Kybernetika

Jiří V. Outrata; Jiří Jarušek

On Fenchel dual schemes in convex optimal control problems

Kybernetika, Vol. 18 (1982), No. 1, 1--21

Persistent URL: `http://dml.cz/dmlcz/125588`

KYBERNETIKA — VOLUME *18* (1982), NUMBER 1

ON FENCHEL DUAL SCHEMES IN CONVEX OPTIMAL CONTROL PROBLEMS

JIŘÍ V. OUTRATA, JIŘÍ JARUŠEK

The problem of stability of a general convex optimal control problem with control and state-space constraints is investigated with respect to "Fenchel type" perturbations. Two closely related dual problems are constructed and examined in detail. Their mutual relation has enabled to study the behavior of appropriate maximizing sequences in concrete applications, namely in a convex variational problem of Bolza and in an optimal control problem with a parabolic system and state-space constraint.

0. INTRODUCTION

The Fenchel duality theorem and its generalizations (cf. e.g. [8]) play an important role in convex programming. Given the paired spaces $\boldsymbol{U}$ and $\boldsymbol{U}^*$, $\boldsymbol{Y}$ and $\boldsymbol{Y}^*$, $\boldsymbol{V}$ and $\boldsymbol{V}^*$, $\boldsymbol{V} \supset \boldsymbol{Y}$, an operator $\Lambda \in \mathscr{L}[\boldsymbol{U}, \boldsymbol{Y}]$, our task is to solve the extremal problem

$$F(u, \Lambda u) \to \inf$$

(0.1) subj. to

$$u \in \boldsymbol{U},$$

where $F[\boldsymbol{U} \times \boldsymbol{V} \to \overline{\mathbb{R}}]$. Using the perturbations involved into the perturbed objective as follows

$$\Phi(u, p) = F(u, \Lambda u - p), \tag{0.2}$$

we obtain by the standard construction the dual problem

$$-F^*(\Lambda^* p^*, -p^*) \to \sup$$

(0.3) subj. to

$$p^* \in \boldsymbol{V}^* .$$

In concrete applications there is often a certain degree of freedom concerning the choice of appropriate spaces and duality pairings. The main task is evidently to achieve the stable situation, where, by the definition

$$(0.4) \qquad \inf_{u \in U} F(u, \Lambda u) = \sup_{p^* \in V^*} - F^*(\Lambda^* p^*, - p^*),$$

and a solution of (0,3) exists, but, simultaneously, problem (0.3) should be simple and defined over some "reasonable" space V^*, preferably reflexive or Hilbert. Unfortunately, these criteria are often in contradiction. If we take the space of perturbations V sufficiently large, V^* may be "reasonable", but the solution of (0.3) may fail to exist. If we set $V = Y$, V^* may be too large, or (by the change of the pairing) we arrive at a dual, where the cost can hardly be evaluated.

In this sequel we are given a coercive convex programming problem defined over a reflexive space. Two "Fenchel type" dualizations are performed and found to be mutually related. This relationship enables to prove certain assertions concerning convergence properties of minimizing sequences in these duals. It seems that it is highly recommendable to solve convex programming problems of the type defined in Sec. 1 by means of their Fenchel duals provided certain assumptions are satisfied. We remove in such a way all state-space and/or control constraints and are finally faced some unconstrained optimization on a reflexive space, easily solvable by some gradient or subgradient technique. This approach is applied for a convex variational problem of Bolza and for an optimal control problem with a parabolic system and state-space constraints, where the appropriate dual problems are investigated in detail.

The following notation is employed: δ_C is the indicatory functional of a set C, η_C is the support functional of a set C, N_r is the negative cone in R^r, i.e. $N_r = \{v \in R^r \mid v^i \leqq 0,\ i = 1, 2, \ldots, r\}$, $B(\bar{x}, \varrho) = \{x \in X \mid \|x - \bar{x}\| \leqq \varrho\}$ is the closed ball in some normed space X, $f(t, x)$ and $f^*(t, x^*)$ are mutually conjugate functions with respect to the second variable, $\mathscr{L}[X, Y]$ is the space of all continuous linear operators mapping X into Y, A^* is the adjoint of a linear map A, $w(X, X^*)$ is the weak topology induced on X by X^*. $C_\varkappa(x^*)$ is the contact set of a set $\varkappa \subset X$ with respect to a direction x^*, i.e. $C_\varkappa(x^*) = \arg\sup_{x \in \varkappa} \langle x^*, x \rangle$, $|\cdot|_n$ is the Eucleidean norm of R^n and θ denotes the zero vector.

1. THE PROBLEM POSED

The object of our investigations will be the solution of the following convex optimal control problem

$$(\mathscr{P}) \qquad \begin{aligned} & J(u, y) \to \inf \\ \text{subj. to} \quad & y = \Pi u, \\ & u \in \omega \subset U, \\ & y \in K \subset V, \end{aligned}$$

by the way of construction of an appropriate dual problem. In $(\mathscr{P})$ we assume that

(i) U is a reflexive Banach space with U^* being its normed dual.

(ii) Y (response space) is a Banach space, Y^* its dual, $\langle \cdot, \cdot \rangle$ the appropriate duality pairing and Y^* is equipped with a topology compatible with the pairing.

(iii) V is a Banach space, V^* its dual endowed with a compatible topology. $J[U \times V \to \bar{R}]$ is a proper convex lower semicontinuous (l.s.c.) functional and $\omega \times K \subset \text{int}(\text{dom } J)$.

(iv) ω, K are closed and convex.

(v) $\Pi \in \mathscr{L}[U, Y]$.

(vi) $G = \omega \cap \{u \in U \mid \Pi u \in K\}$ is nonempty.

Furthermore, we involve the functional

$$f_{\bar{y}}(u) = J(u, \Pi u - \bar{y}) \tag{1.1}$$

depending on the parameter $\bar{y} \in V$ and suppose additionally that:

(vii) Either ω is bounded or $f_{\bar{y}}(u)$ is on ω coercive uniformly on $B(\theta, \varrho_0)$ for a suitable ϱ_0, i.e.

$$\lim_{\substack{\|u\|_U \to +\infty \\ u \in \omega}} f_{\bar{y}}(u) = +\infty \quad \text{uniformly for} \quad \bar{y} \in B(\theta, \varrho_0) \subset V. \tag{1.2}$$

Assumptions listed above ensure that the infimum of $(\mathscr{P})$ is finite and its solution exists due to Prop. 1.2 (II) of [3]. We construct now the perturbed essential objective

$$F(u, p) = J(u, \Pi u - p) + \delta_\omega(u) + \delta_K(\Pi u - p), \tag{1.3}$$

where $p \in V$ are the perturbations of the system equation, and arrive easily at the dual problem

$$(\mathscr{D}) \qquad \begin{aligned} & - \mathscr{S}^*(\Pi^* p^*, - p^*) \to \sup \\ & \text{subj. to} \\ & p^* \in V^*, \end{aligned}$$

where $\mathscr{S}^*$ is the Fenchel-conjugate to the functional

$$\mathscr{S}(u, y) = J(u, y) + \delta_\omega(u) + \delta_K(y). \tag{1.4}$$

Proposition 1.1. Under the assumptions being imposed $\inf(\mathscr{P}) = \sup(\mathscr{D})$.

For the proof cf. e.g. [7].

In many cases $(\mathscr{D})$ possesses very advantageous properties (no constraints, even implicit ones), but sometimes its solution does not exist ($(\mathscr{P})$ is not stable with respect to perturbations (1.3)). The well-known stability condition for $(\mathscr{P})$ can be formulated in our case in the form of the following assertion:

Proposition 1.2. Let there exists $u_0 \in \omega$ such that

$$\Pi u_0 \in \text{int } K. \tag{1.5}$$

Then $(\mathscr{D})$ possesses at least one solution $\hat{p}^* \in V^*$.

Remark: Assumption (1.5) may be somewhat weakened. E.g. according to [8], [2] it suffices to assume that

$$\theta \in \operatorname{int}\,(\Pi\omega - K)\,. \tag{1.5a}$$

The following simple example, where the system behaviour is described by an ordinary differential equation, demonstrates that the solution of $(\mathscr{D})$ may fail to exist even in the simplest situations.

Let us solve the problem

$$\tfrac{1}{2}\|u\|_U^2 \to \inf$$

subj. to

$$y(t) = \int_0^t u(\tau)\,\mathrm{d}\tau\,, \quad t \in [0, 1]\,, \tag{1.6}$$

$$y(t) \leqq p_1(t) = \max\{-t, -\tfrac{1}{2}\} \quad \text{on} \quad [0, 1]\,,$$

where $U = L_2\,[0, 1]$, $V = L_2\,[0, 1]$. This simple control problem is clearly of the type $(\mathscr{P})$, satisfies all assumptions listed above and its solutions is

$$\hat{u} = \begin{cases} -1 & \text{on} \quad [0, \tfrac{1}{2}) \\ 0 & \text{on} \quad [\tfrac{1}{2}, 1]\,. \end{cases} \tag{1.7}$$

The corresponding trajectory $\hat{y}(t) = p_1(t)$ and the optimal cost value $\mu = \frac{1}{4}$. Problem $(\mathscr{D})$ attains the form

$$-\tfrac{1}{2}\|v^*\|_{U^*}^2 - \eta_K(-p^*) \to \sup$$

(1.8) subj. to

$$v^*(t) = \int_t^1 p^*(\tau)\,\mathrm{d}\tau\,, \quad p^* \in L_2[0, 1]\,,$$

where

$$K = \{y \in L_2[0, 1] \mid y(t) \leqq p_1(t) \text{ a.e. on } [0, 1]\}\,.$$

We show now that every maximizing sequence for (1.8) will be unbounded in the L_2 metric, i.e. (1.8) does not possess any solution. To prove this, we first rewrite (1.8) into the equivalent form

$$-\tfrac{1}{2}\|v^*\|_{U^*}^2 - \langle \dot{v}^*, p_1\rangle \to \sup$$

(1.9) subj. to

$$\dot{v}^* \in D\,,$$
$$v^*(1) = 0\,,$$

where

$$D = \{w \in L_2[0, 1] \mid w(t) \leqq 0 \text{ a.e. on } [0, 1]\}\,.$$

This extremal problem is clearly formulated over the space $\boldsymbol{H}^1[0, 1]$. After performing the integration by parts, it attains the form

$$-\tfrac{1}{2}\int_0^1 (v^*(t))^2 \,\mathrm{d}t - \int_0^{1/2} v^*(t)\,\mathrm{d}t \to \sup$$

subj. to

$$\dot{v}^* \in D\,,$$
$$v^*(1) = 0\,.$$

Let us now remove the constraint $\dot{v}^* \in D$ and examine the relaxed problem

$$(1.10)\qquad -\tfrac{1}{2}\int_0^1 (v^*(t))^2 \,\mathrm{d}t - \int_0^{1/2} v^*(t)\,\mathrm{d}t \to \sup$$

subj. to

$$v^*(1) = 0\,.$$

When we substitute in (1.10) the space $\boldsymbol{H}^1$ by $\boldsymbol{L}_2$ as the space of solutions, the constraint $v^*(1) = 0$ may also be removed and we arrive finally at the maximization of the strictly concave coercive function $-\tfrac{1}{2}\int_0^1 (v^*(t))^2 \,\mathrm{d}t - \int_0^{1/2} v^*(t)\,\mathrm{d}t$ over the whole space $\boldsymbol{L}_2[0, 1]$. This problem has of course only one solution, namely $\hat{v}^* = \hat{u}$ (given by (1.7)), for which the cost in (1.10) attains the value $\mu = \frac{1}{4}$. Thus, as $\hat{v}^* \notin \boldsymbol{H}^1[0, 1]$, we see immediately that there does not exist any element $\bar{v}^* \in \boldsymbol{H}^1[0, 1]$ satisfying the constraint of (1.10) and such that the corresponding cost value would be μ. Hence, we should set $\boldsymbol{V} = \boldsymbol{H}^1[0, 1]$, then $(\mathscr{D})$ would be formulated over $\boldsymbol{H}^{1*}[0, 1]$ and would possess the unique solution $-\delta(\frac{1}{2})$ (the distributive derivative of $\hat{u}$). Alternatively, we could set $\boldsymbol{V} = \boldsymbol{H}^1[0, 1]$ as well, but use the Sobolev pairing

$$\langle a, b\rangle = \langle a(0), b(0)\rangle_{R^n} + \int_0^1 \dot{a}(t)\,\dot{b}(t)\,\mathrm{d}t\,.$$

Then $(\mathscr{D})$ attains the form

$$(1.11)\qquad -\tfrac{1}{2}\int_0^1 (v^*(t))^2 \,\mathrm{d}t + \inf_{\substack{y \in H^1[0,1] \\ y \leq p_1, y(0) = 0}} \int_0^1 \dot{y}(t)\,v^*(t)\,\mathrm{d}t \to \sup$$

subj. to

$$v^* \in \boldsymbol{L}_2[0, 1]\,.$$

A solution of (1.11) exists. However, a more regular v^* is needed for a simple evaluation of its cost. Its value remains the same as in (1.9).

2. ALTERNATIVE FENCHEL DUALIZATION

We involve now a little bit more structure into $(\mathscr{P})$ to be able to perform the "Fenchel type" perturbations also in the control space. So, let us assume that the generalized feasible control region ω is in fact the Cartesian product of a control region $\tilde{\omega} \subset W$ and a subset $\varkappa$ of Z; W, Z are some reflexive Banach spaces, $U = W \times Z$. Furthermore, let Π be a continuous isomorphism of $W \times Z$ and Y. We introduce maps $\Pi_1 \in \mathscr{L}[W, Y]$ and $\Pi_2 \in \mathscr{L}[Z, Y]$ by $\Pi : [w, z] \mapsto \Pi_1 w + \Pi_2 z$ and define a homomorphism $\mathscr{D}$ as a map of Y onto W such that $\mathscr{D}|_{\mathscr{R}(\Pi_1)} = \Pi_1^{-1}$ and $\mathscr{D}|_{\mathscr{R}(\Pi_2)} = \theta$ and a homomorphism P as a map of Y onto Z such that $P|_{\mathscr{R}(\Pi_1)} = \theta$ and $P|_{\mathscr{R}(\Pi_2)} = \Pi_2^{-1}$. We assume finally that $\mathscr{D}$ and P are continuous and recast $(\mathscr{P})$ into the form

$$(2.1) \qquad \tilde{J}(\mathscr{D}y, Py, y) \to \inf$$

$$\text{subj. to} \quad \mathscr{D}y \in \tilde{\omega} \subset W, \quad Py \in \varkappa \subset Z, \quad y \in K \subset Y,$$

where $\tilde{J}(w, z, y) = J(u, y)$ if $u = [w, z]$.

Many optimizations of the type $(\mathscr{P})$ can easily be converted into the form (2.1) and many others are directly given by (2.1) and we construct now the alternative perturbed essential objective as follows:

$$(2.2) \quad \tilde{F}(y, v) = \tilde{J}(\mathscr{D}y + v, Py, y) + \delta_\omega(\mathscr{D}y + v) + \delta_\varkappa(Py) + \delta_K(y), \quad v \in W.$$

The appropriate dual problem attains the form

$$(\tilde{\mathscr{D}}) \qquad -\tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^*v^*) \to \sup$$

$$\text{subj. to } v^* \in W^*,$$

where $\tilde{\mathscr{S}}^*$ is the Fenchel-conjugate to the functional

$$(2.3) \quad \tilde{\mathscr{S}}(w, z, y) = \tilde{J}(w, z, y) + \delta_{\tilde{\omega}}(w) + \delta_\varkappa(z) + \delta_K(y) + \delta_\theta(Py - z).$$

The stability condition for $(\mathscr{P})$ taken as an extremal problem on Y with respect to perturbations (2.2) can be formulated as follows:

Proposition 2.1. Let there exists $y_0 \in K$ such that

$$(2.4) \qquad \mathscr{D}y_0 \in \operatorname{int} \tilde{\omega},\ Py_0 \in \varkappa.$$

Then $(\tilde{\mathscr{D}})$ possesses at least one solution.

Assumption (2.4) may be weakened analogically as (1.5) to the form

$$(2.4a) \qquad \theta \in \operatorname{int}(\tilde{\omega} - \mathscr{D}(K \cap P^{-1}\varkappa)).$$

One may now ask what is the advantage of this approach over the case discussed in the previous section. In what follows we will show that in fundamental applications we are able to construct mere "approximations" of $(\mathscr{D})$ and $(\tilde{\mathscr{D}})$ so that rather these approximations should be compared from the numerical point of view. On the other hand, however, the Kuhn-Tucker vectors of $(\mathscr{P})$ with respect to perturbations (1.3) and (2.2) are closely related and especially this relationship enables us to characterize the convergence properties of minimizing sequences in applications of Sec. 3 and Sec. 4.

Proposition 2.2. Let $V = Y$, $p^* \in Y^*$ and $v^* = \Pi_1^* p^* \in W^*$. Then

$$-\mathscr{S}^*(\Pi^* p^*, -p^*) \leqq -\tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^* v^*)\,.$$

In particular, if $\hat{p}^*$ is a Kuhn-Tucker vector for $(\mathscr{P})$ with respect to perturbations (1.3) then $\hat{v}^* = \Pi_1^* \hat{p}^*$ is a Kuhn-Tucker vector for $(\mathscr{P})$ with respect to perturbations (2.2).

Proof. First we involve the structure introduced in this section into the construction of $(\mathscr{D})$. Clearly, this problem attains then the form (for $V = Y$)

$$(\bar{\mathscr{D}}) \qquad -\bar{\mathscr{S}}^*(\Pi_1^* p^*, \Pi_2^* p^*, -p^*) \to \sup$$

$$\text{subj. to } p^* \in Y^*,$$

where

$$\bar{\mathscr{S}}(w, z, y) = \bar{J}(w, z, y) + \delta_{\tilde{\omega}}(w) + \delta_{\varkappa}(z) + \delta_K(y)$$

and for all $p^* \in Y^*$ $\bar{\mathscr{S}}^*(\Pi_1^* p^*, \Pi_2^* p^*, -p^*) = \mathscr{S}^*(\Pi^* p^*, -p^*)$. It may easily be shown that if we denote $z^* = \Pi_2^* p^*$, then $p^* = \mathscr{D}^* v^* + P^* z^*$. Hence,

$$\bar{\mathscr{S}}^*(\Pi_1^* p^*, \Pi_2^* p^*, -p^*) = \sup_{\substack{w \in W \\ z \in Z \\ y \in Y}} [\langle v^*, w\rangle + \langle z^*, z\rangle + \langle -\mathscr{D}^* v^* - P^* z^*, y\rangle -$$

$$- \bar{\mathscr{S}}(w, z, y) = \sup_{W \times Z \times Y} [\langle v^*, w\rangle + \langle z^*, z - Py\rangle - \langle \mathscr{D}^* v^*, y\rangle -$$

$$- \bar{\mathscr{S}}(w, z, y)] \geqq \sup_{W \times Y} [\langle v^*, w\rangle + \langle z^*, \theta\rangle - \langle \mathscr{D}^* v^*, y\rangle - \bar{\mathscr{S}}(w, Py, y)] =$$

$$= \tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^* v^*)\,.$$

For some $\hat{p}^*$ being a Kuhn-Tucker vector for $(\mathscr{P})$ with respect to perturbations (1.3), the above supremum is achieved at some generalized control $\hat{u} = [\hat{w}, \hat{z}] \in \tilde{\omega} \times \varkappa$ and the corresponding trajectory $\hat{y} = \Pi_1 \hat{w} + \Pi_2 \hat{z} \in K$. Then $\hat{z} - P\hat{y} = \hat{z} - P(\Pi_1 \hat{w} + \Pi_2 \hat{z}) = \theta$ due to the assumptions being imposed and we are done. □

Corollary.

$$-\mathscr{S}^*(\Pi^* p^*, -p^*) \leqq -\tilde{\mathscr{S}}^*(v^*, z^*, -\mathscr{D}^* v^*) =$$

$$= -\tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^* v^*) \quad \text{for each} \quad z^* \in Z^*\,.$$

Hence, Z^* may be chosen in such a way to make $(\tilde{\mathscr{D}})$ as simple as possible.

Thus, the stability (normality) of $(\mathscr{D})$ for $Y = V$ implies the stability (normality) of $(\tilde{\mathscr{D}})$. Let us now imagine the situation, where it is not suitable to perform the maximization of $(\mathscr{D})$ over the whole space Y^* but only over some reflexive Banach space $V_1^* \subset Y^*$. Let us suppose that V_1 is a Banach space such that the assumptions (iii), (iv) of Sec. 1 and (1.2) are satisfied, $Y \subset V_1$, $Y \neq V_1$, V_1^* is its dual. Then we obtain a dual problem $(\mathscr{D}')$ and $\sup(\mathscr{D}') = \min(\mathscr{P})$. Let $\mathscr{S}'$ denote the cost of $(\mathscr{D}')$. Taking this into account, assumptions of the following evident proposition specify the case in which the appropriate objectives are equal.

Proposition 2.3. Let $Y = V_1$ (the closure of Y) and $\bar{K} = K'$, where $K' \subset V_1$ is such that $K' \cap Y = K$. Then, for $p^* \in V_1^*$

$$\mathscr{S}^*(\Pi^* p^*, -p^*) = \mathscr{S}'^*(\Pi^* p^*, -p^*).$$

Propositions 2.2 and 2.3 give together the following important statement.

Proposition 2.4. Let the sequence $\{p_i^*\}$ of elements of V_1^* be a maximizing sequence for $(\mathscr{D}')$ and assumptions of Prop. 2.3 be satisfied. Then the sequence $\{\Pi_1^* p_i^*\}$ of elements of W^* is a maximizing sequence for $(\tilde{\mathscr{D}})$. In particular, if $\tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^* v^*)$ is strictly convex over W^* and satisfies one of the stability conditions (2.4) ((2.4a)), then $\{\Pi_1^* p_i^*\}$ converges to the solution of $(\tilde{\mathscr{D}})$ in the $w(W^*, W)$ topology.

Proof. The first part of the assertion is an easy consequence of Prop. 2.2 and 2.3. The satisfaction of (2.4) or (2.4a) implies that the extremal-value functional of $(\mathscr{P})$ with respect to perturbations (2.2) is continuous at the zero vector. Thus, we may apply Prop. 5.3 (I) of [3] to conclude that this functional is in fact Gâteaux differentiable at θ due to the strict convexity of $\tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^* v^*)$. Hence, its conjugate, i.e. the objective of $(\tilde{\mathscr{D}})$ is $w(W^*, W)$-rotund at $\hat{v}^*$, which is the solution of $(\tilde{\mathscr{D}})$, with respect to the zero functional and this implies finally that every maximizing sequence of $(\tilde{\mathscr{D}})$ converges to $\hat{v}^*$ in the $w(W^*, W)$ topology (cf. [1]). □

3. APPLICATION TO CONVEX CONTROL PROBLEMS OF BOLZA

We take as our model a slightly modified "convex" control problem of [10] which attains the form

$$\int_0^1 [f_1(t, w(t)) + f_2(t, x(t))] \, dt \to \inf \tag{3.1}$$

subj. to

$$\dot{x}(t) = A(t)\,x(t) + w(t) \quad \text{a.e. on} \quad [0, 1],$$
$$x(0) = z \in \varkappa,$$
$$q^i(t, x(t)) \leqq 0 \quad \text{on} \quad [0, 1], \quad i = 1, 2, \ldots, m,$$
$$d^j(t, w(t)) \leqq 0 \quad \text{a.e. on} \; [0,1], \quad j = 1,2, \ldots, r,$$

where $w \in \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$, the elements $a_{ij}(t)$ of matrix $A(t)$ belong to $\boldsymbol{L}_\infty[0, 1]$, the functions $q^i(t, \cdot) : \boldsymbol{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$, $d^j(t, \cdot) : \boldsymbol{R}^n \to \mathbb{R}$, $j = 1, 2, \ldots, r$ and $f_1(t, \cdot)$, $f_2(t, \cdot) : \boldsymbol{R}^n \to \mathbb{R}$ are proper convex and continuous over the whole $\boldsymbol{R}^n$ for every $t \in [0, 1]$ and $\varkappa$ is a convex compact subset of $\boldsymbol{R}^n$ with a nonempty interior. Moreover, let the following assumptions be imposed:

(i) Let $f_1(\cdot, w), f_2(\cdot, x)$ be measurable for each fixed w and x.

(ii) $|f_1(t, w)| \leqq g_1(t) + c_1|w|_n^2$, and $|f_2(t, x)| \leqq g_2(t) + c_2|x|_n^2$ on $[0, 1] \times \boldsymbol{R}^n$, where $c_1, c_2 \geqq 0$ and g_1, g_2 are two real-valued functions from $\boldsymbol{L}_1[0, 1]$.

(iii) $f_1(t, w) \geqq h(t) + \varepsilon|w|_n^2$ on $[0, 1] \times \boldsymbol{R}^n$, where $h(t) \in \boldsymbol{L}_1[0, 1]$, $\varepsilon > 0$.

(iv) Functions $q^i(\cdot, x)$, $i = 1, 2, \ldots, m$ and $d^j(\cdot, x)$, $j = 1, 2, \ldots, r$ are measurable for each fixed $x \in \boldsymbol{R}^n$.

(v) There exists a control $w \in \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$ such that for all $t \in [0, 1]$ $w(t) \in \operatorname{int} W(t)$, $W(t) = \{z \in \boldsymbol{R}^n \mid d^j(t, z) \leqq 0, j = 1, 2, \ldots, r\}$.

(vi) There exists a generalized feasible control $\bar{u} = [\bar{w}, \bar{z}]$ such that the corresponding trajectory $\bar{x}$ satisfies for all $t \in [0, 1]$ the inclusion $\bar{x}(t) \in \operatorname{int} X(t)$, where

$$X(0) = \varkappa \cap \{z \in \boldsymbol{R}^n \mid q^i(0, z) \leqq 0, \ i = 1, 2, \ldots, m\},$$
$$X(t) = \{z \in \boldsymbol{R}^n \mid q^i(t, z) \leqq 0, \ i = 1, 2, \ldots, m\} \quad \text{for} \quad t \in (0, 1].$$

(vii) $A(t) \in L_\infty[0, 1, R^{n^2}]$.

(viii) We are able to express in a simple way the Fenchel conjugates of functionals $f_1(t, z) + \delta_{N_r} \circ d(t, z)$ and $f_2(t, z) + \delta_{N_m} \circ q(t, z)$ with respect to the variable z for almost all $t \in [0, 1]$. Here, $d = [d^1, d^2, \ldots, d^r]^T$ and $q = [q^1, q^2, \ldots, q^m]^T$.

Assumptions (i) and (ii) ensure that the integral part of the objective in (3.1) possesses a normal convex integrand. Moreover, this integral part is in fact a proper convex functional and guarantees also that this functional is in fact continuous over $\boldsymbol{L}_2[0, 1, \boldsymbol{R}^n] \times \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$. Assumption (iii) ensures the coercivity of the whole objective with respect to the control. Assumption (iv) implies together with (v) and (vi) that we are allowed to augment the control constraints and the state-space constraints

$$q^i(t, x(t)) \leqq 0 \quad \text{on} \quad [0, 1], \quad i = 1, 2, \ldots, m,$$

into the integrand of the objective and this new integrand will be a normal convex one. Assumption (vi) is in fact nothing else than the stability requirement (1.5), because the interior of the positive cone in $\boldsymbol{H}^1$ is nonempty. The further reason for imposing assumptions (v) and (vi) is given in the Lemma 3.1 below. Assumption (vii) implies that if we introduce $\mathscr{D}x = \dot{x} - Ax$, then the problem $\mathscr{D}x = w$, $x(0) = z$ possesses a (unique) solution for every initial state vector $z \in \boldsymbol{R}^n$ and every control $w \in \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$. Finally, the satisfaction of assumption (viii) is essential from the technical point of view because otherwise we should solve nontrivial extremal subproblems at each evaluation of the dual cost.

Henceforth an additional state-space constraint and control constraint will be imposed, namely the $(m + 1)$-st inequality

$$q^{m+1}(x(t)) \equiv |x(t)|_n - L \leqq 0 \quad \text{on} \quad [0, 1] , \tag{3.2}$$

and the $(r + 1)$-st inequality

$$d^{r+1}(w(t)) \equiv |w(t)|_n - L \leqq 0 \quad \text{on} \quad [0, 1] , \tag{3.3}$$

where L is some "sufficiently" large number. This is certainly allowed, because an optimal control $\hat{w}$ and an optimal absolutely continuous trajectory $\hat{x}$ really exist so that adding of (3.2), (3.3) does not influence problem (3.1) if L is "sufficiently" large. On the other hand, constraints (3.2), (3.3) will very positively act upon the dual problems, or, better to say, on their regularized approximations.

Lemma 3.1. Assumptions (i), (ii), (iv), (v) and (vi) of this section imply that the conjugates to the functionals

$$I_1(w) = \int_0^1 [f_1(t, w(t)) + \delta_{N_{r+1}} \circ d(t, w(t))] \, dt ,$$

$$I_2(x) = \int_0^1 [f_2(t, x(t)) + \delta_{N_{m+1}} \circ q(t, x(t))] \, dt$$

are the functionals

$$I_1^*(w^*) = \int_0^1 (f_1 + \delta_{N_{r+1}} \circ d)^* \, (w^*(t)) \, dt ,$$

$$I_2^*(x^*) = \int_0^1 (f_2 + \delta_{N_{m+1}} \circ q)^* \, (x^*(t)) \, dt , \tag{3.4}$$

where all functionals are defined over $\boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$.

Proof. We shall prove the assertion for only one couple, let us say, for I_2, I_2^*. With respect to the corresponding result of Rockafellar (cf. [9]) and the discussion above, it remains to be proved that there exists an element $x^* \in \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$ such that the integrand in (3.4) is summable in t. From assumption (vi) we know that for all $t \in [0, 1]$ there exists a vector $\tilde{x}(t) \in \boldsymbol{R}^n$ and a positive scalar $r_0(t)$ such that

$$B_t(\tilde{x}, r_0) = \{v \in \boldsymbol{R}^n \mid |v - \tilde{x}(t)|_n \leqq r_0(t)\} \subset X(t) .$$

Thus, assumption (ii) implies that

$$\delta_{B(\theta,L)}(z) - g_2(t) - c_2 L^2 \leqq f_2(t, z) + \delta_{N_{m+1}} \circ q(t, z) \leqq$$
$$\leqq g_2(t) + c_2 |z|_n^2 + \delta_{B_t(\tilde{x},r_0)}(z) .$$

The left-hand side inequality implies immediately that the integrand in (3.4) is less than $g_2(t) + L|x^*(t)|_n + c_2 L^2$. From the right hand side we obtain that for all $t \in [0, 1]$

(3.5) $$(f_2(t, \cdot) + \delta_{N_{m+1}} \circ q(t, \cdot))^* (z^*) \geqq -g_2(t) + \langle \tilde{x}(t), z^* \rangle -$$

$$-c_2|\tilde{x}(t)|_n^2 + \begin{cases} \dfrac{1}{4c_2} |z^* - 2c_2 \tilde{x}(t)|_n^2 & \text{if } |z^* - 2c_2 \tilde{x}(t)|_n \leqq 2c_2 r_0(t) \\ r_0(t) [|z^* - 2c_2 \tilde{x}(t)|_n - c_2 r_0(t)] & \text{otherwise} . \end{cases}$$

The last relation implies that if we take $x^*(t) = z^*(t) = 2c_2 \tilde{x}(t)$ for $t \in [0, 1]$, then the examined integrand is minorized by $-g_2(t) + c_2|\tilde{x}(t)|_n^2$ so that it is indeed summable. □

We apply now the dualization (2.2) to the problem (3.1). Here, $W = L_2[0, 1, R^n]$, $Z = R^n$, $Y = H^1[0, 1, R^n]$, and

(3.6) $$\Pi_1 w(t) = \int_0^t \Phi(t, \tau) w(\tau) \, d\tau , \quad \Pi_2 z = \Phi(t, 0) z ,$$

where Φ is the fundamental matrix of the dynamical system of (3.1), i.e. the solution of the matrix differential equation

(3.7) $$\frac{d}{dt} \Phi(t, \tau) = A(t) \Phi(t, \tau) , \quad t, \tau \in [0, 1] , \quad \Phi(\tau, \tau) = Id .$$

If we define the operator $\mathscr{D}[H^1[0, 1, R^n] \to L_2[0, 1, R^n]] : x \mapsto \dot{x} - Ax$ and the operator $P[H^1[0, 1, R^n] \to R^n] : x \mapsto x(0)$, all requirements of Sec. 2 are met and we are prepared to compute the dual cost $\tilde{\mathscr{S}}^*$:

$$\tilde{\mathscr{S}}^*(v^*, \theta, -\mathscr{D}^* v^*) = \sup_{d(t, w(t)) \in N_{r+1}} \left[\langle v^*, w \rangle_{L_2} - \int_0^1 f_1(t, w(t)) \, dt \right] +$$

(3.8) $$+ \sup_{\substack{x(0) = z \in \varkappa \\ q(t, x(t)) \in N_{m+1}}} \left[\langle -\mathscr{D}^* v^*, x \rangle_{L_2} - \int_0^1 f_2(t, x(t)) \, dt \right] .$$

The first extremal subproblem on the right hand side of (3.8) is, under the assumptions being imposed, easily solvable and we can write

$$\sup_{d(t, w(t)) \in N_{r+1}} \left[\langle v^*, w \rangle_{L_2} - \int_0^1 f_1(t, w(t)) \, dt \right] = \int_0^1 (f_1 + \delta_{N_{r+1}} \circ d)^* (v^*(t)) \, dt .$$
(3.9)

On the other hand, the second subproblem is much harder. The differential operator $\mathscr{D}^*[L_2[0, 1, R^n] \to H^{1*}[0, 1, R^n]]$ can be formally written in the form

(3.10) $$\langle \mathscr{D}^* v^*, x \rangle_{L_2} = \langle x(1), v^*(1) \rangle - \langle x(0), v^*(0) \rangle -$$

$$- \int_0^1 \langle \dot{v}^*(t) + A^*(t) v^*(t), x(t) \rangle \, dt .$$

As we look for the solution of $(\tilde{\mathscr{D}})$ in $\boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$, we have now to work with elements of $\boldsymbol{H}^{1*}[0, 1, \boldsymbol{R}^n]$. On the other hand, when solving $(\tilde{\mathscr{D}})$ numerically, we can approach its solution $\hat{v}^* \in \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$ by functions of $\mathscr{H}[0, 1, \boldsymbol{R}^n] = \{v^* \in \boldsymbol{H}^1[0, 1, R^n] \mid v^*(1) = \theta\}$. Then, for $v^* \in \mathscr{H}[0, 1, \boldsymbol{R}^n]$ the last inner product on the right hand side of (3.10) is the usual $\boldsymbol{L}_2$ product of functions from $\boldsymbol{L}_2$. At this place we have to impose an other requirement, enabling us to decouple the second extremal subproblem of (3.8). Thus, from now on we shall assume that the constraint functions q^i are such that for any $x^* \in \boldsymbol{L}_2[0, 1, \boldsymbol{R}^n]$

$$(3.11)\quad \sup_{\substack{x \in \boldsymbol{H}^1[0,1,\boldsymbol{R}^n] \\ x(0) \in \varkappa}} \left[\langle x^*, x \rangle_{\boldsymbol{L}_2} - \int_0^1 [f_2(t, x(t)) + \delta_{N_{m+1}} \circ q(t, x(t))] \, \mathrm{d}t \right] =$$

$$= \sup_{x \in \boldsymbol{L}_2[0,1,\boldsymbol{R}^n]} \left[\langle x^*, x \rangle_{\boldsymbol{L}_2} - \int_0^1 [f_2(t, x(t)) + \delta_{N_{m+1}} \circ q(t, x(t))] \, \mathrm{d}t \right].$$

Remark. This situation occurs e.g. provided q^i are continuous over $\boldsymbol{R}^n \times \mathbb{R}$, $i = 1, 2, \ldots, m$, and there exists an $\varepsilon > 0$, and a trajectory $\tilde{x}(t)$ such that

$$\max \{q^i(t, \tilde{x}(t)) \mid t \in [0, 1], \; i = 1, 2, \ldots, m + 1\} \leqq -\varepsilon .$$

For $v^* \in \mathscr{H}[0, 1, \boldsymbol{R}^n]$ and under the previous assumption, the second extremal subproblem of (3.8) may be converted into the following much simpler form:

(3.12)

$$\sup_{x \in \boldsymbol{L}_2[0,1,\boldsymbol{R}^n]} \left[\langle \dot{v}^* + A^* v^*, x \rangle_{\boldsymbol{L}_2} - \int_0^1 [f_2(t, x(t)) + \delta_{N_{m+1}} \circ q(t, x(t))] \, \mathrm{d}t \right] +$$

$$+ \sup_{\substack{x(0) \in \varkappa \\ q(0, x(0)) \in N_{m+1}}} \langle v^*(0), x(0) \rangle =$$

$$= \int_0^1 (f_2 + \delta_{N_{m+1}} \circ q)^* (\dot{v}^*(t) + A^*(t) \, v^*(t)) \, \mathrm{d}t + \eta_{\tilde{\varkappa}}(v^*(0)) ,$$

where

$$\tilde{\varkappa} = \varkappa \cap \{x(0) \in \boldsymbol{R}^n \mid q(0, x(0)) \in N_{m+1}\} .$$

So, the solution of the appropriate problem $(\tilde{\mathscr{D}})$ may be approached arbitrarily closely by the solution of the following optimal control problem (taken as a minimization)

$$\mathscr{S}_1(v^*(0)) + \int_0^1 [\Psi_1(t, v^*(t)) + \Psi_2(t, p^*(t))] \, \mathrm{d}t \to \inf$$

subj. to

$$(\tilde{\mathscr{D}}_{\mathrm{ap}})\qquad \dot{v}^*(t) = -A^*(t) \, v^*(t) - p^*(t) \quad \text{for almost every} \quad t \in [0, 1] ,$$

$$v^*(1) = \theta ,$$

where

$$\mathscr{S}_1 = \eta_{\hat{x}}\,,$$

$$\Psi_1(t, v^*(t)) = (f_1 + \delta_{N_{r+1}} \circ d)^* (v^*(t))\,,$$

and

$$\Psi_2(t, p^*(t)) = (f_2 + \delta_{N_{m+1}} \circ q)^* (-p^*(t))\,,$$

provided we prove that $-\inf(\tilde{\mathscr{D}}_{ap}) = \max(\tilde{\mathscr{D}}) = \min(\mathscr{P})$. To show this, we shall construct the dual problem $(\mathscr{D}')$ using the dualization (1.3) and taking $V = L_2[0, 1, R^n]$. This problem attains the form (taken as the minimization)

$(\mathscr{D}')$

$$\int_0^1 (f_1 + \delta_{N_{r+1}} \circ d)^* (\bar{\Pi}_1^* p^*)\, dt + \eta_{\hat{x}}(\bar{\Pi}_2^* p^*) + \int_0^1 (f_2 + \delta_{N_{m+1}} \circ q)^* (-p^*)\, dt \to \inf$$

subj. to

$$p^* \in L_2[0, 1, R^n]\,,$$

where $\bar{\Pi}_1^* = \Pi_1^*/_{L_2[0,1,R^n]}$ and $\bar{\Pi}_2^* = \Pi_2^*/_{L_2[0,1,R^n]}$. This problem clearly need not possess a solution, but $-\inf(\mathscr{D}') = \sup(\mathscr{D})$ due to Prop. 1.1.

Lemma 3.2.

$$(\tilde{\mathscr{D}}_{ap}) \equiv (\mathscr{D}')\,.$$

Proof. Indeed, as

$$\bar{\Pi}_1^* p^* = \int_t^1 \Phi^*(\tau, t)\, p^*(\tau)\, d\tau\,,$$

we may introduce $v^*(t) = \bar{\Pi}_1^* p^* \in H^1[0, 1, R^n]$ as the solution of the differential equation

$$\dot{v}^*(t) = -A^*(t)\, v^*(t) - p^*(t) \tag{3.13}$$

backwards from the terminal condition

$$v^*(1) = \theta\,. \tag{3.14}$$

As

$$\bar{\Pi}_2^* p^* = \int_0^1 \Phi^*(t, 0)\, p^*(t)\, dt\,,$$

we immediately see that $\bar{\Pi}_2^* p^* = v^*(0)$ which was to be proved. □

Thus, we may solve $(\mathscr{D}')$ or $(\tilde{\mathscr{D}}_{ap})$ which is equivalent over the space of dual controls $p^* \in L_2[0, 1, R^n]$. In such a way, we have an optimization over a Hilbert space so that advanced subgradient techniques (cf. e.g. [5], [6]) may be applied. Concerning the convergence, an easy consequence of Prop. 2.4 can be formulated.

Proposition 3.1. Let $\mathscr{F}(v^*) = \int_0^1 (f_1 + \delta_{N_{r+1}} \circ d)^* \, (v^*(t)) \, \mathrm{d}t$ be a strictly convex functional over $L_2[0, 1, R^n]$ and $\{v_i^*\}$ be the sequence of trajectories corresponding to a minimizing sequence $\{p_i^*\}$ of dual controls $(\mathscr{D}')$ $((\tilde{\mathscr{D}}_{\mathrm{ap}}))$. Then, the sequence $\{v_i^*\}$ converges weakly to the Kuhn-Tucker vector of (3.1) with respect to perturbations (2.2) which is unique.

Proof. It suffices to realize that the satisfaction of the Slater constraint qualification (1.5) which is implied by assumption (vi) of this section implies also the satisfaction of the weakened constrained qualification (2.4a). Hence, the extremal-value functional of (3.1) with respect to perturbations (2.2) is continuous at the zero vector. □

Remark. If $\varkappa$ is a singleton x_0, then clearly $\operatorname{int} X(0) = \emptyset$. However, if we perform a simple translation of the state-space $y = x - x_0$ and set $Y = \{y \in H^1[0, 1, R^n] \mid y(0) = \theta\}$, then it suffices to require that there exists a feasible control $\bar{w}$ such that the corresponding trajectory $\bar{y}$ satisfies

$$\bar{y}(t) \in \operatorname{int} Y(t), \quad Y(t) = \{z \in R^n \mid q^i(t, z + x_0) \leqq 0, \; i = 1, 2, \ldots, m\}$$

for the stability condition (1.5) to be satisfied.

We illustrate this situation on the trivial example of Sec. 1. Let us write first the problem (1.6) in the form (3.1), i.e.

$$\tfrac{1}{2} \int_0^1 w^2(t) \, \mathrm{d}t \to \inf$$

subj. to

(3.15)
$$\begin{aligned} \dot{x}(t) &= w(t) \quad \text{a.e. on} \quad [0, 1], \\ x(0) &= 0, \\ x(t) &\leqq p_1(t) \quad \text{on} \quad [0, 1]. \end{aligned}$$

After adding the additional constraint $|x(t)| \leqq 10$ for $t \in [0, 1]$, the problem $(\mathscr{D}')$ $((\tilde{\mathscr{D}}_{\mathrm{ap}}))$ attains the form

(3.16)
$$\int_0^1 [\tfrac{1}{2}(v^*)^2 \, (t) + \eta_\zeta(-p^*(t))] \, \mathrm{d}t \to \inf$$

subj. to

$$\begin{aligned} \dot{v}^*(t) &= -p^*(t) \quad \text{a.e. on} \quad [0, 1], \\ v^*(1) &= 0, \end{aligned}$$

where

$$\zeta = \{x(t) \in L_2[0, 1] \mid -10 \leqq x(t) \leqq p_1(t) \quad \text{a.e. on} \quad [0,1]\}.$$

It can easily be shown that this problem does not possess an optimal control in $L_2[0, 1]$. However, the sequence of dual controls

$$p_k^*(t) = \begin{cases} 0 \text{ on } [0, \frac{1}{2} - 1/k), \\ -\frac{1}{4}\pi k \cos(\frac{1}{2}\pi k(t - \frac{1}{2})) \text{ on } [\frac{1}{2} - 1/k, \frac{1}{2} + 1/k], \\ 0 \text{ on } (\frac{1}{2} + 1/k, 1] \end{cases}$$

is a minimizing sequence of (3.16) and corresponding trajectories

$$v_k^* \xrightarrow{w} \hat{v}^* = \begin{cases} -1 \text{ on } [0, \frac{1}{2}), \\ 0 \text{ on } [\frac{1}{2}, 1] \end{cases}$$

which is the Kuhn-Tucker vector (Lagrange multiplier) of (3.15) with respect to the dualization (2.2). Here, we have immediately the optimal control $\hat{w}(t)$ for (3.15)

$$\hat{w} \in \arg\max_{w \in L_2[0,1]} \int_0^1 [\hat{v}^*(t)\, w(t) - \tfrac{1}{2}w^2(t)]\, dt = \{\hat{v}^*\}.$$

It is important to remark that the dual cost $\tilde{\mathscr{S}}^*$ given by (3.8) is not necessarily defined over the whole space $W^* = L_2[0, 1, R^n]$. We could overcome this objection by adding some suitable constraint on the norm of x, but this would complicate the evaluation of the dual cost substantially. Thus, let us investigate the effective domain of the cost of $(\tilde{\mathscr{D}}_{ap})$ (or $(\mathscr{D}')$) which we take now as a functional of the dual control p^* and denote Ξ.

Proposition 3.2.

$$\operatorname{dom} \Xi = L_2[0, 1, R^n].$$

Proof. It suffices to prove that

$$\lim_{\|w\|_{L_2} \to \infty} \frac{1}{\|w\|_{L_2}} \int_0^1 [f_1(t, w(t)) + \delta_{N_{r+1}} \circ d(t, w(t))]\, dt = +\infty,$$

and

$$\lim_{\|x\|_{L_2} \to \infty} \frac{1}{\|x\|_{L_2}} \int_0^1 [f_2(t, x(t)) + \delta_{N_{m+1}} \circ q(t, x(t))]\, dt = +\infty.$$

The first relation is certainly true due to assumption (iii) of this section. The second relation is true because of the additional state-space constraint (3.2). Indeed,

$$f_2(t, x(t)) + \delta_{N_{m+1}}\, q(t, x(t)) \geqq -g_2(t) - c_2|x(t)|_n^2 + \delta_{N_1}(|x(t)|_n - L) \geqq$$
$$\geqq -g_2(t) - c_2L^2 + \delta_{B_{L_\infty}(0,L)}(x(t))$$

so that the above limit is $+\infty$ as a consequence of the fact that L_∞ norm is stronger than L_2 norm. □

In the rest of this section we shall discuss how to solve $(\mathscr{D}')$ $((\tilde{\mathscr{D}}_{\text{ap}}))$ in the general (nondifferentiable) case. There are various effective numerical methods for minimization of convex functionals over some Hilbert space. However, we have to be able to compute at least one subgradient of the minimized objective at every point. In what follows we will show how to proceed in the case of $(\mathscr{D}')$ $((\tilde{\mathscr{D}}_{\text{ap}}))$.

Proposition 3.3.

$$\partial\Xi(p^*) = \Pi_1 \operatorname*{arg\,sup}_{w \in L_2[0,1,R^n]} \left\{ \langle v^*, w \rangle_{L_2} - \int_0^1 [f_1(\tau, w(\tau)) + \delta_{N_{r+1}} \circ d(\tau, w(\tau))] \, dt \right\} +$$

$$+ \Pi_2 \, C_{\hat{x}}(v^*(0)) - \operatorname*{arg\,sup}_{x \in L_2[0,1,R^n]} \left\{ \langle -p^*, x \rangle_{L_2} - \int_0^1 [f_2(\tau, x(\tau)) + \right.$$

$$\left. + \delta_{N_{m+1}} \circ q(\tau, x(\tau))] \, dt \right\},$$

where $v^*(t)$ is the solution of the equation (3.13) backwards from the terminal condition (3.14).

Proof. As all functionals creating Ξ are defined and continuous over $L_2[0, 1, R^n]$, Moreau-Rockafellar's theorem may be involved to express $\partial\Xi$ as the sum of subdifferentials of single components. From the same reason we are entitled to apply the standard theorem of the subdifferential calculus concerning the subdifferentials of composed maps. Finally, it is easy to see that the individual subdifferentials consist exactly of those elements maximizing the individual decoupled extremal subproblems at the evaluation of the dual cost (3.8). □

The following proposition may in some cases help us to estimate how precise is our approximate solution of the primal problem (3.1) if we solve it by the way of $(\mathscr{D}')$ $((\tilde{\mathscr{D}}_{\text{ap}}))$. However, to apply it, we must be able to guess somehow the minimum of (3.1).

Proposition 3.4. Let $\{p_i^*\}$ be a minimizing sequence for $(\mathscr{D}')$ $((\tilde{\mathscr{D}}_{\text{ap}}))$, and let $\{\varrho_i\}$ be a sequence of nonnegative scalars such that

$$\Xi(p_i^*) < \inf(\mathscr{D}') + \varrho_i = -\min(\mathscr{P}) + \varrho_i \, .$$

Then the optimal control of (3.1)

$$\hat{w} \in \partial_{\varrho_i} \int_0^1 (f_1 + \delta_{N_{r+1}} \circ d)^*(v_i^*(t)) \, dt \,, \tag{3.17}$$

where v_i^* is the trajectory of the system (3.13) corresponding to the terminal condition (3.14) and dual control p_i^*.

The proof consists merely of an obvious modification of Prop. 1.2 (V) in [3].

Corollary. Let $f_1(t, w(t)) = \frac{1}{2}\langle w(t), w(t)\rangle$ and the controls w be unconstrained. Then

(3.18) $$\|w_i - \hat{w}\|_{L_2} \leqq \sqrt{(2\varrho_i)},$$

where w_i is the approximate solution of (3.1) corresponding to v_i^* as the approximate optimal trajectory of $(\mathscr{D}')$, $((\tilde{\mathscr{D}}_{ap}))$.

Proof. Relation (3.17) implies in this case that

$$0 \leqq \frac{1}{2}\int_0^1 \langle v_i^*(t), v_i^*(t)\rangle \, dt + \frac{1}{2}\int_0^1 \langle \hat{w}(t), \hat{w}(t)\rangle \, dt \leqq \int_0^1 \langle v_i^*(t), \hat{w}(t)\rangle \, dt + \varrho_i .$$

As $w_i = v_i^*$, we obtain immediately the estimate (3.18). □

4. AN EXAMPLE FROM CONSTRAINED CONTROL PROBLEMS WITH A PARABOLIC SYSTEM

Let Ω be a bounded domain with a sufficiently smooth boundary and $Q = (0, 1) \times \Omega$. Let $a_{ij} \in W^1_\infty(Q)$ be such functions that $a_{ij}(t, z) = a_{ji}(t, z)$ a.e. in Q; $i, j = 1, 2, \ldots, n$. Let there exist $a_0 > 0$ such that for every $\xi \in \mathbf{R}^n$ $a_{ij}(t, y)\,\xi_i\xi_j \geqq a_0|\xi|_n^2$ on Q. Let us introduce

(4.1) $$\tilde{Y} = \{y \in L_2[0, 1, \overset{\circ}{H}{}^1(\Omega)] \mid \frac{\partial y}{\partial t} \in \mathbf{L}_2(Q), D^2_{zz}y \in \mathbf{L}_2[Q, \mathbf{R}^{n^2}]\},$$

where $D^2_{zz} = \left(\frac{\partial^2}{\partial z_i\, \partial z_j}\right)^n_{i,j=1}$. $\tilde{Y}$ will be provided by the norm

(4.2) $$\|y\|^2_{\tilde{Y}} = \|y\|^2_{L_2(Q)} + \left\|\frac{\partial y}{\partial t}\right\|^2_{L_2(Q)} + \sum_{i,j=1}^n \left\|\frac{\partial^2 y}{\partial z_i\, \partial z_j}\right\|^2_{L_2(Q)}.$$

For $y \in \tilde{Y}$ we denote

(4.3) $$Ay = \sum_{i,j=1}^n \frac{\partial}{\partial z_i}\, a_{ij}\, \frac{\partial y}{\partial z_j}.$$

Put the control space $\mathbf{U} = \mathbf{L}_2(Q)$ and define the cost J on $\mathbf{U} \times \mathbf{Y}$ by

(4.4) $$J(u, y) = \tilde{J}_1(u) + \tilde{J}_2(y) = \tfrac{1}{2}\|u\|^2_{\mathbf{U}} + [\tfrac{1}{2}\varepsilon\|y\|^2_{L_2(Q)} + \tfrac{1}{2}\eta\|y(1, \cdot) - \tilde{x}\|^2_{L_2(\Omega)}],$$

where ε, η are positive parameters and $\tilde{x} \in \mathbf{L}_2(\Omega)$ is a "desired" terminal state.

Let a be a nonnegative function from $C^0(\bar{Q})$ such that $\tilde{x}(z) \leqq a(1, z)$ a.e. in Ω and $y_0 \in \overset{\circ}{H}{}^1(\Omega) \cap H^2(\Omega)$ be a given function such that $y_0(z) \leqq a(0, z)$ a.e. in Ω. Our task is to solve the following optimal control problem:

$$J(u, y) \to \inf$$

(4.5) subj. to

$$\frac{\partial y}{\partial t} = Ay + u \quad \text{a.e. in} \quad Q, \quad y(0, z) = y_0(z) \quad \text{on} \quad \Omega,$$

$$y = \theta \quad \text{a.e. on} \quad [0, 1] \times \partial\Omega,$$

$$y(t, z) \leqq a(t, z) \quad \text{a.e. on} \quad Q.$$

We notice that due to well-known regularity results (cf. e.g. [4]) the solution of the system equation with the homogeneous Dirichlet boundary value condition and given y_0 belongs to $\tilde{\boldsymbol{Y}}$ whenever $u \in \boldsymbol{U}$. For the sake of simplicity we shall suppose that $y_0 = \theta$.

The posed problem can be treated by dual means in the similar framework as problem (3.1). However, its structure is somewhat different and therefore the results of Sec. 2 cannot be applied dierectly. Let $\mathscr{H}_1, \mathscr{H}_2, \boldsymbol{U}$ be three reflexive Banach spaces, $\boldsymbol{Y} := \mathscr{H}_1 \times \mathscr{H}_2$, let Π_1 be a continuous isomorphismus of $\boldsymbol{U}$ and $\mathscr{H}_1$, let C be a continuous epimorphismus from $\mathscr{H}_1$ onto $\mathscr{H}_2$. Denote $\Pi_2 = C\Pi_1$, $\Pi = [\Pi_1, \Pi_2]$. Let $\mathscr{D} = \Pi_1^{-1}$, $D = \mathscr{D} \circ P_1$, where P_1 is the canonical projection of $\mathscr{H}_1 \times \mathscr{H}_2$ onto $\mathscr{H}_1$. We choose first $\boldsymbol{V} = \boldsymbol{Y}$ and introduce the following perturbed essential objective:

$$(4.6) \qquad F(u, p) = J_1(u) + J_2(\Pi_1 u - p_1) + J_3(\Pi_2 u - p_2), \quad p = [p_1, p_2] \in \boldsymbol{Y}.$$

(In the case of (4.5) $\mathscr{H}_1 = \{y \in \tilde{\boldsymbol{Y}} \mid y(0, \cdot) = \theta \text{ on } \Omega, y/_{[0,1]\times\partial\Omega} = \theta\}$, $\mathscr{H}_2 = \boldsymbol{H}^2(\Omega) \cap \overset{\circ}{\boldsymbol{H}}{}^1(\Omega)$, $J_1 = \tilde{J}_1$, $J_2(y) = \frac{1}{2}\varepsilon\|y\|^2_{\boldsymbol{L}_2(Q)} + \delta_K(y)$, $J_3(w) = \frac{1}{2}\eta\|w - \tilde{x}\|^2_{\boldsymbol{L}_2(\Omega)} + \delta_{K_1}(w)$, $K = \{z \in \boldsymbol{L}_2(Q) \mid z \leqq a \text{ a.e. on } Q\}$, $K_1 = \{z \in \boldsymbol{L}_2(\Omega) \mid z \leqq a(1, \cdot) \text{ a.e. on } \Omega\}$, $\mathscr{D} : y \mapsto \partial y/\partial t - Ay$, $C : v \mapsto v(1, \cdot)$.)

The dual problem $(\mathscr{D})$ attains the form

$$(4.7) \qquad -F^*(\theta, p^*) = -J_1^*(\Pi_1^* p_1^* + \Pi_2^* p_2^*) - J_2^*(-p_1^*) - J_3^*(-p_2^*) \to \sup$$
$$\text{subj. to } p_1^* \in \mathscr{H}_1^*, \ p_2^* \in \mathscr{H}_2^*.$$

The perturbed essential objective corresponding to (2.2) is given by

$$(4.8) \qquad \tilde{F}(y, v) = J_1(\mathscr{D}y_1 + v) + J_2(y_1) + J_3(y_2) + \delta_0(Cy_1 - y_2),$$
$$v \in \boldsymbol{L}_2(Q), \quad y = [y_1, y_2] \in \boldsymbol{Y},$$

so that the dual problem $(\tilde{\mathscr{D}})$ attains the form

$$(4.9) \qquad -\tilde{F}^*(\theta, v^*) = -J_1^*(v^*) - \sup_{y_1 \in \mathscr{H}_1} [\langle -\mathscr{D}^* v^*, y_1 \rangle - J_2(y_1) - J_3(Cy_1)] \to \sup$$
$$\text{subj. to } v^* \in \boldsymbol{U}^*.$$

Problem (4.9) is stable, because there exists a feasible trajectory $y \in K$ and no control constraints are imposed (cf. Prop. 2.1). Since Π^* is surjective, we can easily derive for $v^* = \Pi_1^* p_1^* + \Pi_2^* p_2^*$ that

$$(4.10) \quad \tilde{F}^*(\theta, v^*) = J_1^*(\Pi_1^* p_1^* + \Pi_1^* C^* p_2^*) + \sup_{y_1 \in \mathscr{H}_1} [\langle -p_1^*, y_1 \rangle - J_2(y_1) + \langle -p_2^*, Cy_1 \rangle - J_3(Cy_1)] \leqq F^*(\theta, p^*) .$$

In such a way, inequality (4.10) implies a quite analogous assertion to Prop. 2.2. If we now take $\boldsymbol{V} = \boldsymbol{V}_1 := \boldsymbol{L}_2(Q) \times \boldsymbol{L}_2(\Omega)$ in (4.6), we obtain the dual problem $(\mathscr{D}')$ which is normal due to Prop. 1.1. Since $\overline{\mathscr{H}_1 \times \mathscr{H}_2} = \boldsymbol{V}, \overline{K \cap \mathscr{H}_1} = K, \overline{K_1 \cap \mathscr{H}_2} = K_1$, K, K_1 having the role of K' in Prop. 2.3, we conclude analogously to that assertion that the costs of $(\mathscr{D})$ and $(\mathscr{D}')$ are equal for $p^* \in \boldsymbol{V}_1^*$. Hence the convergence statement may be formulated as follows:

Proposition 4.1. Let the sequence $S = \{[p_{1_i}^*, p_{2_i}^*]\}$ of elements of $\boldsymbol{L}_2(Q) \times \boldsymbol{L}_2(\Omega)$ be a maximizing sequence for $(\mathscr{D}')$. Then, S is a maximizing sequence for $(\mathscr{D})$ and $\{v_i^*\} = \Pi^* S$ of elements of $\boldsymbol{U}$ is a maximizing sequence for $(\tilde{\mathscr{D}})$. In particular, if $\tilde{F}^*(\theta, \cdot)$ is strictly convex over $\boldsymbol{L}_2(Q)$ (as in our case), then the sequence $\{v_i^*\}$ tends weakly to a solution of $(\tilde{\mathscr{D}})$.

The proof is quite analogous to the proof of Prop. 2.4. The strict convexity of $\tilde{F}^*(\theta, \cdot)$ in our case is a consequence of the strict convexity of J_1^*.

For $p_1^* \in \boldsymbol{L}_2(Q)$, $p_2^* \in \boldsymbol{L}_2(\Omega)$ it is easy to calculate that $\Pi_1^* p_1^*$, $\Pi_2^* p_2^*$ are solutions of the problems

$$(4.11) \quad \frac{\partial y}{\partial t} + Ay = -p_1^* \quad \text{on} \quad Q, \; y(1, \cdot) = \theta \text{ on } \Omega, \; y/_{[0,1]\times\partial\Omega} = \theta,$$

$$(4.12) \quad \frac{\partial y}{\partial t} + Ay = \theta \quad \text{on} \quad Q, \; y(1, \cdot) = p_2^* \quad \text{on} \quad \Omega, \; y/_{[0,1]\times\partial\Omega} = \theta ,$$

respectively. If we denote v^* the solution of the problem

$$(4.13) \quad \frac{\partial v^*}{\partial t} + Av^* = -p_1^* \text{ on } Q, \; v^*(1, \cdot) = p_2^* \quad \text{on} \quad \Omega, \; v^*/_{[0,1]\times\partial\Omega} = \theta ,$$

$F^*(\theta, p^*)$ can be evaluated in our case as follows:

$$F^*(\theta, p^*) = \frac{1}{2}\|v^*\|^2_{L_2(Q)} + \frac{1}{2\varepsilon} \int\limits_{\{[\tau,\xi]\in Q \,|\, -p_1^*(\tau,\xi) \leqq \varepsilon a(\tau,\xi)\}} (p_1^*(t, z))^2 \, \mathrm{d}t \, \mathrm{d}z +$$

$$(4.14) \quad + \int\limits_{\{[\tau,\xi]\in Q \,|\, -p_1^*(\tau,\xi) > \varepsilon a(\tau,\xi)\}} \left[-p_1^*(t, z)\, a(t, z) - \frac{\varepsilon}{2} (a(t, z))^2 \right] \mathrm{d}t \, \mathrm{d}z +$$

$$+ \int\limits_{\{\xi\in\Omega \,|\, -p_2^*(\xi) \leqq \eta[a(1,\xi) - \hat{x}(\xi)]\}} \left[-p_2^*(z)\, \hat{x}(z) + \frac{1}{2\eta} (p_2^*(z))^2 \right] \mathrm{d}z +$$

$$+ \int\limits_{\{\xi\in\Omega|-p_2^*(\xi)>\eta[a(1,\xi)-\tilde{x}(\xi)]\}} \left[-p_2^*(z)\,a(1,z) - \frac{\eta}{2}\left(a(1,z)-\tilde{x}(z)\right)^2\right] \mathrm{d}z\,.$$

Clearly, $F^*(\theta, \cdot)$ is convex and finite over $\boldsymbol{L}_2(Q) \times \boldsymbol{L}_2(\Omega)$, hence it is continuous and subdifferentiable there. Due to the continuity of $F^*(\theta, \cdot)$ over $\boldsymbol{V}_1^*$, its maximization may be performed over any space dense in $\boldsymbol{V}_1^*$ so that e.g. the problem $(\mathscr{D}'')$ – to maximize $-F^*(\theta, p^*)$ over $\boldsymbol{V}_2^* = \boldsymbol{L}_2(Q) \times (\boldsymbol{H}^{3/2}(\Omega) \cap \mathring{\boldsymbol{H}}^1(\Omega))$ – may also be used as an approximation of $(\mathscr{D})$. Indeed, Π^* is an isomorphismus of the spaces $\boldsymbol{V}_2^*$ and $\boldsymbol{Y}_1 = \{y \in \tilde{\boldsymbol{Y}} \mid y/_{[0,1]\times\partial\Omega} = \theta\}$. If we now define $(\tilde{\mathscr{D}}_{\mathrm{ap}})$ as the maximization of $-\tilde{F}^*$. (θ, v^*) over $\boldsymbol{Y}_1$, we obtain the dual cost in the form

$$F^*(\theta, v^*) = \frac{1}{2}\|v^*\|^2_{L_2(Q)} + \frac{1}{2\varepsilon} \int\limits_{\{[\tau,\xi]\in Q|\tilde{\mathscr{D}}v^*(\tau,\xi)\leq\varepsilon a(\tau,\xi)\}} (\tilde{\mathscr{D}}\, v^*(t,z))^2 \,\mathrm{d}t\,\mathrm{d}z\, +$$

$$(4.15) \qquad + \int\limits_{\{[\tau,\xi]\in Q|\tilde{\mathscr{D}}v^*(\tau,\xi)>\varepsilon a(\tau,\xi)\}} [\tilde{\mathscr{D}}\, v^*(t,z)\, a(t,z) - \tfrac{1}{2}\varepsilon(a(t,z))^2] \,\mathrm{d}t\,\mathrm{d}z\, +$$

$$+ \int\limits_{\{\xi\in\Omega|-v^*(1,\xi)\leq\eta[a(1,\xi)-\tilde{x}(\xi)]\}} \left[-v^*(1,z)\,\tilde{x}(z) + \frac{1}{2\eta}\,(v^*(1,z))^2\right] \mathrm{d}z\, +$$

$$+ \int\limits_{\{\xi\in\Omega|-v^*(1,\xi)>\eta[a(1,\xi)-\tilde{x}(\xi)]\}} [-v^*(1,z)\,a(1,z) - \tfrac{1}{2}\eta\,(a(1,z)-\tilde{x}(z))^2] \,\mathrm{d}z\,,$$

where

$$\tilde{\mathscr{D}} : v^* \mapsto \frac{\partial v^*}{\partial t} + Av^*\,.$$

Thus, $(\mathscr{D}'')$ and $(\tilde{\mathscr{D}}_{\mathrm{ap}})$ are principially the same problems just expressed in different variables p^* and v^* related by the isomorphismus Π^*. Hence, $\sup(\tilde{\mathscr{D}}_{\mathrm{ap}}) = \min(\mathscr{P})$ as well.

As the most subgradient routines are applicable exclusively for optimizations over Hilbert spaces we evaluate here the subdifferential of the cost merely for $(\mathscr{D}')$ defined over $\boldsymbol{L}_2(Q) \times \boldsymbol{L}_2(\Omega)$. The partial subdifferentials have the form:

$$\partial_{p_1^*} F^*(\theta, p^*) = \Pi_1 v^* + \begin{cases} -p_1^*/\varepsilon & \text{on } \{[\tau,\xi] \mid -p_1^*(\tau,\xi) \leq \varepsilon\, a(\tau,\xi)\} \\ a & \text{elsewhere on } Q\,, \end{cases}$$

$$\partial_{p_2^*} F^*(\theta, p^*) = \Pi_2 v^* + \begin{cases} -p_2^*/\eta + \tilde{x}_0 & \text{on } \{\xi \in \Omega \mid -p_2^*(\xi) \leq \eta[a(1,\xi) - \tilde{x}(\xi)]\} \\ a(1,\cdot) & \text{elsewhere on } \Omega. \end{cases}$$

Clearly, $\partial F^*(\theta, p^*) = [\partial_{p_1^*} F^*(\theta, p^*), \partial_{p_2^*} F^*(\theta, p^*)]$ is a singleton for every $p^* \in V_1^*$ so that $F^*(\theta, \cdot)$ is in fact Gâteaux differentiable over V_1^*. Concerning the estimates of the convergence in the "primal" control space, an analogue of Prop. 3.4 may evidently also be formulated.

Remark. Clearly, for the parabolic system we could construct some general theory like in Sec. 3. In order to avoid any repetition we have confined ourselves to the preceding example.

CONCLUSION

Undoubtedly, the way of solving a convex programming problem by the way of its suitable dual problem may be very effective provided the primal problem happens to be stable or at least normal with respect to the appropriate perturbations. In this sequel, two types of perturbations were applied for a general convex optimal control problem with state – and /or control constraints. Due to an easy evaluation of dual objectives, certain "approximative" dual problems were involved. Convergence properties of maximizing (minimizing) sequences in these approximative problems were studied in detail. To be able to construct such sequences we need some information about the local behaviour of the appropriate costs. This information is provided by subdifferentials which were (in Secs 3 and 4) constructed in two important applications of the presented approach.

(Received June 25, 1981.)

REFERENCES

[1] E. Asplund, R. T. Rockafellar: Gradients of convex functions. Trans. Amer. Math. Society *189* (1969), 443–467.

[2] J. P. Aubin: Gradients généralisés de Clarke. Ann. Sci. Math. Québec *2* (1978), 197–252.

[3] I. Ekeland, R. Temam: Analyse Convexe et Problèmes Variationnels. Dunod, Paris 1974.

[4] O. A. Ladyzhenskaya: Boundary Value Problems of Mathematical Physics (in Russian). Nauka, Moscow 1973.

[5] C. Lemaréchal: Bundle methods in nonsmooth optimization. In: Proc. 1-st Workshop Nonsmooth Optimiz., Pergamon Press, Oxford 1978.

[6] C. Lemaréchal: Nonsmooth Optimization and Descent Methods. Inter. Inst. for Appl. Systems Analysis, RR-78-4, 1978.

[7] J. V. Outrata, O. F. Kříž: An application of conjugate duality for numerical solution of continuous convex optimal control problems. Kybernetika *16* (1980), 6, 477–497.

[8] R. T. Rockafellar: Conjugate Duality and Optimization. SIAM/CBMS monograph series No. 16, SIAM Publications 1974.

[9] R. T. Rockafellar: Integrals which are convex functionals. Pac. J. Math. *24* (1968), 525–539.

[10] R. T. Rockafellar: State constraints in convex control problems of Bolza. SIAM J. Control *10* (1972), 691–715.

Ing. Jiří V. Outrata, CSc., RNDr. Jiří Jarušek, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.