bounded so are φ and ψ.

Using (26) and (29), the equation (23) can be rewritten as

$$\int_a^b \widehat{\mathscr{H}}_u(t, p(t), \varphi(t), \psi(t))u(t)\,dt = 0$$

for all $u \in U$. This yields (19) at once.

The first inequality in (21) can be rephrased in the form

$$v^*(v) = \int_a^b \varphi^T(t)v(t)\,dt \le 0 \quad \text{for } v \in \Gamma - \widehat{G},$$

that is, v^* is a supporting functional of $\Gamma - \widehat{G}$. Using the result of [11] (Example 10.5, p. 76), this conditions, yields that, for a.e. $t \in [a, b]$,

$$\varphi^T(t)(v - \widehat{G}(t)) \le 0 \quad \text{for } v \in \Gamma.$$

Thus the first inequality of (16) is also proved. The domain of the latter inequality can be extended to the larger set $\Gamma_{\widehat{G}(t)} = \overline{\text{cone}}(\Gamma - \widehat{G}(t))$. On the other hand, the relation $\widehat{G}_x\xi + \widehat{G}_u\eta \in \Gamma_{\widehat{G}}$ yields

$$\widehat{G}_x(t)\xi(t) + \widehat{G}_u(t)\eta(t) \in \Gamma_{\widehat{G}(t)}$$

for a.e. $t \in [a, b]$. Therefore, we have

$$\varphi^T(t)(\widehat{G}_x(t)\xi(t) + \widehat{G}_u(t)\eta(t)) \le 0$$

for a.e. t. The second relation in (21) says that the integral of the left-hand side of this inequality is zero, hence the inequality sign here can be replaced by equality almost everywhere. Thus the second part of (16) is also proved.

The inequality stated in (20) follows from (24) if one makes use of (27) and (29) and the definition of the support function δ^*.

In the last part of the proof we have to interpret the inequality (22). Using (25) and (29), (22) reduces to
(30)

$$\sum_{i=1}^{\alpha} \int_{T_i} \hat{f}_{i[T_i]}^{\circ}(t; x(t))\,d\mu_i(t) + \sum_{i=1}^{\beta} \int_{S_i} \hat{g}_{i[S_i]}^{\circ}(s; x(s))\,d\nu_i(s) + \sum_{i=1}^{\kappa} \lambda_i \hat{k}_{ix} x(r_i)$$

$$+ \int_a^b \widehat{\mathscr{H}}_x(t, p(t), \varphi(t), \psi(t))x(t)\,dt - \int_a^b x^T(t)\,d\zeta(t) + p^T(a)x(a) \ge 0$$

for all $x \in X$.

The mappings $x \mapsto \hat{f}_{i[T_i]}^{\circ}(t, x(t))$ and $x \mapsto \hat{g}_{i[S_i]}^{\circ}(s, x(s))$ are sublinear functionals on $X = C_n[a, b]$ for all fixed $t \in T_i$ and $s \in S_i$, respectively. Therefore, using Lemma 1 (the disintegration theorem cited in Remark 3 of Theorem 1), we can find functions $\hat{\alpha}_i \colon T_i \times X \to \mathbf{R}$, $(i = 1, \dots, \alpha)$ and $\hat{\beta}_i \colon S_i \times X \to \mathbf{R}$, $(i = 1, \dots, \beta)$ that are Borel measurable in the first and linear in the second variable; furthermore, the following relations are also satisfied

(31) $\hat{\alpha}_i(t, x) \le \hat{f}_{i[T_i]}^{\circ}(t, x(t))$, $(t \in T_i)$, $\hat{\beta}_i(s, x) \le \hat{g}_{i[S_i]}^{\circ}(s, x(s))$, $(s \in S_i)$,

and
(32)

$$\sum_{i=1}^{\alpha} \int_{T_i} \hat{\alpha}_i(t, x)\,d\mu_i(t) + \sum_{i=1}^{\beta} \int_{S_i} \hat{\beta}_i(s, x)\,d\nu_i(s) + \sum_{i=1}^{\kappa} \lambda_i \hat{k}_{ix} x(r_i)$$

$$+ \int_a^b \widehat{\mathscr{H}}_x(t, p(t), \varphi(t), \psi(t))x(t)\,dt - \int_a^b x^T(t)\,d\zeta(t) + p^T(a)x(a) \ge 0$$

for all $x \in X$. The right-hand side of (32) is a linear functional x; therefore, the inequality sign can be replaced by equality here.

Let now $t = (t_1, \dots, t_{\vartheta_i}) \in T_i$ be fixed. The Riesz representation theorem yields an integral form of the linear functional $\hat{\alpha}_i(t, \cdot)$ with a vector valued Borel measure. However, it follows from (31) that $\hat{\alpha}_i(t, x)$ vanishes if the support of the function x does not intersect $\{(t_1, \dots, t_{\vartheta_i})\}$. Thus the representing measure must be supported in this one point set. The integral representation of $\hat{\alpha}_i(t, \cdot)$ is therefore nothing else than a linear function of $(x(t_1), \dots, x(t_{\varphi_i}))$ i.e. there exists $\alpha_i(t) \in \mathbf{R}^{\vartheta_i \times n}$ such that

$$\hat{\alpha}_i(t, x) = \langle \alpha_i(t), x(t) \rangle = \operatorname{tr}(\alpha_i(t) \cdot x(t))$$

on the other hand, by (31), we have $\alpha(t) \in \partial_{[T_i]}\hat{f}_i(t)$. The Borel measurability of α_i is obvious.

A similar argument shows that $\hat{\beta}_i(s, x)$ can also be expressed in the form

$$\hat{\beta}_i(s, x) = \beta_i(s)x(s),$$

where $\beta_i(s) \in \partial_{[S_i]}\hat{g}_i(s)$ holds for all $s \in S_i$.

Now (32) reduces to
(33)

$$\sum_{i=1}^{\alpha} \int_{T_i} \langle \alpha_i(t), x(t) \rangle \, d\mu_i(t) + \sum_{i=1}^{\beta} \int_{S_i} \langle \beta_i(s), x(s) \rangle \, d\nu_i(s) + \sum_{i=1}^{\kappa} \lambda_i \hat{k}_{ix} x(r_i)$$

$$+ \int_a^b \widehat{\mathscr{H}}_x(t, p(t), \varphi(t), \psi(t))x(t) \, dt - \int_a^b x^T(t) \, d\zeta(t) + p^T(a)x(a) = 0$$

for $x \in C_n[a, b]$.

Observe that if a sequence of functions $x_i \in X$ tends pointwise to a function $x: [a, b] \to \mathbf{R}^n$ such that these functions have a common constant upper bound, then (33) remains valid for the limit function x as well. Therefore, we can make the following substitution in (33): $x(t) := \overline{x}\chi_\Omega(t)$, where $\Omega \subset [a, b]$ is an arbitrary subinterval and $\overline{x} \in \mathbf{R}^n$ is a fixed element. Observe that by the known properties of the trace function,

$$\langle \alpha_i(t), x(t) \rangle = \operatorname{tr}(\alpha_i(t) \cdot \overline{x}\chi_\Omega(t)) = \operatorname{tr}(\chi_\Omega(t)\alpha_i(t) \cdot \overline{x}) = \chi_\Omega(t)\alpha_i(t) \cdot \overline{x}$$

for $t \in T_i$, $i = 1, \dots, \alpha$, and similarly

$$\langle \beta_i(s), x(s) \rangle = \chi_\Omega(s)\beta_i(s) \cdot \overline{x} \quad \text{for } s \in S_i, \ i = 1, \dots, \beta.$$

Now taking $\Omega =]\tau, b]$, where $\tau \in [a, b]$, and using

$$p(b) = 0, \quad \int_a^b x^T(t) \, d\zeta(t) = [p(\tau) - p(b)]^T \overline{x} \quad \text{and} \quad p^T(a)x(a) = 0,$$

the equation (33) yields (18).

Substituting $\Omega = \{a\}$, (33) leads to (17). Here one should also use the continuity of p at the endpoint $t = a$, whence

$$\int_a^b x^T(t) \, d\zeta(t) = [p(a) - p(a+)]^T \overline{x} = 0 \quad \text{and} \quad p^T(a)x(a) = p^T(a)\overline{x}.$$

Thus we arrive at (17). Therefore, the proof is complete. $\quad\square$

Remarks. 1. As we have seen in the proof, the multipliers φ and ψ are completely determined by p in the form $\varphi^T(t) := -p^T(t)\widehat{K}_u(t)A(t)$ and $\psi^T(t) := -p^T(t)\widehat{K}_u(t)B(t)$, where A and B are submatrices of the right inverse R of J. This property is essential, when one has to find the solutions of problem $(\widehat{\mathscr{P}})$.

2. Let $\mathscr{V}$ be the $\gamma \times q$-matrix function defined by $\mathscr{V} := (v_1, \ldots, v_q)$ and let

$$\mathscr{R} := \begin{pmatrix} A & B \\ -\mathscr{V}C & -\mathscr{V}D \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & -\mathscr{V} \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where A, B, C, and D are the matrix functions given via R in the proof. Then $\mathscr{R}$ is the right inverse of the matrix function

$$S = \begin{pmatrix} \widehat{G}_u & -I \\ \widehat{H}_u & 0 \end{pmatrix},$$

since R is the right inverse of J and $S\mathscr{R} = JR$. On the other hand, the operator from $V \times W$ to $U \times V$ defined by $\mathscr{R}$ has its image in $U \times \Lambda(\Gamma_{\widehat{G}})$ by the modulus property of $\Lambda(\Gamma_{\widehat{G}})$ and the following identity:

$$\mathscr{R}\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & -\mathscr{V} \end{pmatrix}\begin{pmatrix} u \\ c_1 \\ \vdots \\ c_q \end{pmatrix} = \begin{pmatrix} u \\ -c_1 v_1 - \cdots - c_q v_q \end{pmatrix}.$$

Hence, the determinant condition imposed on J implies the strong full rank property displayed in Section 4 after the proof of Theorem 3.

6. APPLICATION TO AN OPTIMAL CONTROL PROBLEM WITH VARIOUS CONSTRAINTS

We formulate two propositions that are important to handle the optimal control problem below.

Lemma 6. *Let*

$$\Gamma = \{u \in \mathbf{L}^{(\infty)}[a, b] | u(t) \le 0 \text{ for a.e. } t \in [a, b]\}$$

and $u \in \Gamma$. Then in order that a bounded measurable function $v: [a, b] \to \mathbf{R}$ be in $\Gamma_u = \overline{\text{cone}}(\Gamma - u)$ it is necessary and sufficient that the following condition be satisfied:

For all $\varepsilon > 0$ there exists $\lambda > 0$ such that

(34) $v(t) + \lambda u(t) \le \varepsilon \quad \text{for a.e. } t \in [a, b].$

If (34) holds, then

(35) $v(t) \le 0, \quad \text{whenever } u(t) = 0 \text{ for a.e. } t \in [a, b].$

Conversely, if (35) and in addition

(36) $\dfrac{v^2}{u}$ *is essentially bounded on $\{t \in [a, b] | u(t) < 0, v(t) > 0\}$,*

then $v \in \Gamma_u$ is valid.

Proof. By definition, $v \in \overline{\text{cone}}(\Gamma - u)$ if and only if for all $\varepsilon > 0$ there exists $\lambda > 0$ and $w \in \Gamma$ such that

$$\|v - \lambda(w - u)\| \le \varepsilon,$$

i.e.,

$$v(t) + \lambda u(t) - \varepsilon \le \lambda w(t) \le v(t) + \lambda u(t) + \varepsilon \quad \text{for a.e. } t \in [a, b].$$

In order that w could be chosen in Γ it is necessary and sufficient that the function on the left-hand side be nonpositive. Thus the first statement of the lemma is verified.

To see that (35) and (36) is already sufficient, choose $\varepsilon > 0$ arbitrarily. We have to construct λ such that (34) is valid. (34) is obvious if $v(t) \le 0$. The set $\{t: v(t) > 0, u(t) = 0\}$ is of measure zero by (35); therefore, we have to prove (34) only for those values of t for which $u(t) < 0$ and $v(t) > 0$. By (36), $-v^2/u$ is essentially bounded above on this set. Denote by K an upper bound and let $\lambda := K/4\varepsilon$. Then

$$v(t) + \lambda u(t) \le -\frac{v^2(t)}{4\lambda u(t)} = -\frac{v^2(t)}{u(t)} \cdot \frac{\varepsilon}{K} \le \varepsilon$$

for a.e. t. $\quad\square$

Remarks. 1. One can see that the inclusion $v \in \text{cone}(\Gamma - u)$ is satisfied if and only if (35) and

$$\frac{v}{u} \text{ is essentially bounded on } \{t \in [a, b] | u(t) < 0, v(t) > 0\}.$$

This condition obviously implies (36), but not conversely; therefore, the conditions (35) and (36) describe a cone bigger than $\text{cone}(\Gamma - u)$ and smaller than $\overline{\text{cone}}(\Gamma - u)$. As we shall see below, (36) plays an essential role in determining those directions v for which $\Gamma°(u, v)$ is not empty.

The condition (35) alone does not describe Γ_u, the set defined by (35) is much bigger than Γ_u.

2. Taking the function $v = r\sqrt{|u|}$, $r \in \mathbf{R}$, it is easily seen that (35) and (36) are satisfied. Therefore, $\pm\sqrt{|u|} \in \Gamma_u$, i.e., $\sqrt{|u|} \in \Lambda(\Gamma_u)$.

Lemma 7. *Let $u \in \Gamma$ and $v \in \mathrm{L}^{(\infty)}[a, b]$. Then in order that $\Gamma°(u, v)$ be nonempty it is necessary and sufficient that the conditions (35) and (36) of Lemma 6 be satisfied.*

If $\varphi: [a, b] \to \mathbf{R}$ is a nonnegative integrable function such that $\varphi(t)u(t) = 0$ and $\varphi(t)v(t) = 0$ for a.e. t and the conditions (35) and (36) are satisfied, then

$$(37) \qquad \sup\left\{\int_a^b \varphi(t)w(t)\,dt \,|\, w \in \Gamma°(u, v)\right\} = 0.$$

Proof. By the definition of $\Gamma°$, a bounded measurable function w is in $\Gamma°(u, v)$ if and only if there exists $\bar{\varepsilon} > 0$ such that, for $0 < \varepsilon < \bar{\varepsilon}$,

$$u(t) + \varepsilon v(t) + \varepsilon^2(w(t) + \bar{\varepsilon}) \le 0 \quad \text{for a.e. } t \in [a, b].$$

In other words, for each $0 < \varepsilon < \bar{\varepsilon}$, there exists a set $A_\varepsilon \subset [a, b]$ of full measure such that

$$u(t) + \varepsilon v(t) + \varepsilon^2(w(t) + \bar{\varepsilon}) \le 0 \quad \text{for } t \in A_\varepsilon.$$

However, this is equivalent to the existence of a set A of full measure such that

$$u(t) + \varepsilon v(t) + \varepsilon^2(w(t) + \bar{\varepsilon}) \le 0 \quad \text{for } t \in A, \varepsilon \in \,]0, \bar{\varepsilon}[.$$

For, let $\{\varepsilon_n | n \in \mathbf{N}\}$ be a dense subset of the interval $]0, \bar{\varepsilon}[$ and denote by A the set $\bigcap_{n=1}^{\infty} A_{\varepsilon_n}$. Then A is of full measure and

$$u(t) + \varepsilon_n v(t) + \varepsilon_n^2 (w(t) + \bar{\varepsilon}) \leq 0 \quad \text{for } t \in A, \, n \in \mathbf{N}.$$

Since the sequence ε_n forms a dense subset in $]0, \bar{\varepsilon}[$, hence, the above statement follows. Therefore, a function w is in $\Gamma^\circ(u, v)$ if and only if there exists $\bar{\varepsilon} > 0$ and a set $A \in [a, b]$ of full measure such that,

$$w(t) \leq w_{\bar{\varepsilon}}(t) := -\bar{\varepsilon} + \inf_{0 < \varepsilon < \bar{\varepsilon}} -\frac{u(t) + \varepsilon v(t)}{\varepsilon^2} \quad \text{for } t \in A.$$

An easy calculation for $w_{\bar{\varepsilon}}$ yields:

$$w_{\bar{\varepsilon}}(t) = \begin{cases} -\infty, & \text{if } u(t) = 0, \, v(t) > 0, \\ -\bar{\varepsilon} + \frac{v^2(t)}{4u(t)}, & \text{if } u(t) < 0, \, 2u(t) + \bar{\varepsilon}v(t) > 0, \\ -\bar{\varepsilon} - \frac{u(t) + \bar{\varepsilon}v(t)}{\bar{\varepsilon}^2}, & \text{if } 2u(t) + \varepsilon v(t) \leq 0. \end{cases}$$

Assume $\Gamma^\circ(u, v) \neq \varnothing$, then there exists a bounded measurable function w such that, for some $\bar{\varepsilon}$, $w(t) \leq w_{\bar{\varepsilon}}(t)$ holds almost everywhere. Then we have $-\|w\| \leq w_{\bar{\varepsilon}}(t)$ a.e. Thus, the set of those points t where $u(t) = 0$ and $v(t) > 0$ is of measure 0. Furthermore, for almost all t,

$$4(\bar{\varepsilon} - \|w\|) \leq v^2(t)/u(t) \leq 0 \quad \text{if } u(t) < 0, \, v(t) > 0, \text{ and } 2u(t) + \bar{\varepsilon}v(t) > 0.$$

On the other hand,

$$-2v(t)/\bar{\varepsilon} \leq v^2(t)/u(t) \leq 0 \quad \text{if } u(t) < 0, \, v(t) > 0, \text{ and } 2u(t) + \bar{\varepsilon}v(t) \leq 0.$$

Hence (35) and (36) hold true.

Conversely, assume now that (35) and (36) are satisfied. Then, for all positive $\bar{\varepsilon}$, the function $w_{\bar{\varepsilon}}$ is essentially bounded. Thus, $w_{\bar{\varepsilon}} \in \Gamma^\circ(u, v)$ for all $\bar{\varepsilon}$. Therefore, $\Gamma^\circ(u, v)$ is obviously not empty.

To show the last assertion of the lemma, note that for any w in $\Gamma^\circ(u, v)$ there exists $\bar{\varepsilon}$ such that

$$\int_a^b \varphi(t)w(t)\,dt \leq \int_a^b \varphi(t)w_{\bar{\varepsilon}}(t)\,dt = \int_a^b -\bar{\varepsilon}\varphi(t)\,dt \leq 0.$$

Thus the left-hand side of (37) is nonpositive. On the other hand, we have $w_{\bar{\varepsilon}} \in \Gamma^\circ(u, v)$ and

$$\int_a^b \varphi(t)w_{\bar{\varepsilon}}(t)\,dt = -\bar{\varepsilon}\int_a^b \varphi(t)\,dt \to 0 \quad \text{as } \bar{\varepsilon} \to 0,$$

from which (37) follows. $\square$

Remark. If the set Γ is the set of nonpositive continuous functions in $\mathscr{C}[a, b]$ and u and v are also continuous functions, then using the above approach, one can see that $\Gamma^\circ(u, v)$ is not empty if and only if (35) and (36) hold true; however, the analogue of (37) is not valid in this case. The reason is that the members of the optimizing sequence $w_{\bar{\varepsilon}}$ found in the above proof are not continuous functions. Therefore they cannot be used to approach the supremum. In fact, for any $w \in \Gamma^\circ(u, v)$ the relation $w(t) \leq \sigma_{u,v}(t)$ can be proved; therefore, the right-hand side of (37) should be replaced by $\int_a^b \varphi(t)\sigma_{u,v}(t)\,dt$ in this case.

Consider the following optimal control problem $(\mathscr{O}\mathscr{C})$.

Minimize $f(x(a), x(b))$ subject to

  (i) $g(t, x(t)) \leq 0$ for $t \in [a, b]$;
  (ii) $G(t, x(t), u(t)) \leq 0$ for a.e. $t \in [a, b]$;
  (iii) $H(t, x(t), u(t)) = 0$ for a.e. $t \in [a, b]$;
  (iv) $x'(t) = K(t, x(t), u(t))$ for a.e. $t \in [a, b]$;
  (v) $l(x(a), x(b)) \leq 0$;
  (vi) $k(x(a), x(b)) = 0$,

where $x: [a, b] \to \mathbf{R}^n$ is absolutely continuous, $u: [a, b] \to \mathbf{R}^m$ is essentially bounded, and the ranges of f, g, G, H, K, l, and k are respectively in $\mathbf{R}$, $\mathbf{R}^\beta$, $\mathbf{R}^\gamma$, $\mathbf{R}^\delta$, $\mathbf{R}^n$, $\mathbf{R}^r$, and $\mathbf{R}^\kappa$.

If (x, u) satisfy (i)–(vi), then it is said to be admissible. An arc $(\hat{x}, \hat{u})$ is *regular* if for some $\varepsilon > 0$:

- f is locally Lipschitz on $\mathscr{T}_\varepsilon(\hat{x}(a)) \times \mathscr{T}_\varepsilon(\hat{x}(b))$;
- $g = (g_1, \ldots, g_\beta)$ is defined in $\mathscr{T}_\varepsilon(\hat{x}; [a, b])$, $g_i(\cdot, x)$ $(i = 1, \ldots, \beta)$ is upper semicontinuous on $\{s \in [a, b] | x \in \mathscr{T}_\varepsilon(\hat{x}(s))\}$ for fixed $x \in \mathbf{R}^n$, and $g_i(s, \cdot)$ is Lipschitz on $\mathscr{T}_\varepsilon(\hat{x}; [a, b])$ uniformly in s;
- G, H, and K satisfy the regularity conditions of the previous section;
- For the matrix function

$$J(t) := \begin{pmatrix} \widehat{G}_{1u}(t) & \sqrt{-\widehat{G}_1(t)} & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ \widehat{G}_{\gamma u}(t) & 0 & \cdots & \sqrt{-\widehat{G}_\gamma(t)} \\ \widehat{H}_u(t) & 0 & \cdots & 0 \end{pmatrix}$$

  we have

$$\det[J(t) \cdot J^T(t)] \geq \tau > 0$$

  for a.e. $t \in [a, b]$;
- k and l are strictly Fréchet differentiable at $(\hat{x}(a), \hat{x}(b))$.

A pair (ξ, η) is *critical* for $(\mathscr{O}\mathscr{C})$ at $(\hat{x}, \hat{u})$ if

- $\hat{f}^\circ(\xi(a), \xi(b)) \leq 0$;
- $\hat{g}^\circ_{i[a,b]}(s; \xi(s)) \leq 0$, whenever $\hat{g}_i(s) = 0$, $(i = 1, \ldots, \beta)$;
- $\hat{l}'_i(\xi(a), \xi(b)) \leq 0$, whenever $\hat{l}_i = 0$, $(i = 1, \ldots, r)$;
- $\widehat{G}'_i(t; \xi, \eta) := \widehat{G}_{ix}(t)\xi(t) + \widehat{G}_{iu}(t)\eta(t) \leq 0$ whenever $\widehat{G}_i(t) = 0$, $(i = 1, \ldots, \gamma)$;
- $\hat{k}'(\xi(a), \xi(b)) = 0$;
- $\widehat{H}_x(t)\xi(t) + \widehat{H}_u(t)\eta(t) = 0$ and $\xi'(t) = \widehat{K}_x(t)\xi(t) + \widehat{K}_u(t)\eta(t)$ hold for a.e. $t \in [a, b]$.

An arc (ξ, η) is *regular* for $(\mathscr{O}\mathscr{C})$ at $(\hat{x}, \hat{u})$ if

- $\hat{f}^{\circ\circ}(\xi(a), \xi(b))$ is finite;
- $\hat{g}^{\circ\circ}_{[a,b]}(s; \xi(s))$ and $\hat{\sigma}_{[g]}(s; \xi(s))$ are finite for $s \in [a, b]$;
- The mapping $L(t, x, u) := (G(t, x, u), H(t, x, u), K(t, x, u))$ satisfies the conditions in the previous section;
- The function

$$t \mapsto \frac{(\widehat{G}'_i(t; \xi, \eta))^2}{\widehat{G}_i(t)}$$

is essentially bounded on $\{t \mid \widehat{G}_i(t) < 0 \text{ and } \widehat{G}'_i(t; \xi, \eta) > 0\}$ for $i, \ldots, \gamma$;

- k and l are twice strongly directionally differentiable at $(\hat{x}(a), \hat{x}(b))$ in the direction $(\xi(a), \xi(b))$.

Define the Hamiltonian of the problem $(\mathscr{O}\mathscr{C})$ similarly as that of $(\widehat{\mathscr{P}})$.

Theorem 5. *Let* $(\hat{x}, \hat{u})$ *be a regular local minimum for the problem* $(\mathscr{O}\mathscr{C})$. *Then, for every regular critical arc* (ξ, η), *there exist a vector valued Borel regular measure* $\nu = (\nu_1, \ldots, \nu_\beta) \in (\mathscr{M}([a, b]))^\beta$, *two integrable functions* $\varphi : [a, b] \to \mathbf{R}^\gamma$ *and* $\psi : [a, b] \to \mathbf{R}^\delta$, *an absolutely continuous function* $p : [a, b] \to \mathbf{R}^n$, $\lambda_0 \in \mathbf{R}$, $\lambda \in \mathbf{R}^r$, *and* $\mu \in \mathbf{R}^\kappa$ *not all zero; furthermore, there exist a Borel measurable function* $\overline{\beta} : [a, b] \to \mathbf{R}^{\beta \times n}$ *and a vector* $\overline{\alpha} \in \partial \hat{f}$ *such that*

$$(38) \qquad \overline{\beta}(s) \in \partial_{[a, b]} g(s, \hat{x}(s)) \qquad \text{for } s \in [a, b],$$

$$(39) \qquad \int_a^b g^T(s, \hat{x}(s)) \, d\nu(s) = 0$$

$$(40) \qquad \lambda_0 \geq 0, \quad \lambda \geq 0, \quad \text{and} \quad \lambda^T \hat{l} = 0,$$

(41)

$$\varphi(t) \geq 0, \quad \varphi^T(t)\widehat{G}(t) = 0 \quad \text{and} \quad \varphi^T(t)\widehat{G}'(t; \xi, \eta) = 0 \quad \text{for a.e. } t \in [a, b],$$

furthermore,

$$(42) \qquad \begin{pmatrix} -p(a) \\ p(b) \end{pmatrix} = \lambda_0 \overline{\alpha}^T + \hat{k}'^T \mu + \hat{l}'^T \lambda + \begin{pmatrix} \int_a^b \overline{\beta}^T(s) \, d\nu(s) \\ 0 \end{pmatrix},$$

$$(43) \qquad -\dot{p}^T(t) = \widehat{\mathscr{H}}_x \left(t, p(t) + \int_{]t, b]} \overline{\beta}^T(s) \, d\nu(s), \varphi(t), \psi(t) \right)$$

and

$$(44) \qquad \widehat{\mathscr{H}}_u \left(t, p(t) + \int_{]t, b]} \overline{\beta}^T(s) \, d\nu(s), \varphi(t), \psi(t) \right) = 0$$

for a.e. $t \in [a, b]$,

(45)

$$(\lambda_0 \hat{f}^{\circ\circ} + \lambda^T \hat{l}'' + \mu^T \hat{k}'')(\xi(a), \xi(b)) + \int_a^b (\hat{g}^{\circ\circ}_{[a, b]} - 2\hat{\sigma}_{[g]})(s; \xi(s)) \, d\nu(s)$$

$$+ \int_a^b \widehat{\mathscr{H}}'' \left(t, p(t) + \int_{]t, b]} \overline{\beta}^T(s) \, d\nu(s), \varphi(t), \psi(t); \xi(t), \eta(t) \right) dt \geq 0,$$

where $\mathscr{H}''$ *denotes the second-order strong directional derivative of* $\mathscr{H}$ *with respect to the variable* (x, u).

If $\widehat{K}_u$ *is essentially bounded so are* φ *and* ψ.

Remarks. 1. The equations in (39), (40), (41) imply

$$\operatorname{supp} \nu_i \subset \{s \in [a, b] \mid \hat{g}_i(s) = 0\}, \quad \text{for } i = 1, \ldots, \beta,$$

$$\lambda_i = 0, \quad \text{whenever } \hat{l}_i < 0, \ (i = 1, \ldots, r),$$

for a.e. t, $\varphi_i(t) = 0$, whenever $\widehat{G}_i(t) < 0$, or $\widehat{G}'_i(t; \xi, \eta) \neq 0$ $(i = 1, \ldots, \gamma)$.

2. The result given in Theorem 5 is a generalization of all existing results that deal with first and/or second variation for the optimal control problem with constraints. For instance, Theorem 5.2.1 of [6] is the maximum principle for the problem $(\mathscr{O}\mathscr{C})$, where (ii) and (iii) are absent and g in (i) is unilateral. Thus Theorem 5 generalizes that result to a more general setting and provides the second variation of the problem. In [24] and [21] second variations derived for the problem where respectively constraints of type (i), (ii), (v), and (i), (v) are absent. These results are a special case of Theorem 5 above.

Proof of Theorem 5. The problem $(\mathscr{O}\mathscr{C})$ is a special case of $(\widehat{\mathscr{P}})$ presented in Section 5. Here $\alpha = 1$, $\vartheta_1 = 2$, $T_1 = \{(a, b)\}$; for $i = 1, \dots, \beta$, $\pi_i = 1$ and $S_i = [a, b]$, and for $i = 1, \dots, r$ take $\pi_i = 2$ and $S_i = \{(a, b)\}$; for $i = 1, \dots, \kappa$, $\varrho_i = 2$ and $r_{i1} = a$, $r_{i2} = b$; and $\Gamma = \{y \in \mathbf{R}^\gamma : y \le 0\}$.

Since $(\hat{x}, \hat{u})$ is regular for $(\mathscr{O}\mathscr{C})$, then it is regular for $(\widehat{\mathscr{P}})$ if we can show that the rank property of $(\widehat{\mathscr{P}})$ is satisfied. For, it suffices to show that the bounded measurable functions

$$v_i := (0, \dots, 0, \sqrt{-\widehat{G}_i}, 0, \dots, 0)^T$$

satisfy

(46) $v_i \in \Lambda(\Gamma_{\widehat{G}})$,   for $i = 1, \dots, \gamma$.

One can easily see that

$$\Gamma_{\widehat{G}} = \Gamma^{(1)}_{\widehat{G}_1} \times \cdots \times \Gamma^{(\gamma)}_{\widehat{G}_\gamma}$$

and, as an obvious consequence,

$$\Lambda(\Gamma_{\widehat{G}}) = \Lambda(\Gamma^{(1)}_{\widehat{G}_1}) \times \cdots \times \Lambda(\Gamma^{(\gamma)}_{\widehat{G}_\gamma}),$$

where $\Gamma^{(i)} :=]-\infty, 0]$ for all $i = 1, \dots, \gamma$. Since 0 and $\sqrt{-\widehat{G}_i}$ are in $\Lambda(\Gamma^{(i)}_{\widehat{G}_i})$ by Remark 2 to Lemma 6, hence (46) holds trivially.

One can also see that a regular critical direction (ξ, η) for $(\mathscr{O}\mathscr{C})$ is also a regular critical direction for $(\widehat{\mathscr{P}})$. Therefore, we are able to apply Theorem 4.

Thus there exist ν, φ, ψ, λ_0, $\lambda \in \mathbf{R}^r$, $\mu \in \mathbf{R}^\kappa$, and a function $\overline{p}: [a, b] \to \mathbf{R}^n$ of bounded variation, which is right continuous on $[a, b[$ and $\overline{p}(b) = 0$, $\nu \in \mathscr{M}([a, b])^\beta$, φ and ψ are integrable, such that they are not all zero, and there exist $\overline{\alpha} = (\overline{\alpha}_1, \overline{\alpha}_2) \in \partial \hat{f}$ and a measurable function $\overline{\beta}$, satisfying (38),(39), and (40); furthermore,

(47)
$$\sup_{v \le 0} \varphi^T(t)v \le \varphi^T(t)\widehat{G}(t) \quad \text{and} \quad \varphi^T(t)\widehat{G}'(t, \xi, \eta) = 0 \quad \text{for a.e. } t \in [a, b],$$

(48) $\overline{p}(a) = -\lambda_0\overline{\alpha}_1^T - \overline{\beta}^T(a)\nu(\{a\}) - \hat{l}_{x_1}^T\lambda - \hat{k}_{x_1}^T\mu,$

(49) $\overline{p}(t) = \lambda_0\overline{\alpha}_2^T + \hat{l}_{x_2}^T\lambda + \hat{k}_{x_2}^T\mu + \int_{]t, b]} \overline{\beta}^T(s)\, d\nu(s)$

$$+ \int_t^b \widehat{\mathscr{H}}_x^T(s, \overline{p}(s), \varphi(s), \psi(s))\, ds,$$

for all $t \in [a, b[$,

$$\text{(50)} \qquad \widehat{\mathscr{H}}_u(t, \overline{p}(t), \varphi(t), \psi(t)) = 0$$

for a.e. $t \in [a, b]$, and

$$
(\lambda_0 \hat{f}^{\circ\circ} + \lambda^T \hat{l}'' + \mu^T \hat{k}'')(\xi(a), \xi(b)) + \int_a^b (\hat{g}^{\circ\circ}_{[a,b]} - 2\hat{\sigma}_{[g]})(s; \xi(s)) \, d\nu(s)
$$

$$
\text{(51)} \qquad + \int_a^b \widehat{\mathscr{H}}''(t, \overline{p}(t), \varphi(t), \psi(t); \xi(t), \eta(t)) \, dt
$$

$$
\geq 2 \sup \left\{ \int_a^b \varphi^T(t) v(t) \, dt \, \middle| \, v \in \Gamma^\circ(\widehat{G}, \widehat{G}_x \xi + \widehat{G}_u \eta) \right\}.
$$

Moreover, if $\widehat{K}_u$ is essentially bounded so are φ and ψ.

The relations in (47) yield (41) at once. Now define

$$
p(t) := \overline{p}(t) - \int_{]t,b]} \overline{\beta}^T(s) \, d\nu(s), \quad \text{if } t \in [a, b[
$$

and

$$
\text{(52)} \qquad p(b) = \lim_{\tau \to b-} p(\tau) = \lambda_0 \overline{\alpha}_2^T + \hat{l}_{x_2}^T \lambda + \hat{k}_{x_2}^T \mu.
$$

Then by (49), on $[a, b[$,

$$
p(t) = \lambda_0 \overline{\alpha}_2^T + \hat{l}_{x_2}^T \lambda + \hat{k}_{x_2}^T \mu + \int_t^b \widehat{\mathscr{H}}_x(s, \overline{p}(s), \varphi(s), \psi(s)) \, ds,
$$

and by the definition of $p(b)$ this holds true also for $t = b$. Hence p is absolutely continuous and (43) is satisfied for a.e. $t \in [a, b]$. The equation (42) easily follows from (48) and (52). Similarly, (50) implies (44).

In the last step of the proof we show that (51) is equivalent to (45). To see this, we have only to prove that the right-hand side of (51) is zero.

It follows from the first inequality of (47) that, for a.e. t,

$$
\varphi_i(t) \geq 0 \quad \text{and} \quad \varphi_i(t) = 0, \quad \text{whenever } \widehat{G}_i(t) < 0.
$$

By the criticality of (ξ, η) we also have

$$
\widehat{G}'_i(t, \xi, \eta) \leq 0 \quad \text{whenever } \widehat{G}_i(t) = 0.
$$

Therefore

$$
\text{(53)} \qquad \varphi_i(t) \widehat{G}'_i(t, \xi, \eta) \leq 0
$$

for a.e. $t \in [a, b]$. Adding up these inequalities, we arrive at an inequality where the left-hand side is zero a.e. by the second relation of (47). Therefore we must have equality instead of inequality in (53).

Now we are in the position to apply the last statement of Lemma 7. Then we have

$$
\sup \left\{ \int_a^b \varphi_i(t) v_i(t) \, dt \, \middle| \, v_i \in \Gamma^{(i)\circ}(\widehat{G}_i, \widehat{G}_{ix} \xi + \widehat{G}_{iu} \eta) \right\} = 0
$$

for $i = 1, \ldots, \gamma$, whence we get the conclusion that the right-hand side of (51) is zero. Therefore (51) is equivalent to (45).

The nontriviality of the multipliers follows from that of Theorem 4. □

REFERENCES

1. V. M. Alekseev, S. V. Fomin, and V. M. Tihomirov, *Optimal control*, "Nauka", Moscow, 1979. (Russian)

2. J. P. Aubin, *Applied functional analysis*, Wiley-Interscience, New York, 1978.

3. J. P. Aubin and H. Frankowska, *Set-valued analysis, systems and control: Foundations and applications*, Vol. 2, Birkhäuser-Verlag, Boston, Basel, and Berlin, 1990.

4. A. Ben-Tal, *Second order theory of extremum problems*, Extremal Methods and System Analysis (A. V. Fiacco and K. Kortanek, eds.), Springer-Verlag, Berlin, 1980, pp. 336–356.

5. A. Ben-Tal and J. Zowe, *A unified theory of first and second order conditions for extremum problems in topological vector spaces*, Math. Programming Study **19** (1982), 39–76.

6. F. H. Clarke, *Optimization and nonsmooth analysis*, Canad. Math. Soc. Series of Monographs and Advanced Texts, Wiley, New York, 1983.

7. R. Cominetti, *Metric regularity, tangent sets, and second-order optimality conditions*, Appl. Math. Optim. **21** (1990), 265–287.

8. A. Ya. Dubovitskii and A. A. Milyutin, *Extremum problems with constraints*, Dokl. Akad. Nauk SSSR **149** (1963), 759–762 = Soviet Math. Dokl. **4** (1963), 452–455.

9. _____, *Second variations in extremal problems with constraints*, Dokl. Akad. Nauk SSSR **160** (1965), 18–21.

10. B. Fuchssteiner and W. Lusky, *Convex cones*, North-Holland Math. Studies 56, North-Holland, Amsterdam, New York, and Oxford, 1981.

11. I. V. Girsanov, *Lectures on mathematical theory of extremum problems*, Lecture Notes in Economics and Mathematical Systems 67, Springer-Verlag, Berlin, Heidelberg, and New York, 1972.

12. A. D. Ioffe, *On some recent developments in the theory of second order optimality conditions*, Optimization, (S. Dolecki, ed.), Lectures Notes in Math., vol. 1405, Springer-Verlag, New York and Berlin, 1989, pp. 55–68.

13. _____, *Variational analysis of a composite function: a formula for the lower second order epi-derivative*, J. Math. Anal. Appl. **160** (1991), 379–405.

14. A. D. Ioffe and V. M. Tihomirov, *Theory of extremal problems*, North-Holland, Amsterdam, 1979.

15. H. Kawasaki, *An envelope like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems*, Math. Programming **41** (1988), 73–96.

16. _____, *The upper and second order directional derivatives for a sup-type function*, Math. Programming **41** (1988), 327–339.

17. _____, *Second order necessary optimality conditions for minimizing a sup type function*, Math. Programming **49** (1991), 213–229.

18. _____, *Second order necessary and sufficient optimality conditions for minimizing a sup type function*, Appl. Math. Optim. **26** (1992), 195–220.

19. E. S. Levitin, A. A. Milyutin, and N. P. Osmolovskii, *Higher order conditions for a local minimum in problems with constraints*, Uspehi Mat. Nauk **33** (1978), 85–148.

20. H. Maurer and J. Zowe, *First and second order necessary and sufficient conditions for infinite dimensional programming problems*, Math. Programming **16** (1979), 98–110.

21. N. P. Osmolovskii, *Second order conditions for weak local minimum in an optimal control problem (necessity, sufficiency)*, Soviet Math. Dokl. **16** (1975), 1480–1484.

22. Zs. Páles and V. M. Zeidan, *Nonsmooth optimum problems with constraints*, SIAM J. Control Optim. (to appear).

23. M. Schechter, *Principles of functional analysis*, Academic Press, New York, 1971.

24. G. Stefani and P. Zezza, *Optimal control with mixed state-control constraints*, preprint.

25. V. Zeidan and P. Zezza, *The conjugate point condition for smooth control sets*, J. Math. Anal. Appl. **132** (1988), 572–589.

26. E. Zeidler, *Nonlinear functional analysis and its applications*, Vol. I, Springer-Verlag, Berlin, Heidelberg, and New York, 1984.

INSTITUTE OF MATHEMATICS, L. KOSSUTH UNIVERSITY, H-4010 DEBRECEN, PF. 12, HUNGARY
E-mail address: `pales@borel.math.klte.hu`

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY, EAST LANSING, MICHIGAN 48824
E-mail address: `zeidan@math.msu.edu`