Control
and Cybernetics
VOL. 2 (1973) No. 1/2

A local maximum principle for operator constraints and its application to systems with time lags

by

STANISŁAW KURCYUSZ

Technical University of Warsaw
Institute of Automatics

The paper presents a version of the local maximum principle based on the theory of Dubovitskii-Milyutin. Systems of operator constraints satisfying so called Green formula are introduced. A Green formula and a local maximum principle for systems with delayed argument are proved, extending earlier results.

Introduction

In Section 1 of this paper, a variational theorem based on the theory of Dubovitskii-Milyutin [5] is proven. The theorem is formulated for the general case of nonlinear operator constraints and extends similar results obtained in [2], [6], [9] for the linear case; it is also connected with the saddle point theorem proven in [1], but the assumptions made here are more straightforward. Only equality constraints are considered; because of this limitation more detailed discussion of the assumptions is possible.

These results are restated for the case of the constraints operator being defined implicitly by the abstract state equations which satisfy a Green formula. The results obtained are the continuation of some concepts and theorems of Aubin [2]; a local maximum principle in a distinct form is derived, generalizing local maximum principles known for systems described by ordinary [8] and partial differential equations [9].

Not until recently Banks and Kent [3] using the results of Neustadt [10] proved a very general maximum principle for systems described by functional differential equations; however, it was not possible to establish the nontriviality of adjoint variables. Jacobs and Kao [7] applied the multipliers rule to the systems with time lags and proved the local maximum principle in the normal form ($\lambda_0 \neq 0$) in the absence of constraints for control and under the assumption of complete controllability.

The theorems of Section 1, are applied to the case of systems with lags in Section 2. The results of Jacobs and Kao are extended to cover the case of constrained control and systems which are not completely controllable. Section 2 contains also the comparison of results known for systems with delayed argument.

In Sections 3 and 4 the problem when the attainable subspace is closed is discussed and two examples are presented. Section 5 contains final conclusions which apply to infinite dimensional systems in general and can be of interest in numerical work.

Notation

If X, Y are Banach spaces, then X^* will denote the dual of X and $\mathscr{L}(X, Y)$ the set of all continuous linear operators from X to Y. For $A \in \mathscr{L}(X, Y)$, A^* will denote its adjoint, $ker\, A$ its null space and $im\, A$ — its range. If $\mathcal{O} \subset X$ is an open set and $S: \mathcal{O} \to Y$ is a Frechet differentiable operator, then its Frechet derivative at $x_0 \in \mathcal{O}$ will be denoted by $S_x(x_0)$ or briefly S_{x_0}. If U is another Banach space and $\mathcal{O}_x$, $\mathcal{O}_u$ denote the open subsets of X, U and $F: \mathcal{O}_x \times \mathcal{O}_u \to Y$ is Frechet differentiable, then the derivative of F at $(x_0, u_0) \in \mathcal{O}_x \times \mathcal{O}_u$ with respect to x will be denoted by $F_x(x_0, u_0)$ or briefly by F_{x_0}, when u_0 is fixed.

$\langle \cdot, \cdot \rangle$ will denote the duality between X and X^*; that is, for $x^* \in X^*$, $x \in X$, $\langle x^*, x \rangle$ is the value of x^* at the point x. If $K \subset X$, then K^* is defined by

$$K^* = \{x^* \in X^*: \langle x^*, x \rangle \geqslant 0 \ \forall x \in K\}.$$

If K is a subspace, then K^* is and we will write $K^* = K^\perp$ in this case.

Finally, the following convention is adopted concerning the derivatives of a map $f: R^n \times R^r \times [t_0, t_1] \to R^m$ where $[t_0, t_1]$ is an interval of R:

$$\frac{\partial f(x_1, x_2, t)}{\partial x_i} = D_i f(x_1, x_2, t), \ i = 1, 2;$$

if $x_i(\cdot)$ are R^n and R^r, respectively, valued functions defined on $[t_0, t_1]$, then

$$\frac{\partial f(x_1(t), x_2(t), t)}{\partial x_i} = D_i f(x_1(t), x_2(t), t)$$

will be denoted briefly by $D_i f(t)$.

A vector from R^n and its transpose are not distinguished, while A^T denotes the transpose of matrix A.

1. Basic theorems

In this Section we shall introduce the basic notions which will be used throughout the paper.

The following theorem will be of constant use in the sequel.

THEOREM 0 (Banach). Let U, L be Banach spaces and $A \in \mathscr{L}(U, L)$. Then

(i) $(\overline{im\, A})^\perp = ker\, A^*$ and $im\, A^* \subset (ker\, A)^\perp$.

(ii) $im\, A$ is closed in L if, and only if $im\, A^*$ is closed in U^*. If $im\, A$ is closed, then

$$im\, A^* = (ker\, A)^\perp.$$

Part (i) of the theorem follows immediately from the definition of the adjoint operator. Part (ii) can be found in [4] (Chapts. VI.6.2 and VI.6.4).

We assume that the following will be satisfied from now on: U and L are real Banach spaces, $\mathcal{O}$ and M are the subsets of U, $\mathcal{O}$ is nonempty and open while M is closed, convex and of nonempty interior; $S: \mathcal{O} \to L$ and $J: \mathcal{O} \to R$ are continuously Frechet differentiable mappings.

The space U will be called the space of controls, L — the target space, M — the set of admissible controls.

The following basic problem will be considered:

$$\text{(BP)} \quad \begin{cases} \text{minimize } J(u) \\ \text{on the set } M \cap \{u \in \mathcal{O}: S(u) = 0\}. \end{cases}$$

For $u_0 \in O$, the subspace $im\, S_{u_0}$ of L will be called the attainable subspace of S at u_0, or simply attainable subspace, when u_0 is fixed.

THEOREM 1. Let $\hat{u}$ be a local solution to the problem (BP). Then

(i) if the attainable subspace $im\, S_{\hat{u}}$ at $\hat{u}$ is not a proper subspace dense in L, then there exist a number $\lambda_0 \geqslant 0$ and a functional $l^* \in L^*$, $(\lambda_0, l^*) \neq (0, 0)$, satisfying

$$\langle -\lambda_0 J_{\hat{u}} + S_{\hat{u}}^* l^*, \hat{u} - u \rangle \geqslant 0 \quad \forall u \in M. \tag{1}$$

(ii) if $im\, S_{\hat{u}}$ is a closed subspace of L, the tangent subspace to the set $S^{-1}(0)$ at $\hat{u}$ is equal to $ker\, S_{\hat{u}}$ and $int\,(M - \hat{u}) \cap ker\, S_{\hat{u}}$ is nonempty, then $\lambda_0 \neq 0$ in (1).

Proof.

ad (i). Suppose that $J_{\hat{u}} = 0$; then the pair $(1, 0)$ satisfies (1). If $im\, S_{\hat{u}} \neq L$, then by hypothesis $im\, S_{\hat{u}} \not\supset L$; hence by Hahn-Banach theorem or by (i) of Theorem 0, there exists a $l^* \neq 0$, $l^* \in ker\, S_{\hat{u}}^*$, and (1) holds with $(0, l^*)$. Therefore one can assume that $J_{\hat{u}} \neq 0$, $im\, S_{\hat{u}} = L$.

Define the cones

$$K_1 = \{u \in U: \langle J_{\hat{u}}, u \rangle < 0\}$$

$K_2 = \{u \in U: \exists \varepsilon_0, \delta > 0 \ \ \forall \bar{u} \in U \ \ \forall 0 < \varepsilon \leqslant \varepsilon_0 \ \ (\|\bar{u} - u\| < \delta \Rightarrow \hat{u} + \varepsilon \bar{u} \in M)\}$ $K_3 = ker\, S_{\hat{u}}$.

From theorems 7.5, 9.1 and 6.1 of [5] it follows that there exist functionals $u_1^*, u_2^*, u_3^* \in U$, $u_i^* \in K_i$ $(i = 1, 2, 3)$, $(u_1^*, u_2^*, u_3^*) \neq (0, 0, 0)$, such that

$$u_1^* + u_2^* + u_3^* = 0. \tag{2}$$

Theorems 10.2 and 10.5 of [5] imply that $u_1^* = -\lambda_0 J_{\hat{u}}$ for certain $\lambda_0 \geqslant 0$ and $K_3^* = (ker\, S_{\hat{u}})^\perp$. Since $im\, S_{\hat{u}} = L$, it is closed and by (ii) of Theorem 0 it obtains that $u_3^* = S_{\hat{u}}^* l^*$ for certain $l^* \in L^*$.

From theorems 8.2 and 10.1 of [5] and from (2) one has $\langle -\lambda_0 J_{\hat{u}} + S_{\hat{u}}^* l^*, \hat{u} - u \rangle = \langle u_2^*, u - \hat{u} \rangle \geqslant 0 \ \forall u \in M$ so that (1) holds. λ_0 and l^* cannot vanish simultaneously, since then $u_1^* = u_3^* = 0$ and from (2) $u_2^* = 0$, contradicting the nontriviality of (u_1^*, u_2^*, u_3^*).

ad (ii). Arguing as above, we prove the existence of a nonzero triple of functionals (u_1^*, u_2^*, u_3^*) satisfying (2). Suppose that $\lambda_0 = 0$; hence $u_1^* = -\lambda_0 J_{\hat{u}} = 0$. (1) implies:

$$\langle S_{\hat{u}}^* l^*, u \rangle \leq 0 \quad \forall u \in M - \hat{u}. \tag{3}$$

By hypothesis, there is $\bar{u} \in int\,(M - \hat{u}) \cap ker\, S_{\hat{u}}$; hence

$$\langle S_{\hat{u}}^* l^*, \bar{u} \rangle = \langle l^*, S_{\hat{u}} \bar{u} \rangle = 0, \quad \bar{u} \in int\,(M - \hat{u}) \tag{4}$$

From (3) and (4) we deduce that $u_2^* = S_{\hat{u}}^* l^* = 0$. By (2), also $u_3^* = 0$. This contradiction proves (ii).

Thus the theorem is proved.

The basic problem (BP) and Theorem 1 can be easily generalized to the case when the constraint set is equal to $M \cap \{u \in \mathcal{O} : S(u) \in K\}$, where $K \subset L$ is a convex, closed cone. Such a problem with $M = \mathcal{O} = U$ and S affine was considered in [2] and [9], Theorem 13.1, Chapter 3. The general problem with S being K-convex was studied in detail by Golshteyn [6]; however, the case of equality constraints $(K = \{0\})$ was investigated under the assumption that S be linear and surjective (Theorem 2.1, Chapter 3 of [6]). Note that the very general results of Neustadt [10] do not allow to establish the nontriviality of the multipliers in the case of operator equality constraints, unless $L = R^n$.

The really restrictive assumption here is that $int\, M \neq \emptyset$. In the course of the above proof, it is verified that $K_1 \cap K_2 \cap K_3 = \emptyset$ and then the fundamental lemma of Dubovitskii-Milyutin [5; Lemma 5.11] is applied to show the existence of u_1^*, u_2^*, u_3^*. If $int\, M = \emptyset$, this argument cannot be utilised. In this case, $K_2 = \emptyset$ and the conical approximation $K_{\hat{u}}$ of M at $\hat{u}$ should be defined as

$$K_{\hat{u}} = \{u \in U : \exists \varepsilon > 0 \quad \varepsilon u \in M - \hat{u}\}.$$

It can be proved that the following assumptions would do in this case:

— $K_{\hat{u}}^* + K_3^\perp$ is *-weakly closed (which can be viewed as an analogue of Minkowski-Farkaš Lemma);

— $\overline{K_{\hat{u}} \cap K_3} = \overline{K_{\hat{u}}} \cap K_3$ (which in fact is an analogue of Kuhn-Tucker regularity conditions: see [1] and [11] Chapter 2.1).

These assumptions must be verified directly in each case.

Another way of proving the local maximum principle without assuming $int\, M \neq \emptyset$ is to consider the linearized attainable set $S_{\hat{u}}(M)$ and to prove that its conical approximation at $\hat{u}$, which is equal to $S_{\hat{u}}(K_{\hat{u}})$ cannot be dense in L. Then the theorem on tangent functionals ([4] Chapt. V.9.10) will yield the existence of a nonzero l^*.

Both methods are frequently used but they require rather detailed assumptions on the constraints; in the sequel, we shall stick to the assumption that $int\, M \neq \emptyset$.

The case when the attainable subspace is a proper subspace dense in L, is essentially singular and the assumptions of Theorem 1 here cannot be weakened, as the following Lemma shows.

LEMMA 1. Let $S \in \mathcal{L}(U, L)$ and $im\, S$ be a proper dense subspace of L. If U is a Hilbert space, then there exists a continuously Frechet-differentiable functional $J : U \to R$ such that the following minimization problem:

$$\text{(SP)} \begin{cases} \text{minimize } J(u) \\ \text{on the set } ker\, S = \{u \in U : S_u = 0\} \end{cases}$$

has a unique solution $\hat{u}$, and if $\lambda_0 \geq 0$, $l^* \in L^*$ satisfy

$$-\lambda_0 J_{\hat{u}} + S^* l^* = 0 \tag{5}$$

then $\lambda_0 = 0$ and $l^* = 0$.

Proof. By theorem 0, $im\, S^*$ cannot be closed and hence there is an $v_0 \in (ker\, S)^\perp \setminus \setminus im\, S^*$. Fix an $\hat{u} \in ker\, S$ and set $v = v_0 + \hat{u}$, $J(u) = \frac{1}{2}\langle u - v, u - v \rangle$.

Then $\hat{u}$ is clearly the unique solution of the problem (SP); moreover, $J_{\hat{u}} = -v_0$. Since $v_0 \notin im\, S^*$, (5) cannot be satisfied with $\lambda_0 \neq 0$. But $\lambda_0 = 0$ implies $S^* l^* = 0$ and $l^* = 0$, because of (i) Theorem 0 and the density of $im\, S$.

Theorem 1 and Lemma 1 show that in practical applications the choice of the target space L plays a very important role. The topology in L cannot be too weak, otherwise the above mentioned singularity would occur. On the other hand, the topology should not be too strong because this will usually result in a complicated form of Lagrange multipliers. These problems will be discussed in the next sections in the case of differential equations with delayed argument, but the conclusions are general and apply to partial differential equations as well.

In the remaining part of this section, we shall apply Theorem 1 to the case when the operator S is defined implicitly by the abstract state equations satisfying so called Green formula. In fact, we shall build a model of many dynamical optimization problems and obtain the local maximum principle in a general, yet distinct form, together with adjoint equations and transversality conditions. Aubin [2] was first to investigate this problem in the linear case and applied it to partial differential equations.

Assume that

— X, W_0, W_1, Y, L are real Banach spaces,

— $F : X \times U \to Y$, $G : W_1 \to L$, $Q : X \times U \to R$, $Q_1 : W_1 \to R$ are continuously Frechet-differentiable mappings and $B_i \in \mathcal{L}(X, W_i)$, $i = 0, 1$.

The problem is:

$$\text{(P)} \begin{cases} \text{minimize } Q(x, u) + Q_1(B_1 x) \\ \text{on the set of } (x, u) \in X \times U \text{ such that } u \in M \text{ and} \\ \left.\begin{array}{c} F(x, u) = 0 \\ B_0 x = b_0 \end{array}\right\} \quad (6) \\ G(B_1 x) = 0 \quad (7) \end{cases}$$

where b_0 is a given element of W_0.

Assume further that:

(H.1) — there is a nonempty open subset $\mathcal{O} \subset U$ such that to every $u \in \mathcal{O}$ there corresponds unique $x = \mathcal{F}(u)$ satisfying (6) and that the mapping $\mathcal{F}: \mathcal{O} \to X$ is Frechet-differentiable;

(H.2) — for any $u_0 \in \mathcal{O}$, $x_0 = \mathcal{F}(u_0)$ there exists operators $D^+_{x_0} \in \mathcal{L}(Y^*, X^*)$, $T_{x_0} \in \mathcal{L}(Y^*, W_1^*)$ satisfying:

(i) $\langle F_{x_0} x, \psi \rangle = \langle B_1 x, T_{x_0} \psi \rangle + \langle x, D^+_{x_0} \psi \rangle \; \forall\, x \in \ker B_0 \; \forall\, \psi \in Y^*$,

(ii) for any $\lambda \in R$, $w^* \in W_1^*$ there exist an $\psi \in Y^*$ such that

$$\langle D^+_{x_0} \psi, x \rangle = \langle \lambda Q_{x_0}, x \rangle \qquad \forall x \in \ker B_0$$

$$T_{x_0} \psi = w^*.$$

Here, F_{x_0} and Q_{x_0} denote the Frechet derivatives with respect to x of F, Q respectively, evaluated at the point (x_0, u_0).

Under the assumption (H.1) problem (P) can be converted to problem (BP) by defining J and S in the following manner:

$$J: \mathcal{O} \to R, \qquad S: \mathcal{O} \to L \qquad (\mathcal{O} \text{ as in (H.1)})$$

$$J(u) = Q(\mathcal{F}(u), u) + Q_1(B_1 \circ \mathcal{F}(u)),$$

$$S(u) = G(B_1 \circ \mathcal{F}(u)).$$

Assumption (H.1) guarantees the existence of a solution $x = \mathcal{F}_{u_0} u$ to the equations

$$\begin{aligned} F_{x_0} x + F_{u_0} u &= 0, \\ B_0 x &= 0, \end{aligned} \tag{8}$$

whatever is $u \in U$. Let $w_0 = B_1 x_0$ and consider the "terminal condition"

$$G_{w_0} \circ B_1 x = 0. \tag{9}$$

Equations (8), (9) are the linearization of (6), (7). The attainable subspace of the system (6), (7) at u_0 consists of all $l \in L$ such that there exist an $u \in U$ and $x = \mathcal{F}_{u_0}(u)$ being the solution of (8) satisfying $G_{w_0} \circ B_1 x = l$; note that $S_{u_0} = G_{w_0} \circ B_1 \circ \mathcal{F}_{u_0}$, hence the above assertion simply describes $im\, S_{u_0}$ and justifies the name "attainable".

System (6), (7) will be called regularly linearized at u_0 if the subspace tangent to the set $S^{-1}(0)$, where $S = G \circ B_1 \circ \mathcal{F}$ is equal to $\ker S_{u_0}$. Clearly, any affine system (6), (7) is regularly linearized.

Now we are ready to state the local maximum principle for the problem (P).

THEOREM 2. Let $\hat{u} \in \mathcal{O}$ be a local solution to the problem (P) and set $\hat{x} = \mathcal{F}(\hat{u})$, $\hat{w} = B_1 \hat{x}$. If the corresponding attainable subspace $im\, S_{\hat{u}} = im\,(G_{\hat{w}} \circ B_1 \circ \mathcal{F}_{\hat{u}})$ of the system (6), (7) at $\hat{u}$ is not a proper subspace dense in L, then:

(i) there exist a number $\lambda_0 \geq 0$ and $l^* \in L^*$, $(\lambda_0, l^*) \neq (0, 0)$ such that the solution ψ of the equations:

$$\langle D^+_{\hat{x}} \psi, x \rangle = \langle \lambda_0 Q_{\hat{x}}, x \rangle \;\; \forall x \in \ker B_0, \;\; T_{\hat{x}} \psi = \lambda_0 Q_{1\hat{w}} - G^*_{\hat{w}} l^* \tag{10}$$

satisfies the maximum condition:

$$\langle -\lambda_0 Q_{\hat{u}} + F^*_{\hat{u}} \psi, \hat{u} - u \rangle \geq 0 \qquad \forall u \in M; \tag{11}$$

(ii) if, additionally, $im\, G_{\hat{w}}$ is dense in L, then $(\lambda_0, \psi) \neq (0, 0)$;

(iii) if the system (6), (7) is regularly linearized at $\hat{u}$, the attainable subspace at $\hat{u}$ is closed in L and there is $\bar{u} \in int\,(M - \hat{u})$ satisfying (8) and (9) ($u_0 = \hat{u}$, $x_0 = \hat{x}$, $u_0 = \hat{u}$, $x = \mathcal{F}_{\hat{u}} \bar{u}$) then $\lambda_0 \neq 0$.

Proof.

ad (i). It suffices to compute $J_{\hat{u}}$, $S^*_{\hat{u}}$ and apply Theorem 1.

$$J_{\hat{u}} = Q_{\hat{x}} \circ \mathcal{F}_{\hat{u}} + Q_{\hat{u}} + Q_{1\hat{w}} \circ B_1 \circ \mathcal{F}_{\hat{u}}$$

Let ψ_1 be a solution to

$$\langle D^+_{\hat{x}} \psi_1, x \rangle = \langle Q_{\hat{x}}, x \rangle \qquad \forall x \in \ker B_0, \qquad T_{\hat{x}} \psi_1 = Q_{1\hat{w}}. \tag{12}$$

Then, for any $u \in U$

$$\langle (Q_{\hat{x}} \circ \mathcal{F}_{\hat{u}} + Q_{1\hat{w}} \circ B_1 \circ \mathcal{F}_{\hat{u}}), u \rangle = \langle Q_{\hat{x}}, \mathcal{F}_{\hat{u}} u \rangle + \langle Q_{1\hat{w}}, B_1 \circ \mathcal{F}_{\hat{u}} u \rangle =$$

$$= \langle D^+_{\hat{x}} \psi_1, \mathcal{F}_{\hat{u}} u \rangle + \langle T_{\hat{x}} \psi_1, B_1 \circ \mathcal{F}_{\hat{u}} u \rangle = \langle F_{\hat{x}} \circ \mathcal{F}_{\hat{u}} u, \psi_1 \rangle =$$

$$= \langle -F_{\hat{u}} u, \psi_1 \rangle = \langle -F^*_{\hat{u}} \psi_1, u \rangle..$$

in virtue of (H.2 (i)), (8), (9) and (12). Hence

$$J_{\hat{u}} = Q_{\hat{u}} - F^*_{\hat{u}} \psi_1 .$$

Take any $u \in U$ and $l^* \in L^*$, and let ψ_2 be a solution to

$$\langle D^+_{\hat{x}} \psi_2, x \rangle = 0 \qquad \forall x \in \ker B_0, \qquad T_{\hat{x}} \psi_2 = -G^*_{\hat{w}} l^*. \tag{12a}$$

Just like above, we have:

$$\langle S_{\hat{u}} u, l^* \rangle = \langle G_{\hat{w}} \circ B_1 \circ \mathcal{F}_{\hat{u}} u, l^* \rangle = \langle B_1 \circ \mathcal{F}_{\hat{u}} u, G^*_{\hat{w}} l^* \rangle =$$

$$= -\langle B_1 \circ \mathcal{F}_{\hat{u}} u, T_{\hat{x}} \psi_2 \rangle - \langle \mathcal{F}_{\hat{u}} u, D^+_{\hat{x}} \psi_2 \rangle = -\langle F_{\hat{x}} \circ \mathcal{F}_{\hat{u}} u, \psi_2 \rangle =$$

$$= \langle F_{\hat{u}} u, \psi_2 \rangle = \langle u, F^*_{\hat{u}} \psi_2 \rangle,$$

so that $S^*_{\hat{u}} l^* = F^*_{\hat{u}} \psi_2$ where ψ_2 is a solution to (12a). Now, let $\psi = \lambda_0 \psi_1 + \psi_2$ where ψ_1, ψ_2 satisfy (12), (12a) and (λ_0, l^*) are as in Theorem 1, (i). Then ψ is a solution to (10) and the maximum condition (1) yields (11).

ad (ii). Suppose the contrary, i.e. $(\lambda_0, \psi) = (0, 0)$. Then (10) implies $l^* \in \ker G^*_{\hat{w}}$; by hypothesis $im\, G_{\hat{w}}$ is dense, which is equivalent to $\ker G^*_{\hat{w}} = \{0\}$. Hence $l^* = 0$ and $(\lambda_0, l^*) = (0, 0)$, contrary to (i).

ad (iii). Follows immediately from part (ii) of Theorem 1 and the definition of a regularly linearized system.

Note that Theorem 2 will remain valid if one substitutes assumption (H.2) by the following:

(H.2′) — for any $u_0 \in \mathcal{O}$, $x_0 = \mathcal{F}(u_0)$ there exist operators $D_{x_0}^{+} \in \mathcal{L}(Y^*, X^*)$, $T_{x_0}^{i} \in \mathcal{L}(Y^*, W_i^*)$, $i=0, 1$, satisfying:

(i) $\langle F_{x_0} x, \psi\rangle = \langle B_0 x, T_{x_0}^0 \psi\rangle + \langle B_1 x, T_{x_0}^1 \psi\rangle + \langle x, D_{x_0}^{+} \psi\rangle \ \forall x \in X, \ \forall \psi \in Y^*$;

(ii) for any $\lambda \in R$, $w^* \in W_1^*$ there exist an $\psi \in Y^*$ such that

$$D_{x_0}^{+} \psi = \lambda Q_{x_0}, \quad T_{x_0}^1 \psi = w^*.$$

Clearly, (H.2′) implies (H.2) with $T_{x_0} = T_{x_0}^1$.

In order to illustrate the meaning of the above formalism let us consider briefly the following example:

$$\begin{cases} \text{minimize } \int_{t_0}^{t_1} q(x(t), u(t), t)\, dt \\ \text{on the trajectories of the system} \\ \dot{x}(t) = A(t)\, x(t) + B(t)\, u(t) \ \text{ a.e. in } [t_0, t_1] \\ x(t_0) = x^0, \\ g(x(t_1)) = 0, \end{cases}$$

where $x(t) \in R^n$, $u(t) \in R^r$, A, B are $n \times n$ and $n \times r$ matrices and $q\colon R^n \times R^r \times [t_0, t_1] \to R$, $g\colon R^n \to R^m$.

The adjoint equations in this case are given by

$$\begin{cases} -\dot{\psi}(t) - A^T \psi(t) = -\lambda_0 D_1 q(t), \\ \psi(t_1) = D_g^T g l, \quad l \in R^m \end{cases}$$

($D_1 q(t) = \dfrac{\partial q}{\partial x}(\hat{x}(t), \hat{u}(t), t)$ where $\hat{u}$, $\hat{x}$ are the optimal control and solution). These equations are related to the state equations by means of the integration by parts formule:

$$\int_{t_0}^{t_1} \psi(t)\,(\dot{x}(t) - A(t)\, x(t))\, dt = \psi(t_1)\, x(t_1) - \psi(t_0)\, x(t_0) + \\ + \int_{t_0}^{t_1} x(t)\,(-\dot{\psi}(t) - A^T(t)\, \psi(t))\, dt.$$

This is a special case of (H.2′ (ii)) where $(F_{x_0} x)(t) = \dot{x}(t) - A(t)\, x(t)$, $(D_{x_0}^{+} \psi)(t) = -\dot{\psi}(t) - A^T(t)\, \psi(t)$, $B_0 x = x(t_0)$, $B_1 x = x(t_1)$, $T_{x_0}^0 \psi = -\psi(t_0)$, $T_{x_0}^1 \psi = \psi(t_1)$ and the spaces X, Y are defined in a suitable way. It is easy to see how the other operators should be defined, and that the maximum principle (11) is here equivalent to

$$-\lambda_0 D_2 q(t) + \psi(t)\, B(t) = 0 \quad \text{a.e. in } [t_0, t_1]$$

($\lambda_0 \geq 0$), that is,

$$\left.\frac{\partial H(t)}{\partial u}\right|_{u=\hat{u}(t)} \equiv 0.$$

Green formulas can be written for systems described by discrete and partial differential equations [9]. It is also possible to write such a formula for systems with time lags.

2. Systems with lags

In the following, we shall deal with the problem

$$\text{(DP)}\ \begin{cases} \text{minimize } q_1(x(t_1)) + \int_{t_0}^{t_1} q(x(t), u(t), t)\, dt \\ \text{with constraints:} \\ \quad \dot{x}(t) = f(x(t), x(\alpha(t)), u(t), t) & \text{a.e. on } [t_0, t_1] \\ \quad x(t_0) = x^0 \\ \quad \dot{x}(t) = \varphi(t) & \text{a.e. on } [\alpha(t_0), t_0] \\ \quad g_1(x(t_1)) = 0 \\ \quad g(\dot{x}(t), t) = 0 & \text{a.e. on } [\alpha(t_1), t_1] \\ \text{and with the additional restriction} \\ \quad u \in M \subset U([t_0, t_1]) \\ \text{where } M \text{ is a closed, convex set of nonempty interior in the space } U([t_0, t_1]) \\ \text{of control functions defined on } [t_0, t_1]. \end{cases}$$

We assume that:

(A.1) — for any $t \in [\alpha(t_0), t_1]$, $x(t)$ is a R^n vector, and $u(t) \in R^r$, $t \in [t_0, t_1]$;

(A.2) — the functions q_1, q, f, g_1, g are defined on the following spaces:

$$q_1\colon R^n \to R, \quad q\colon R^n \times R^r \times [t_0, t_1] \to R$$

$$f\colon R^n \times R^n \times R^r \times [t_0, t_1] \to R^n$$

$$g_1\colon R^n \to R^m, \quad g\colon R^n \times [\alpha(t_1), t_1] \to R^p;$$

Function $f(x_1, x_2, u, t)$ is assumed to be affine in u, while g is assumed to be affine in the first argument. (See (A.4) and (A.5) below).

(A.3) — the map $\alpha\colon [t_0, t_1] \to R$, representing the argument deviation, is increasing in $[t_0, t_1]$ and $\alpha(t) \leq t - d$ for certain $d > 0$ and all $t \in [t_0, t_1]$, and $\alpha(t_1) > t_0$; moreover, there is an absolutely continuous map $\gamma\colon [\alpha(t_0), \alpha(t_1)] \to [t_0, t_1]$ such that $\alpha(\gamma(t)) = t$ and $\alpha(\gamma(t)) = t$ a.e. in $[\alpha(t_0), \alpha(t_1)]$ and $[t_0, t_1]$, respectively.

This problem will be solved below following Jacobs and Kao [7]. Before applying Theorem 2 it is necessary to define the spaces and operators in a suitable way. Set

$$X = W_1^2([\alpha(t_0), t_1]; R^n).$$

$W_1^2([\alpha(t_0), t_1]; R^n)$ is a Sobolev space of absolutely continuous functions, having first derivative square integrable, endowed with the scalar product

$$\langle \varphi_1, \varphi_2 \rangle = \varphi_1(t_1)\, \varphi_2(t_1) + \int_{\alpha(t_0)}^{t_1} \dot\varphi_1(t)\, \dot\varphi_2(t)\, dt.$$

This Hilbert space is isometrically isomorphic to $L^2(\alpha(t_0), t_1; R^n) \times R^n$, any element $\varphi \in W_1^2([\alpha(t_0), t_1]; R^n)$ being in a one-to-one correspondence with the pair $(\dot\varphi, \varphi(t_1))$. Therefore in the sequel we shall assume that

$$X = L^2(\alpha(t_0), t_0; R^n) \times L^2(t_0, t_1; R^n) \times R^n$$

identifying any element $x \in X$ with the triple $(x', \bar x, x_1)$ satisfying:

$$x' = \dot x|_{[\alpha(t_0), t_0]}, \quad \bar x = \dot x|_{(t_0, t_1)}, \quad x_1 = x(t_1).$$

Often, we shall write simply $x = (x', \dot x, x(t_1))$. Also the elements w of the space of "terminal conditions", $W_1 = W_1^2([\alpha(t_1), t_1]; R^n)$ will be identified with the pairs $w = (\bar w, w_1) = (\dot w, w(t_1))$.

The elements of the space $Y = L^2(t_0, t_1; R^n) \times L^2(\alpha(t_1), t_1; R^n) \times R^n$ will be, however, treated as triples $(\bar\psi, \psi', \psi_1)$ only, and no "global" meaning will be assigned to them.

Finally, set $U = L^2(t_0, t_1; R^r)$, $W_0 = W_1^2([\alpha(t_0), t_0]; R^n)$ and $L = L^2(\alpha(t_1), t_1; R^p) \times R^m$.

The operators will be defined as follows:

(i) $F\colon X \times U \to Y$, $F(x, u) = (\bar F(x, u), 0, 0)$,

where
$$\bar F(x, u)(t) = \dot x(t) - f(x(t), x(\alpha(t)), u(t), t), \quad t \in [t_0, t_1].$$

(ii) $B_0\colon X \to W_0$, $\quad B_0 x = (x', x(t_0))$.

(iii) $B_1\colon X \to W_1$, $\quad B_1 x = (\dot x|_{[\alpha(t_1), t_1]}, x(t_1))$.

(iv) $G\colon W_1 \to L$, $G(w) = G((\bar w, w_1)) = (g(\bar w(\cdot), \cdot), g_1(w_1))$.

Clearly, $B_i \in \mathscr{L}(X, W_i)$, $i = 0, 1$.

Define the functionals: $Q\colon X \times U \to R$, $Q_1\colon W_1 \to R$

$$Q(x, u) = \int_{t_0}^{t_1} q(x(t), u(t), t)\, dt, \quad Q_1(w) = Q_1((\bar w, w_1)) = q_1(w_1).$$

We need that F, G, Q, Q_1 be continuously Frechet differentiable.

(A.4) — f is of the form

$$f(x_1, x_2, u, t) = f_1(x_1, x_2, t) + f_2(x_1, x_2, t)\, u,$$

$x_1, x_2 \in R^n$, $u \in R^r$, $t \in [t_0, t_1]$, where functions $f_i(x_1, x_2, \cdot)$ are measurable $\forall x_1, x_2$, functions $f_i(\cdot, \cdot, t)$ are of class C^1 for almost every t and the following is satisfied

$$|f_1(x_1, x_2, t)| + |D_1 f_1(x_1, x_2, t)| + |D_2 f_1(x_1, x_2, t)| \leqslant M_1(h, t)$$

$$|f_2(x_1, x_2, t)| + |D_1 f_2(x_1, x_2, t)| + |D_2 f_2(x_1, x_2, t)| \leqslant M_2(h)$$

$\forall h > 0$, $\forall x_1, x_2 \in R^n$, $|x_1|$, $|x_2| \leqslant h$, $\forall t \in [t_0, t_1]$ where

$$M_1(h, \cdot) \in L^2(t_0, t_1), \quad M_2(h) < +\infty \ \forall h > 0.$$

(A.5) — g is of the form $g(x, t) = a(t) + b(t)\, x$, where

$$|a| \in L^2(\alpha(t_1), t_1), \quad |b| \in L^\infty(\alpha(t_1), t_1).$$

g_1 is of the class C^1.

(A.6) — Function $q(x, u, \cdot)$ is measurable $\forall x, u$, function $q(\cdot, \cdot, t)$ is of the class C^1 for almost every $t \in [t_0, t_1]$ and the following holds

$$|q(x, u, t)| + |D_1 q(x, u, t)| \leqslant M_3(h, t) + M_2(h)\, |u|^2$$

$$|D_2 q(x, u, t)| \leqslant M_3(h, t) + M_2(h)\, |u|$$

$\forall h > 0$, $\forall\, |x| \leqslant h$, $\forall\, u$ and almost every t where $M_3(h, \cdot) \in L^1(t_0, t_1)$ $\forall h \geqslant 0$. Function q_1 is C^1.

With these assumptions it can be shown that F, G, Q, Q_1 are continuously Frechet differentiable. For the details see [15] and [14]. The very restrictive assumption that f be affine in u and g in x cannot be omitted, otherwise F and G would not be Frechet differentiable at any point, see [16] [1]). In the sequel, for brevity, we shall not use the functions f_1, f_1, a, b, but refer to f, g as a whole. Thus, for instance, $D_3 f = f_2$.

Thus problem (DP) appears to be a special case of (P) with the operators F, G, etc, defined as above. Now we proceed to checking the hypotheses (H.1) and (H.2).

Note that the equations (6) are equivalent to:

$$\left.\begin{aligned} &\bar F(x, u)(t) = \dot x(t) - f(x(t), x(\alpha(t)), u(t), t) = 0 && \text{a.e. in } [t_0, t_1] \\ &x(t_0) = x^0 && . \\ &\dot x(t) = \varphi(t) && \text{a.e. in } [\alpha(t_0), t_0) \end{aligned}\right\} \tag{13}$$

Assume that there is $u_0 \in L^2(t_0, t_1; R^n)$ such that the solution $x_0(\cdot)$ of (13) that exists on $[t_0, t_1]$. (13) is equivalent to the following operator equation

$$\mathscr{A}(x, u) = 0,$$

where $\mathscr{A}(x, u) = (\bar F(x, u), B_0 x - (\varphi, x_0))$. We have that $\mathscr{A}$ is Frechet continuously differentiable, $\mathscr{A}(x_0, u_0) = 0$ and the Frechet derivative $\mathscr{A}_{x_0} (= \mathscr{A}_x(x_0, u_0))$ is an invertible operator (since it is defined by linearized equations (13)). Hence by the implicit operator theorem there are neighbourhoods V_{x_0}, V_{u_0} of x_0, u_0 in X, U such that (13) defines the Frechet-differentiable map $\mathscr{F}\colon V_{u_0} \to V_{x_0}$,

$$F(\mathscr{F}(u), u) = 0,$$

$$\mathscr{F}(u)(t_0) = x^0,$$

$$\mathscr{F}(u)(t) = \varphi(t) \ \text{ a.e. in } [\alpha(t_1), t_1).$$

Since any solution of (13) is unique, we can consider the map $\mathscr{F}$ as being defined from V_{u_0} into X. Thus we proved that the set of all u_0 which define by (13) a solution defined on the whole interval, is open and the map $\mathscr{F}$ exists in a neighbourhood of

---

[1]) Thus, the results obtained in [7] are correct only under the assumptions given above and not those in [7], which allow nonlinearity of f in u.

any such u_0; hence (H.1) is satisfied, provided there is at least one such u_0. Before considering (H.2) assume that

(A.7) — $\dot\gamma$ is essentially bounded on $[\alpha(t_0), \alpha(t_1)]$.

Let $u_0 \in U$ and x_0 be a solution to (13). Define the operators $D^+_{x_0}: Y \to X$ and $T_{x_0}: Y \to W_1$ in the following manner:

$$D^+_{x_0}\psi = D^+_{x_0}(\bar\psi, \psi', \psi_1) = (0, \bar D^+_{x_0}\psi, 0)$$

where

$$\bar D^+_{x_0}\psi(t) = \begin{cases} \bar\psi(t) - \int_t^{t_1} (D_1 f(\tau))^T \bar\psi(\tau)\,d\tau - \int_t^{\alpha(t_1)} (D_2 f(\gamma(\tau)))^T \gamma(\tau)\,\bar\psi(\gamma(\tau)) \times \\ \qquad \times\, d\tau - \psi_1, \quad t_0 \leqslant t \leqslant \alpha(t_1) \\ \bar\psi(t) - \int_t^{t_1}(D_1 f(\tau))^T\bar\psi(\tau)\,d\tau - \psi'(t) - \psi_1, \quad \alpha(t_1) < t \leqslant t_1\end{cases} \tag{14}$$

$$T_{x_0}\psi = T_{x_0}(\bar\psi, \psi', \psi_1) = (\psi', \psi_1).$$

Now let $x \in ker\, B_0$, i.e. $x' = 0$, $x(t_0) = 0$, and $\psi = (\bar\psi, \psi', \psi_1) \in Y$ $(= Y^*)$. Then

$$\begin{aligned} \langle F_{x_0} x, \psi\rangle_Y &= \int_{t_0}^{t_1} \bar F_{x_0} x(t)\,\bar\psi(t)\,dt = \int_{t_0}^{t_1} \big(\dot x(t) - D_1 f(t)\,x(t) - \\ &\quad - D_2 f(t)\,x(\alpha(t))\big)\,\bar\psi(t)\,dt = \int_{t_0}^{t_1} \dot x(t)\,\bar\psi(t)\,dt - \int_{t_0}^{t_1} x(t)\,(D_1 f(t))^T\,\bar\psi(t)\,dt - \\ &\quad - \int_{t_0}^{\alpha(t_1)} x(t)\,(D_2 f(\gamma(t)))^T\,\dot\gamma(t)\,\bar\psi(\gamma(t))\,dt = \int_{t_0}^{t_1} \dot x(t)\,\bar\psi(t)\,dt - \int_{t_0}^{t_1} \dot x(t) \times \\ &\quad \times \int_t^{t_1} (D_1 f(\tau))^T\,\bar\psi(\tau)\,d\tau\,dt - \int_{t_0}^{\alpha(t_1)} \dot x(t) \int_t^{\alpha(t_1)} (D_2 f(\gamma(\tau)))^T\,\dot\gamma(\tau)\,\bar\psi(\gamma(\tau))\,d\tau\,dt + \\ &\quad - \int_{\alpha(t_1)}^{t_1} \dot x(t)\,\psi'(t)\,dt + \int_{\alpha(t_1)}^{t_1} \dot x(t)\,\psi'(t)\,dt - \int_{t_0}^{t_1} \dot x(t)\,\psi_1\,dt + x(t_1)\,\psi_1 = \\ &= \int_{t_0}^{t_1} \dot x(t)\,\bar D^+_{x_0}\psi(t)\,dt + \langle B_1 x, T_{x_0}\psi\rangle_Y = \langle x, D^+_{x_0}\psi\rangle_Y + \langle B_1 x, T_{x_0}\psi\rangle_Y . \end{aligned}$$

Thus the Green formula (H.2. (i)) holds.
In order to prove (ii) take $x \in ker\, B_0$ and note that:

$$\langle Q_{x_0}, x\rangle = \int_{t_0}^{t_1} \dot x(t) \int_t^{t_1} D_1\, q(x_0(\tau), u_0(\tau), \tau)\,d\tau\,dt .$$

Hence (H.2 (ii)) will be proved, if for any $w \in W_1$ there is a $\psi \in Y$ satisfying:

$$\bar D^+_{x_0}\psi(t) = \lambda \int_t^{t_1} D_1\, q(x_0(\tau), u_0(\tau), \tau)\,d\tau \quad \text{a.e. in } [t_0, t_1]$$

$$\psi' = \bar w$$

$$\psi_1 = w_1$$

This is equivalent to the following pair of equations (compare (14)):

$$\bar\psi(t) - \int_t^{t_1} (D_1 f(\tau))^T\,\bar\psi(\tau)\,d\tau = \bar w(t) + w_1 + \lambda \int_t^{t_1} D_1\, q(\tau)\,d\tau \tag{15}$$

$$\text{a.e. in } (\alpha(t_1), t_1)$$

$$-\dot{\bar\psi}(t) = (D_1 f(t))^T\,\bar\psi(t) + (D_2 f(\gamma(t)))^T\,\dot\gamma(t)\,\bar\psi(\gamma(t)) + \lambda D_1\, q(t) \tag{16}$$

$$\text{a.e. in } (t_0, t_1)$$

and the terminal conditions for (16) are determined by (15) and (14):

$$\bar\psi(\alpha(t_1)) = \int_{\alpha(t_1)}^{t_1} \big((D_1 f(t))^T\,\bar\psi(t) + \lambda D_1\, q(t)\big)\,dt + w_1 . \tag{17}$$

Since (15) is a Volterra equation of second kind, it has a solution $\bar\psi$ for any $(\bar w, w_1) \in W_1$. Similarly, (16) can be solved by the method of steps yielding the absolutely continuous solution.

Before stating the result of this section, one must find necessary Frechet derivatives. If $w_0 = (\bar w_0, w_{10})$ is fixed in W_1, and $w \in W_1$, then

$$\langle Q_{1w_0}, w\rangle = Dq_{1w_0} \cdot w_{10};$$

hence

$$Q_{1w_0} = (0, Dq_{1w_0}) \in W_1 .$$

Similarly,

$$G_{w_0} = (Dg(\cdot), Dg_{1w_0})$$

where $Dg(t) = Dg(\bar w(t), t)$ as above.

The attainable subspace at $\hat u_0$ of the operator S, defined implicitly by the constraints in (DP), consists of all points $l = (\bar l(\cdot), l_1) \in L$ such that there is $u \in M$ and $x \in X$ satisfying the linearized equations:

$$\begin{aligned} &\dot x(t) - D_1 f(t)\,x(t) - D_2 f(t)\,x(\alpha(t)) - D_3 f(t)\,u(t) = 0 \quad \text{a.e. in } [t_0, t_1] \\ &x(t_0) = 0 \\ &\dot x(t) = 0 \qquad \text{a.e. in } [\alpha(t_0), t_0) \end{aligned} \tag{18}$$

$$\left.\begin{aligned} &D g_1\, x(t_1) = l_1 \\ &D g(t)\,\dot x(t) = \bar l(t) \qquad \text{a.e. in } [\alpha(t_1), t_1] \end{aligned}\right\} \tag{19}$$

The application of Theorem 2 yields the following.

Theorem 3. Suppose that (A.1)—(A.7) are valid. If $\hat u \in M$ is a local solution to problem (DP) and the attainable subspace at $\hat u$ is not a proper subspace dense in $L^2(\alpha(t_1), t_1; R^p) \times R^m$, then:

(i) There exist a number $\lambda_0 \geqslant 0$, a vector $l_1 \in R^m$ and a function $l \in L^2(\alpha(t_1), t_1; R^p)$, not all equal to zero, and functions $\psi \in L^2(t_0, t_1; R^n)$, $\psi' \in L^2(\alpha(t_1), t_1; R^n)$, and a vector $\psi_1 \in R^m$ such that:

$$\psi_1 = \lambda_0 Dq_1 - (Dg_1)^T l_1, \tag{20}$$

$$\psi'(t) = -(Dg(t))^T l(t) \quad \text{a.e. in } [\alpha(t_1), t_1], \tag{21}$$

$$\psi(t) - \int_t^{t_1} (D_1 f(\tau))^T \psi(\tau)\, d\tau = \psi'(t) + \psi_1 + \lambda_0 \int_t^{t_1} D_1 q(\tau)\, d\tau, \quad \text{a.e. in } [\alpha(t_1), t_1], \tag{22}$$

$$\psi(\alpha(t_1)) = \psi_1 + \int_{\alpha(t_1)}^{t_1} \big((D_1 f(t))^T \psi(t) + \lambda_0 D_1 q(t)\big)\, dt, \tag{23}$$

$$-\dot{\psi}(t) = (D_1 f(t))^T \psi(t) + \big(D_2 f(\gamma(t))\big)^T \dot{\gamma}(t)\, \psi(\gamma(t)) + \lambda_0 D_1 q(t), \quad \text{a.e. in } [t_0, \alpha(t_1)], \tag{24}$$

and the following maximum condition holds:

$$\int_{t_0}^{t_1} (-\lambda_0 D_2 q(t) + \psi(t) D_3 f(t))\, (\hat{u}(t) - u(t))\, dt \geqslant 0 \quad \forall u \in M. \tag{25}$$

Note: all the derivatives here are evaluated along the trajectory $\hat{x}(\cdot)$, corresponding to $\hat{u}(\cdot)$, so that for example

$$D_1 f(t) = D_1 f\big(\hat{x}(t), \hat{x}(\alpha(t)), \hat{u}(t), t\big).$$

(ii) If, in addition, matrix Dg_1 has rank m and matrix $Dg(t)$ has rank p for almost every $t \in [\alpha(t_1), t_1]$, then $(\lambda_0, \psi) \neq (0, 0)$.

(iii) If the system (6), (7) of section 1 with F, B_0, B_1, G defined as above is regularly linearized at $\hat{u}$ (if the state equations and the terminal constraints are affine, this assumption is always satisfied), the attainable subspace is closed and there exists an $\bar{u} \in int\,(M - \hat{u})$ such that the corresponding solution $\bar{x}$ of (18) satisfies

$$Dg_1 \cdot \bar{x}(t_1) = 0$$

$$Dg(t) \cdot \dot{\bar{x}}(t) = 0 \quad \text{a.e. in } [\alpha(t_1), t_1],$$

then $\lambda_0 \neq 0$.

Points (i) and (iii) are immediate corollaries to Theorem 2, (i), (iii). To prove (ii) observe that $(\lambda_0, \psi) = (0, 0)$ implies $\psi_1 = 0$, $\psi' = 0$ in virtue of (22) and (23). Hence if $(\lambda_0, \psi) = (0, 0)$, then $(\lambda_0, (\psi, \psi', \psi_1)) = (0, 0)$, contrary to Theorem 2, point (ii) (since by hypothesis G_w^* is injective, $im\, G_{\hat{w}}$ must by dense in L — see (i), Theorem 0).

The problem when the attainable subspace is closed, will be discussed in Section 3.

Observe first that the pair (ψ', ψ_1) can be identified with a function $\mu \in W_1^2([\alpha(t_1), t_1]; R^n)$. Then $p = \psi|_{[\alpha(t_1), t_1]} - \dot{\mu} = \psi|_{[\alpha(t_1), t_1]} - \psi' \in W_1^2([\alpha(t_1), t_1]; R^n)$ and equation (22) takes the form

$$-\dot{p}(t) = D_1 f(t)\, \psi(t) - \lambda_0 D_1 q(t) \quad \text{a.e. in } [\alpha(t_1), t_1] \tag{26}$$

with terminal condition

$$p(t_1) = \psi_1 = \lambda_0 Dq_1 - (Dg_1)^T l_1.$$

Equations (26), (24) are identical with those obtained by Jacobs and Kao [7], while (22), (24) are the same as those in [3]. The difference between our result and that of [7] as far, as adjoint equations are concerned, is due to the fact that the additional Lagrange multiplier μ can be represented by $\dot{\mu} = \psi'$ and its value either at t_1 or at $\alpha(t_1)$; this results in minor changes in terminal condition for p. Note also that $\dot{\mu} = \psi'$ being an element of $L^2(\alpha(t_1), t_1; R^n)$ is an equivalence class of functions equal almost everywhere and therefore it has no value $\psi'(t)$ at any point $t \in [\alpha(t_1), t_1]$. It can happen, however, that ψ' is equivalent to the function right-continuous at $\alpha(t_1)$. In this case, also $\lim_{t \to \alpha(t_1)+0} \psi(t) = \psi(\alpha(t_1)+0)$ exists and in virtue of (22), (23) we have the jump condition:

$$\psi(\alpha(t_1)+0) - \psi(\alpha(t_1)) = \psi'(\alpha(t_1)) = \dot{\mu}(\alpha(t_1)). \tag{27}$$

The equation (26) can be easily transformed to contain p and ψ' only. Since ψ' is given by (21), this would be an ordinary differential equation for p. Solving numerically this equation is easier than the corresponding integral equation (22).

If $M = U$, $L = W_1$, $g_1(w_1) = w_1$ for $w_1 \in R^n$, $g(\bar{w}(\cdot), \cdot) = \bar{w}(\cdot)$ for $\bar{w} \in L^2(\alpha(t_1), t_1; R^n)$ and the linearized system (18) is completely controllable, then from (ii) and (iii) it follows that $\lambda_0 \neq 0$. Thus the result of Jacobs and Kao appears to be a special case of Theorem 3.

Observe finally that unlike other necessary conditions, Theorem 3 can be applied to the problems of control to targets in both finite-dimensional, and function space. If one is interested in controlling $x(t_1)$ only, it suffices to put $g(y, t) \equiv 0$, $t \in [\alpha(t_1), t_1]$. Then from (21) we have $\psi' = 0$, hence $\psi|_{[\alpha(t_1), t_1]} = p$ and the equations (26), (27) imply that ψ is absolutely continuous in $[t_0, t_1]$ and satisfies the well known adjoint equations [12]

$$\psi(t_1) = \lambda_0 Dq_1 - (Dg_1)^T l_1$$

$$-\dot{\psi}(t) = (D_1 f(t))^T \psi(t) - \lambda_0 D_1 q(t) \quad \text{a.e. in } [\alpha(t_1), t_1]$$

$$-\dot{\psi}(t) = (D_1 f(t))^T \psi(t) + \big(D_2 f(\gamma(t))\big)^T \dot{\gamma}(t)\, \psi(\gamma(t)) - \lambda_0 D_1 q(t), \quad \text{a.e. in } [t_0, \alpha(t_1)).$$

As mentioned in Section 2, the requirement that $\text{int}\, M \neq \emptyset$ is rather restrictive. The typical example of such a set M is given by

$$M = \{u \in L^2(t_0, t_1; R^r) : \int_{t_0}^{t_1} k(t)\, |u(t)|^2\, dt \leqslant K\}$$

where $K \geqslant 0$ and $k(t) \geqslant 0$, $\frac{1}{k(t)} \leqslant K_1 < +\infty$ whenever $k(t) \neq 0$.

Theorem 3 does not cover the classical case of the set M being defined by

$$M = \{u : u \text{ measurable}, \ u(t) \in \Omega \quad \text{a.e. in } [t_0, t_1]\} \tag{28}$$

where Ω is a compact subset of R^r, since this set has no interior in the topology of L^2. In the framework of Theorems 1 and 2 only one thing can be done — to strengthen the topology of $U([t_0, t_1])$. Preferably, one should use L^∞ instead of L^2, then if $\operatorname{int}\Omega \neq \emptyset$, the set M defined by (28) would be of nonempty interior. But in that case also other spaces should be changed, X to $L^\infty(\alpha(t_0), t_0; R^n) \times W_1^\infty([t_0, t_1]; R^n)$, W_1 to $W_1^\infty([\alpha(t_1), t_1]; R^n)$ etc. in order to assure that the attainable subspace would not be a proper dense subspace of L. However, the proof of assumption (H.2) leans on many properties of adjoint spaces, X^*, Y^* and W_1^*. These spaces are isomorphic to spaces of finitely additive bounded set functions, vanishing on sets of Lebesque measure zero ([4] Chapt. IV.8.16). The proof of Green formula (H.2 (ii)) would require many facts known for measures, analogues of theorem of Radon-Nikodym, Fubini, or similar. We do not know whether these theorems are valid for finitely additive set functions.

Much more is known about $U([t_0, t_1]) = C([t_0, t_1]; R^r)$ and its dual. A reasoning similar to the proof of Theorem 3, but more complicated leads to the following.

THEOREM 4. Let $\Omega \subset R^r$ be closed convex and of nonempty interior, $M = \{u \in C([t_0, t_1]; R^r) : u(t) \in \Omega \ \forall t\}$ and $\hat{u} \in M$ be a local solution to the problem (DP) with $U([t_0, t_1]) = C([t_0, t_1]; R^r)$ and $f, q_1, q, g_1, g, \gamma$ satisfying suitable continuity and differentiability assumptions [2]). Suppose that the attainable subspace of the linearized system (18), (19) is not a proper subspace dense in $L = C([\alpha(t_1), t_1]; R^p) \times R^m$. Then there exist a number $\lambda_0 \geq 0$, a vector $l_1 \in R^m$ and a R^p — valued function $\bar{l}$, defined, left continuous and of bounded variation in $[\alpha(t_1), t_1]$, not all equal to zero; there exist a vector $\eta_1 \in R^n$ and functions $\psi \in L^\infty(t_0, \alpha(t_1); R^n)$, $\rho \in L^\infty(\alpha(t_1), t_1; R^n)$ satisfying the following equations:

$$\eta_1 = \lambda_0 Dq_1 - Dg_1 l_1$$

$$\rho(t) - \int_t^{t_1} (D_1 f(\tau))^T \rho(\tau)\, d\tau = \eta_1 + \int_t^{t_1} (D_1 f(\tau))^T d\eta'(\tau) + \lambda_0 \int_t^{t_1} D_1 q(\tau)\, d\tau$$

$$\text{a.e. in } (\alpha(t_1), t_1] \quad (29)$$

$$\psi(t) - \int_t^{t_1} (D_1 f(\tau))^T \psi(\tau)\, d\tau - \int_t^{\alpha(t_1)} \big(D_2 f(\gamma(\tau))\big)^T d\bar{\eta}(\gamma(\tau)) = \eta_1 + \lambda_0 \int_t^{t_1} D_1 q(\tau)\, d\tau$$

$$\text{a.e. in } [t_0, \alpha(t_1)]$$

where

$$\eta'(t) = \int_t^{t_1} (D_1 g(\tau))^T d\bar{l}(\tau), \qquad t \in [\alpha(t_1), t_1]$$

[2]) The assumptions needed here are much weaker than (A.4) and (A.5); f and g may be nonlinear [14].

$$\bar{\eta}(t) = \begin{cases} 0 & t = t_1 \\ -\int_t^{t_1} \rho(\tau)\, d\tau + \eta'(t) & t \in (\alpha(t_1), t_1) \\ -\int_{\alpha(t_1)}^{t_1} \rho(\tau)\, d\tau & t = \alpha(t_1) \\ -\int_{\alpha(t_1)}^{t_1} \rho(\tau)\, d\tau - \int_t^{\alpha(t_1)} \psi(\tau)\, d\tau & t \in [t_0, \alpha(t_1)) \end{cases}$$

and such that the following maximum principle holds:

$$\int_{t_0}^{t_1} (\hat{u}(t) - u(t))\, d\left\{\lambda_0 \int_t^{t_1} D_2 q(\tau)\, d\tau - \int_t^{t_1} (D_3 f(\tau))^T d\bar{\eta}(\tau)\right\} =$$

$$= -\lambda_0 \int_{t_0}^{t_1} D_2 q(t)(\hat{u}(t) - u(t))\, dt + \int_{t_0}^{t_1} (\hat{u}(t) - u(t))(D_3 f(t))^T d\bar{\eta}(t) \geq 0$$

$$\forall u \in M.$$

Points (ii) and (iii) of Theorem 3 can be also formulated in this case.

Let us compare briefly the earlier results [3], [7] and Theorems 3 and 4. Banks and Kent [3] worked in the target space $C([\alpha(t_1), t_1]; R^n)$ in which the attainable subspace consisting of absolutely continuous functions cannot be closed (unless it is finite-dimensional). In the case of complete controllability the attainable subspace is a proper dense subspace of this target space. Lemma 1 explains why it was not possible to establish the nontriviality of (λ_0, ψ); however, ψ' was left continuous and of bounded variation. Jacobs and Kao [7] used smaller target space and the assumption of complete controllability guaranteed the closedness of the attainable subspace. But diminishing the target space, one enlarges the space of Lagrange multipliers; hence both Theorem 4.1 of [7] and Theorem 3 yield the existence of nonzero multipliers, but ψ' is only square integrable. Taking smaller control space $C([t_0, t_1]; R^r)$, as in Theorem 4, is connected with enlarging the space multipliers once again, in view of Lemma 1. The result, is that in Theorem 4 neither ρ nor $\psi|_{[t_0\,\alpha(t_1)]}$ are absolutely continuous; the space of terminal conditions is here $C^1([\alpha(t_1), t_1]; R^n)$ and its dual is isomorphic to some space of rather irregular functions.

The spaces $C([\alpha(t_1), t_1]; R^n)$, $W_1^2([\alpha(t_1), t_1]; R^n)$ and $C^1([\alpha(t_1), t_1]; R^n)$ are by no means the only target spaces allowing the solution of problem (DP) in particular cases. In the next section, an example will be presented showing that while $W_1^2([\alpha(t_1), t_1]; R^n)$ cannot be used (since the only existing multipliers are zero), some other choice of the target space will result in normal Lagrange multipliers.

It seems that new results could be obtained under stronger assumptions concerning the performance index Q and its relation to the subspace $\operatorname{im} S_u^*$, where S is the operator as in Theorem 1.

3. The attainable subspace

Consider the linear system

$$\left.\begin{aligned} &\dot{x}(t)=A(t)\,x(t)+A_1(t)\,x(\alpha(t))+C(t)\,u(t) && \text{a.e. in } [t_0, t_1]\\ &x(t_0)=0 && \\ &x(t)=0 && \text{a.e. in } [\alpha(t_0), t_0)\end{aligned}\right\} \quad (30)$$

$$\left.\begin{aligned} &H_1\,x(t_1)=l_1 && \\ &H(t)\,\dot{x}(t)=\bar{l}(t) && \text{a.e. in } [\alpha(t_1), t_1]\end{aligned}\right\} \quad (31)$$

where A, A_1, C, H_1, H are matrices of suitable dimensions and α is as in Section 2, $u(\cdot)\in U([t_0, t_1])$, $(\bar{l}(\cdot), l_1)\in L(\alpha(t_1), t_1)$. The problem is, when the attainable subspace of the system (36), (37) is closed in $L=L([\alpha(t_1), t_1])$, that is, when the operator $S\colon U\to L$ defined by

$$Su=(H(\cdot)\,x(\cdot),\ H_1\,x(t_1))$$

where x is the solution of (36), has closed range $im\ S$. Note that $L=\bar{L}\times R^m$; it suffices to investigate the range of the operator $\bar{S}\colon U\to\bar{L}$,

$$(\bar{S}u)(t)=H(t)\,\dot{x}(t) \qquad t\in[\alpha(t_1), t_1],$$

as the following lemma shows:

LEMMA. Let $\bar{S}\colon U\to\bar{L}$, $S_1\colon U\to R^m$ be continuous operators, and set $S=(\bar{S}, S_1)\colon U\to\bar{L}\times R^m$. Then $im\ S$ is closed in $\bar{L}\times R^m$ if, $im\ \bar{S}$ is closed in $\bar{L}$.

Proof. By the theorem 0, $im\ S$ is closed if $im\ S^*$ is. We have

$$im\ S^*=im\ \bar{S}^*+im\ S_1^*.$$

Since $im\ S_1^*$ is a finite-dimensional subspace, then $im\ S^*$ is closed if $im\ \bar{S}^*$ is ([5] Lemma 2.6), the latter condition being equivalent to the closednes of $im\ \bar{S}$.

We shall find the operator $\bar{S}$. To this purpose denote by $Y(s, t)$ the $n\times m$ matrix function defined for $t_0\leqslant s\leqslant t$, $\alpha(t_0)\leqslant t\leqslant t_1$, satisfying the following conditions:

$$Y(t, t)=\mathrm{I}\ \ \text{(identity)}$$

$$Y(s, t)=0, \qquad s>t$$

$$\frac{\partial}{\partial s}Y(s, t)=-Y(s, t)\,A(s), \qquad \alpha(t)\leqslant s<t$$

$$\frac{\partial}{\partial s}Y(s, t)=-Y(s, t)\,A(s)-Y(\gamma(s), t)\,\dot{\gamma}(s)\,A_1(\gamma(s));\ \ t_0\leqslant s<\alpha(t).$$

Then

$$x(t)=\int_{t_0}^{t}Y(s, t)\,C(s)\,u(s)\,ds, \qquad t\in[t_0, t_1];$$

hence

$$\dot{x}(t)=C(t)\,u(t)+\int_{t_0}^{t}\frac{\partial}{\partial t}Y(s, t)\,C(s)\,u(s)\,ds.$$

($Y(s, t)$ is absolutely continuous with respect to any of the variables in the set $\{s, t\colon s\leqslant t\}$).

Therefore for $t\in[\alpha(t_1), t_1]$

$$\dot{x}(t)=\int_{t_0}^{\alpha(t_1)}\frac{\partial}{\partial t}Y(s, t)\,C(s)\,u(s)\,ds+C(t)\,u(t)+\int_{\alpha(t_1)}^{t}\frac{\partial}{\partial t}Y(s, t)\,C(s)\,u(s)\,ds.$$

Observe that the elements of $U([t_0, t_1])$ can be treated as pairs (u_1, u_2), $u_1\in U_1=U([t_0, \alpha(t_1)])$, $u_2\in U_2=U([\alpha(t_1), t_1])$; in the case of continuous controls $u_1(\alpha(t_1))=u_2(\alpha(t_1))$. Then

$$im\ \bar{S}=H\,(im\ E+im\ V\circ C)$$

where:

$$\left.\begin{aligned} &(Hw)(t)=H(t)\,w(t),\\ &(Eu_1)(t)=\int_{t_0}^{\alpha(t_1)}\frac{\partial}{\partial t}Y(s, t)\,C(s)\,u(s)\,ds,\\ &(Vl)(t)=l(t)+\int_{\alpha(t_1)}^{t}\frac{\partial}{\partial t}Y(s, t)\,l(s)\,ds,\\ &(Cu_2)(t)=C(t)\,u_2(t).\end{aligned}\right|\ t\in[\alpha(t_1), t_1]$$

Assume first that $W_1=L$, $H=\mathrm{I}$ (identity); then

$$im\ \bar{S}=im\ E+im\ V\circ C.$$

We face considerable difficulties when trying to establish whether $im\ \bar{S}$ is closed. Even if $im\ E$ and $im\ V\circ C$ are closed, their algebraic sum may be not (see [13] Chapter 4, § 4). But $im\ E$ is, in general, not closed [14] since E is a Fredholm operator of the first kind. The investigation of $im\ V\circ C$ is much simpler, because V, being a Volterra operator of the second kind, is a topological isomorphism of $\bar{L}$ onto itself in the case $\bar{L}=L^2(\alpha(t_1), t;\ R^n)$ and $\bar{L}=C([(\alpha(t_1), t_1];\ R^n)$ [14]. Thus $im\ V\circ C$ is closed if, and only if, $im\ C$ is. Thus, while we are not able to give a general answer to the question when $im\ S$ is closed, it is possible to give the following, obvious sufficient condition:

— If $im\ E\subset im\ V\circ C$ and $im\ C$ is closed, then $im\ \bar{S}$ is closed. The inclusion $im\ E\subset im\ V\circ C$ takes place if, for instance, $C(t)=0$, $t\in[t_0, \alpha(t_1))$, or if the operator C is surjective. In the latter case $im\ C$ is obviously closed. The necessary and sufficient condition for C to be surjective in the case $U=L^2(t_0, t_1;\ R^n)$, $W_1=W_1^2([\alpha(t_1), t_1];R^n)$ is [7] that the matrix $C(t)$ has rank n a.e. in $[\alpha(t_1), t_1]$ and the function $t\to|(C(t)\,C^T(t))^{-1}|^2$ is integrable on $[\alpha(t_1),\ t_1]$.

It is much easier to give a sufficient condition for the operator C to have *im* C closed.

To fix ideas, assume that $U=L^2(t_0, t_1; R^r)$, $\bar{L}=L^2(\alpha(t_1), t_1; R^n)$, $C(\cdot)$ is an essentially bounded $n\times r$ matrix function on $[\alpha(t_1), t_1]$. Let $\Gamma'\subset[\alpha(t_1), t_1]$ be the subset of measure zero on which $C(\cdot)$ is not defined, and

$$\Gamma_0=\Gamma\cup\{t\in[\alpha(t_1), t_1]: C(t)=0\}.$$

Define on $[\alpha(t_1), t_1]$ the $r\times r$ matrix function $J(\cdot)$ in the following way

$$J(t)=\begin{cases} 0 & t\in\Gamma_0 \\ \text{matrix of orthogonal projection in } R^r \text{ onto } (\ker C(t))^{\perp} & t\in[\alpha(t_1), t_1]\setminus\Gamma_0 . \end{cases}$$

It can be shown that the function $J(\cdot)$ is measurable; it is bounded, since $J(t)$ is a projection matrix for any $t\in[\alpha(t_1), t_1]$.

The range of the operator $C\colon U_2\to\bar{L}$ is closed if, and only if ([4] Chapt. VI.6.1 and VI.9.15) the following condition is satisfied: (C) — there is a constant $k>0$ such that to each $u\in U_2$ there corresponds an $\tilde{u}\in U_2$ with $\|\tilde{u}\|_{U_2}\leqslant k\,\|Cu\|_{\bar{L}}$ and $C\tilde{u}=Cu$. By definition, for each $u\in U_2=L^2(\alpha(t_1), t_1; R^r)$ we have

$$Cu=C(\cdot)\,u(\cdot)=C(\cdot)\,J(\cdot)\,u(\cdot)==C(\cdot)\,\tilde{u}(\cdot)=C\tilde{u}$$

where $\tilde{u}(\cdot)=J(\cdot)\,u(\cdot)$; clearly,

$$\tilde{u}(\cdot)=J(\cdot)\,u(\cdot)\in U_2 .$$

Therefore the following condition is sufficient for (C) to hold:

— there exists a constant $k>0$ such that for each $v\in R^r$ (C) and almost every $t\in[\alpha(t_1), t_1]$

$$|J(t)\,v|\leqslant k\,|C(t)\,v| .$$

Condition (C) is satisfied in the case $C(t)=C_0=\text{const}$; the existence of the constant k follows from condition (C) applied to C_0 as an operator in R^r.

If $n=r=1$, and there is a constant k such that

$$|C(t)|\geqslant k$$

for every t such that $C(t)\neq 0$, then (C) also holds. This condition is also necessary for *im* C to be closed in this case.

Similar conditions can be given for the case of continuous controls. Note that if the system with H, $H_1=$identity is completely controllable, then the attainable subspace of the same system with H, $H_1\neq$identity is closed if *im* H is. The above considerations apply to the mapping H as well.

4. Examples

The first example shows that the attainable subspace can be a closed proper subspace of the target space. The other one presents the case of the attainable subspace being a proper dense subspace of the target space. Nonzero Lagrange multipliers do not exist in this case, but the same problem set in a different target space becomes normal $(\lambda_0\neq 0)$ (compare Lemma 1).

Example 1

$$\begin{cases} \dot{x}_1=x_2 \\ \dot{x}_2=x_2(t-1)+u, \quad t\in[0, 2], \end{cases} \tag{32}$$

$x(t)=0$, $t\in[-1, 0]$. The control space is $U=L^2(0, 2)$, the target space $L=W_1= =W_1^2(1, 2)\times W_1^2(1, 2)$. Solving (32) by the method of steps one obtains

$$x_2(t)=\int_0^t u(s)\,ds,\ x_1(t)=\int_0^t\int_0^s u(r)\,dr\,ds, \qquad 0\leqslant t\leqslant 1,$$

$$\dot{x}_2(t)=\int_0^{t-1} u(s)\,ds+u(t), \qquad \dot{x}_1(t)=\int_0^t u(s)\,ds+\int_0^{t-1}\int_0^s u(r)\,dr\,ds, \qquad 1\leqslant t\leqslant 2.$$

According to what was said in the preceding section, it suffices to prove that the operator

$$\bar{S}\colon u\mapsto\begin{pmatrix} \dot{x}_1|_{[1, 2]} \\ \dot{x}_2|_{[1, 2]} \end{pmatrix}$$

has the closed range in $L^2(1, 2)\times L^2(1, 2)$. Take the sequence (x_1^n, x_2^n) of solutions of (32) such that

$$\dot{x}_1^n|_{[1, 2]}\xrightarrow[L^2]{}\bar{w}_1 , \tag{33}$$

$$\dot{x}_2^n|_{[1, 2]}\xrightarrow[L^2]{}\bar{w}_2 . \tag{34}$$

Since $\dot{x}_1^n(t)=\dot{x}_1^n(1)+\int_1^t \dot{x}_2^n(s)\,ds$, $1\leqslant t\leqslant 2$, then (33) and (34) imply that the sequence $\{\dot{x}_2^n(1)\}\subset R^n$ satisfies the Cauchy condition and is therefore convergent to $w^0\in R^n$. Hence we conclude that $\bar{w}_1$ is absolutely continuous, $\bar{w}_1(1)=w^0$ and $\dot{\bar{w}}_1=\bar{w}_2$. Take $u^0\in U$ defined by

$$u^0(t)=\begin{cases} w^0, & 0\leqslant t\leqslant 1 \\ w_2(t)-(t-1)\,w^0, & 1<t\leqslant 2 . \end{cases}$$

Denote by (x_1^0, x_2^0) the solution of (32) corres-ponding to u^0. We have:

$$\dot{x}_2^0|_{[1, 2]}=\bar{w}_2$$

$$\dot{x}_1^0(1)=w^0=\bar{w}_1(1)$$

$$\ddot{x}_1^0|_{[1, 2]}=\bar{w}_2=\dot{\bar{w}}_1 ,$$

then

$$\dot{x}_1^n|_{[1, 2]}\xrightarrow[L^2]{}\dot{x}_1^0|_{[1, 2]}$$

$$\dot{x}_2^n|_{[1, 2]}\xrightarrow[L^2]{}\dot{x}_2^0|_{[1, 2]}$$

so that the attainable subspace of (32) is closed. It is different from the whole target space, since $x_1|_{[1, 2]}$ having the absolutely continuous derivative $\dot{x}_1|_{[1, 2]}$ cannot be an arbitrary function from $W_1^2(1, 2)$.

Example 2.

$$\begin{cases} \dot{x}_1 = u \\ \dot{x}_2 = x_1(t-1), \quad t \in [0, 3], \end{cases} \tag{35}$$

$$x_1(t) = x_2(t) = 0, \quad t \in [-1, 0].$$

The control $u \in L^2(0, 3)$ is sought, steering the system (35) to the terminal condition

$$x_1(t) = t - 1,$$
$$x_2(t) = \tfrac{1}{2}(t-2)^2, \quad t \in [2, 3],$$

and minimizing the functional

$$Q(x, u) = J(u) = \frac{1}{2} \int_0^3 (u(t) - v(t))^2 \, dt$$

where

$$v(t) = \begin{cases} 0, & 0 \leqslant t \leqslant 3/2 \\ 1, & 3/2 \leqslant t \leqslant 3. \end{cases}$$

The solution of (35) is

$$x_1(t) = \int_0^t u(s)\, ds, \quad x_2(t) \int_0^{t-1} \int_0^s u(r)\, dr\, ds, \quad t \in [2, 3] \tag{36}$$

Set $U = L^2(0, 3)$, $W_1 = W_1^2(2, 3) \times W_1^2(2, 3)$; the operator $S: U \to W_1$ is given by (36). It is easy to see that if $Su_1 = Su_2$, then $u_1(t) = u_2(t)$, $t \in [1, 3]$, and that the control $\hat{u}$,

$$\hat{u}(t) = \begin{cases} 0, & 0 \leqslant t < 1, \\ 1, & 1 < t \leqslant 3, \end{cases}$$

is the only optimal control; on the interval $[1, 3]$ it is defined uniquely by the terminal condition, on $[0, 1]$ by the minimization of Q. It can be verified that if $(l_1, l_2) \in W_1^* (= W_1)$, then

$$S^* \begin{pmatrix} l_1 \\ l_2 \end{pmatrix}(t) = \begin{cases} l_1(3) + (3-t)\, l_2(3) - \dot{l}_2(2), & 0 \leqslant t < 1, \\ l_1(3) + (3-t)\, l_2(3) - \dot{l}_2(t+1), & 1 \leqslant t < 2, \\ l_1(3) + \dot{l}_1(t), & 2 \leqslant t \leqslant 3. \end{cases} \tag{37}$$

S^* is injective. Indeed, let

$$S^* \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = 0. \tag{38}$$

Then

$$l_1(3) + (3-t)\, l_2(3) - \dot{l}_2(2) = 0, \quad 0 \leqslant t < 1, \tag{39}$$

$$-l_2(3) - \ddot{l}_2(t+1) = 0, \quad 1 < t < 2, \tag{40}$$

$$l_1(3) + \dot{l}_1(t) = 0, \quad 2 \leqslant t \leqslant 3. \tag{41}$$

Equation (40) was obtained by differentiating (38) for $1 < t < 2$. (39) implies $l_2(3) = 0$, hence from (40) $\dot{l}_2 = 0$ and $l_2 = 0$; this and (39) yields $l_1(3) = 0$ and from (41) it obtains that also $l_1 = 0$.

Ker $S^* = \{0\}$ implies *im* S is dense in W_1; *im* $S \neq W_1$ since $x_2|_{[2,3]}$ has the derivative absolutely continuous.

Since the set of admisible controls $M = U$, the Lagrange multipliers λ_0, (l_1, l_2) should satisfy (compare Theorem 1)

$$\lambda_0 J_{\hat{u}} - S^* \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = 0. \tag{42}$$

If $\lambda_0 = 0$, then by injectivity of S^* also $\begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. But λ_0 must be zero, otherwise (42) cannot be satisfied. Indeed,

$$J_{\hat{u}}(t) = \hat{u}(t) - v(t) = \begin{cases} 0, & 0 \leqslant t < 1, \\ 1, & 1 \leqslant t < 3/2, \\ 0, & 3/2 \leqslant t \leqslant 3, \end{cases}$$

and $\lambda_0 J_{\hat{u}}$ is not absolutely continuous in $[1, 2]$ unless $\lambda_0 = 0$. On the other hand, the function $S^* \begin{pmatrix} l_1 \\ l_2 \end{pmatrix}$ is always absolutely continuous in $[1, 2]$. Therefore the only λ_0, (l_1, l_2) satisfying (42) must be zero.

Take $\tilde{W}_1 = W_1^2(2, 3) \times W_2^2(2, 3)$; $W_2^2(2, 3)$ is the space of functions with second derivative square integrable, endowed with the scalar product

$$\langle w_1, w_2 \rangle = w_1(3)\, w_2(3) + \dot{w}_1(3)\, \dot{w}_2(3) + \int_2^3 \ddot{w}_1(t)\, \ddot{w}_2(t)\, dt.$$

The operator S, defined by (36) can be considered as an operator $\tilde{S}$ from U to $\tilde{W}_1$; its adjoint is equal to

$$\tilde{S}^* \begin{pmatrix} l_1 \\ l_2 \end{pmatrix}(t) = \begin{cases} l_1(3) + (2-t)\, l_2(3) + \dot{l}_2(3), & 0 \leqslant t < 1 \\ l_1(3) + (2-t)\, l_2(3) + \dot{l}_2(3) + \ddot{l}_2(t+1), & 1 \leqslant t < 2 \\ l_1(3) + \dot{l}_1(t), & 2 \leqslant t \leqslant 3. \end{cases}$$

Put $\hat{l}_1(3) = \hat{l}_2(3) = \dot{\hat{l}}_2(3) = 0$ and $\dot{\hat{l}}_1(t) = 0$,

$$\ddot{\hat{l}}_2(t) = \begin{cases} 1, & 2 \leqslant t < 5/2 \\ 0, & 5/2 \leqslant t \leqslant 3. \end{cases}$$

Then

$$J_{\hat{u}} - \tilde{S}^* \begin{pmatrix} \hat{l}_1 \\ \hat{l}_2 \end{pmatrix} = 0.$$

It is interesting to observe that *im* $\tilde{S}$ is a proper (closed) subspace of $\tilde{W}_1$; it can be verified letting $l_1(3) = 1$, $\dot{l}_2(3) = -1$, $\dot{l}_1(t) \equiv -1$, $l_2(3) = 0$, $\ddot{l}_2(t) \equiv 0$:

$$\tilde{S}^* \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} = 0.$$

hence *im* S cannot be dense in $\tilde{W}_1$, by (i) theorem 0.

5. Conclusions

The necessary conditions presented in [3], [7] and in this paper are rather cumbersome and it seems impossible to apply them to solving analytically more complicated problems which can arise in an engineer's practice. It seems therefore that the problems of control to targets in function spaces should be solved numerically by the methods using the penalty on the terminal constraints. It is here that the theorems like Theorem 3 and 4 can be helpful, since they contain information about the adjoint equations and suggest the suitable choice of the target space. Moreover many convergence theorems concerning the penalty function methods require the existence of nonzero Lagrange multipliers.

It seems clear, for instance, from the discussion at the end of Section 2 that the penalty on the terminal constraints should be of the form

$$|x(t_1)|^2+\int_{\alpha(t_1)}^{t_1}|\dot{x}(t)|^2\,dt$$

rather than

$$\int_{\alpha(t_1)}^{t_1}|x(t)|^2\,dt$$

or

$$\max_{\alpha(t_1)\leqslant t\leqslant t_1}|x(t)|\,.$$

In special cases, some other penalty functions can be used, depending on the shape of the norm in the space in which the attainable subspace is closed.

These remarks apply to partial differential equations as well. In general the target space should be chosen to be the largest space in which the attainable subspace is closed to avoid the singularities (see Lemma 1) and too complicated and irregular multipliers on the other hand.

ACKNOWLEDGMENT. The author wishes to thank doc. dr hab. A. P. Wierzbicki to whom he owes an improvement of the basic Theorem 1.

Note added in proof

The following recent papers are relevant to the topics pursued in the paper:

1) H. T. Banks, M. Q. Jacobs, An attainable sets approach to optimal control of functional differential equations with function space boundary conditions, J. Diff Equat. **13** (1973), 127—149.
2) H. T. Banks, M. Q. Jacobs, C. E. Langenhop — Characterisation of the controlled states in $W^{(1)}_2$ of linear hereditary systems, to appear in SIAM J. Control.

In the first paper the linear — quadratic case is studied; in the other one, authors obtained some results concerning the closure of the attainable subspace. The problem, when the attainable subspace is closed in W_1^2 was recently solved in

3) S. Kurcyusz, A. W. Olbrot — On the closure of the attainable subspace of linear time-lag systems, to appear.

References

1. Arrow K. J., Hurwicz L., Uzawa H., Studies in linear and nonlinear programming. Stanford (Cal.) 1958.
2. Aubin J. P., Characterization of the set of constraints for which the necessary conditions for optimization hold. *SIAM J. Contr.* **8** (1970) 148—162.
3. Banks H. T., Kent G. A., Control of functional differential equations of retarded and neutral type to target sets in function space. *SIAM J. Contr.* **10** (1972) 567—593.
4. Dunford N., Schwartz J., Linear operators. Pt. 1. New York 1958.
5. Girsanow I. W., Lectures on mathematical theory of extremal problems (in Russian). Moscow 1970.
6. Golshteyn J. G., Duality theory in mathematical programming and its applications (in Russian). Moscow 1971.
7. Jacobs M. Q., Kao T. J., An optimum settling problem for time lag systems *J. Math. Anal. Appl.* **40** (1972), 1—21.
8. Lee E. B. Markus L., Foundations of optimal control theory. New York 1967 (Russian translation Moscow 1972).
9. Lions J. L., Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles. Paris 1968 (Russian translation Moscow 1972).
10. Neustadt L. W., A general theory of extremals. *J. Comput. a. Syst. Sci.* **3** (1969) 57—92.
11. Fiacco A.V., McCormick G. P., Nonlinear programming New York 1968 (Russian translation Moscow 1972).
12. Gabasov R., Kirillova F., Qualitative theory of optimal processes (in Russian). Moscow 1971.
13. Kato T., Perturbation theory for linear operators. Berlin 1966 (Russian translation Moscow 1972).
14. Krasnosielski M. A. et. al., Integral equations. (in Russian). Moscow 1968 (Polish translation Warszawa 1972).
15. Vainberg M. M., Variational methods for the investigation of nonlinear operators (in Russian), Moscow 1956.
16. Vainberg M. M., Some problems in the differential calculus in linear spaces (in Russian), Uspekhi Mat. Nauk, **7** (1972), N° 4.

Lokalna zasada maksimum przy ograniczeniach operatorowych i jej zastosowanie do układów z opóźnieniem

W części pierwszej podano warunki konieczne, jakie musi spełniać rozwiązanie zadania

$$\begin{cases}\text{minimalizuj } J(u),\\ \text{przy ograniczeniach } u\in M\subset U,\ S(u)=0\in L,\end{cases}\tag{1}$$

gdzie U, L oznaczają przestrzenie Banacha, J — funkcjonał, S — operator. Zadanie takie było rozważane wielokrotnie od czasu zjawienia się pracy [1]. Przedstawiona w artykule wersja warunków koniecznych jest nieco silniejsza niż dotychczasowe. Podstawowy rezultat części pierwszej jest następujący.

Twierdzenie 1. Niech M będzie zbiorem wypukłym, domkniętym i o niepustym wnętrzu, a operatory $J: \mathcal{O}\to R$, $S: \mathcal{O}\to L$ w sposób ciągły różniczkowalne według Frecheta na niepustym otwartym zbiorze $\mathcal{O}\subset U$. Przypuśćmy, że $\hat{u}$ jest lokalnym rozwiązaniem zadania (1). Wówczas, jeżeli $im\, S_{\hat{u}}'$ (obraz pochodnej Frecheta operatora S w $\hat{u}$) nie jest właściwą podprzestrzenią gęstą w L, to:

(i) istnieją $\lambda_0\geqslant 0$, $l^*\in L^*$, $(\lambda_0, l^*)\neq(0,0)$ takie, że

$$\langle -\lambda_0 J_{\hat{u}}'+S_{\hat{u}}'^{*}l^*,\ \hat{u}-u\rangle\geqslant 0 \quad \forall u\in M \tag{2}$$

(ii) jeżeli $im\, S'_{\hat{u}}$ jest domkniętą podprzestrzenią L, podprzestrzeń styczna do zbioru $S^{-1}(0)$ w punkcie $\hat{u}$ jest równa $ker\, S'_{\hat{u}}$ oraz $int\,(M-\hat{u})\cap ker\, S'_{\hat{u}} \neq \emptyset$, to $\lambda_0 \neq 0$ w (2).

Pokazano również, że założenie o gęstości $im\, S'_{\hat{u}}$ w L nie może być osłabione.

W dalszym ciągu części pierwszej rozważono problem sterowania optymalnego zapisany abstrakcyjnie przy użyciu równań operatorowych spełniających tzw. formułę Greena. Klasa takich równań operatorowych obejmuje równania różnicowe, różniczkowe zwyczajne i cząstkowe, wreszcie różniczkowe z opóźnieniem. Dla tej ogólnej klasy równań wyprowadzono z twierdzenia 1 warunki optymalności: równania sprzężone, warunki transwersalności i nierówność wariacyjną, będące dosyć przejrzystym uogólnieniem relacji znanych w teorii sterowania optymalnego poszczególnych typów układów.

W części drugiej artykułu przedstawiono zastosowanie wyników otrzymanych w części pierwszej do optymalizacji układów z opóźnieniem przy ograniczeniu równościowym na końcowy stan zupełny. Kwestia istnienia zmiennych sprzężonych (mnożników Lagrange'a) była badana najpierw w [3]: wyprowadzono tam ogólne warunki optymalności, nie gwarantujące jednak nietrywialności mnożników Lagrange'a. W pracy [7] dowiedziono niezerowości zmiennych sprzężonych przy założeniu zupełnej sterowalności i przy braku ograniczeń na sterowanie.

W artykule niniejszym podano warunki konieczne optymalności, gwarantujące niezerowość zmiennych sprzężonych, dla problemu ogólniejszego. Przeprowadzono też dyskusję otrzymanych wyników. W świetle założeń twierdzenia 1 istotny staje się dobór mocy topologii w przestrzeni zupełnych stanów końcowych L oraz zgodność między U a L. Przedstawiono dwie wersje słabej (lokalnej) zasady maksimum dla dwóch różnych układów U i $L-U=L^2(t_0, t_1; R^n)$, $L=W_1^2([t_1-h, t]; R^n)$ oraz $U=C(t_0, t_1; R^*)$, $L=C^1(t_1-h, t_1; R^n)$.

W punkcie (ii) twierdzenia 1 występuje warunek domkniętości podprzestrzeni $im\, S'_{\hat{u}}$, która dla układów z opóźnieniem równa jest podprzestrzeni sterowalnych (osiągalnych) stanów zupełnych. Podano wstępną dyskusję tego warunku dla układów z opóźnieniem.

Przytoczono również dwa przykłady. Jeden z nich ilustruje zależność między istnieniem niezerowych mnożników Lagrange'a a doborem takiej przestrzeni stanów zupełnych, w której spełniony byłby warunek domkniętości $im\, S'_{\hat{u}}$.

Локальный принцип максимума при операторных ограничениях и применение его к системам с запаздыванием

Статья состоит из двух частей.

В первой части даются необходимые условия, которые должно удовлетворять решение задачи.

$$\begin{cases} \text{минимизация } J(u) \\ \text{при ограничениях } u \in M \subset U,\ S(u)=0 \in L, \end{cases} \tag{1}$$

где U, L обозначают банаховы пространства, J — функционал, S — оператор. Эта задача рассматривалась неоднократно с момента появления работы [1]. Представленная в статье версия необходимых условий несколько сильнее предыдущих. Основной результат первой части является следующим.

Теорема 1. Пусть M будет выпуклым замкнутым и внутри непустым множеством, а операторы J: $\mathcal{O}\to R$, S: $\mathcal{O}\to L$ непрерывно дифференцируемы по Фрешету в непустом открытом множестве $\mathcal{O}\subset U$. Предположим, что $\hat{u}$ является локальным решением задачи (1). Тогда, если $im\, S'_{\hat{u}}$ (образ производной Фрешета оператора S в $\hat{u}$) не является собственным плотным подпространством b L, то:

(i) Существуют $\lambda_0 \geq 0$, $l^* \in L^*$, $(\lambda_0, l^*) \neq (0, 0)$ такие, что

$$\langle -\lambda_0 J'_{\hat{u}} + S'^{*}_{\hat{u}} l^*, \hat{u}-u \rangle \geq 0 \quad \forall u \in M \tag{2}$$

(ii) Если $im\, S'_{\hat{u}}$ является замкнутым подпространством L, подпространство касательное к множеству $S^{-1}(0)$ в точке $\hat{u}$ равно $ker\, S'_{\hat{u}}$ а также $int\,(M-\hat{u})\cap ker\, S'_{\hat{u}} \neq \emptyset$, то $\lambda_0 \neq 0$ b (2)

Показано также, что предпосылка о плотности $im\, S'_{\hat{u}}$ в L не может быть ослаблена.

Далее в первой части рассмотрена проблема оптимального управления, абстрактно записанная с помощью операторных уравнений удовлетворяющих так называемую формулу Грина. Класс таких операторных уравнений охватывает разностные уравнения, дифференциальные уравнения обыкновенные и с частными производными и наконец дифференциальные уравнения с запаздыванием. Для этого общего класса уравнений выведены из Теоремы 1 условия оптимальности: сопряженные уравнения, условия трансверсальности и вариационное неравенство, являющиеся довольно ясным обобщением соотношений, известных из теории оптимального управления отдельных типов систем.

Вторая часть статьи представляет применение результатов полученных в первой части для оптимизации систем с запаздыванием при ограничениях в виде равенств на конечное полное состояние. Проблема существования сопряженных переменных (множителей Лагранжа) исследовалась в начале в [3]: там выведены общие условия оптимальности, не гарантирующие однако нетривиальность множителей Лагранжа. В работе [7] доказано существование ненулевых сопряженных переменных при предположении полной управляемости и при отсутствии ограничений на управление.

В данной статье приведены необходимые условия оптимальности, гарантирующие ненулевые значения сопряженных переменных, для общей проблемы. Приведено также рассмотрение полученных результатов. Учитывая предположения Теоремы 1 существенным становится подбор мощности топологии в пространстве полных конечных состояний L, а также согласованность между U и L. Представлены две версии ослабленного (локального) принципа максимума для двух разных систем U и $L-U=L^2(t_0, t_1; R^n)$, $L=W_1^2([t-h, t_1]; R^n)$ а также $U=C(t_0, t_1; R^n)$, $L=C^1(t_1-h, t_1; R^n)$.

В пункте (ii) Теоремы 1 имеется условие замкнутости подпространства $im\, S'_{\hat{u}}$, которые для систем с запаздыванием эквивалентно подпространству управляемых (достигаемых) полных значений. Дается предварительное рассмотрение этого условия для систем с запаздыванием.

Приведены также два примера. Один из них иллюстрирует зависимость между существованием ненулевых множителей Лагранжа и подбором такого пространства полных состояний в котором выполнялось бы условия замкнутости $im\, S'_{\hat{u}}$.