THE MAXIMUM PRINCIPLE FOR RELAXED HEREDITARY DIFFERENTIAL SYSTEMS WITH FUNCTION SPACE END CONDITION*

FRITZ COLONIUS†

Abstract. This paper contains a proof of the global pointwise maximum principle for relaxed hereditary differential systems with general function space end condition. First a multiplier theorem establishes the existence of Lagrange multipliers (l_0, l), where $l_0 \in \mathbb{R}_+$ and l is in the dual of the Sobolev space $W^{n,\infty}[-r, 0]$. Then l can be identified with an element of $W^{n,\infty}[-r, 0]$ provided that the optimal trajectory satisfies a certain regularity condition. This yields two equivalent forms of the maximum principle. Using the results on regular reachability obtained in a companion paper, the maximum principle is shown to be—in a certain sense—generically valid.

Introduction. This paper deals with necessary optimality conditions in the following control problem for hereditary differential systems:

$$(0.1) \qquad \text{Minimize} \int_{t_0}^{t_1} g(x(t), u(t), t) \, dt$$

subject to

$$(0.2) \qquad \dot{x}(t) = f(x_t, u(t), t) \qquad (t \in T := [t_0, t_1]),$$

$$(0.3) \qquad x_{t_0} = \varphi_0,$$

$$(0.4) \qquad h(x(t), t) = 0 \qquad \text{all } t \in T_1 := [t_1 - r, t_1],$$

$$\widetilde{(0.5)} \qquad u(t) \in \Omega(t) \qquad (t \in T),$$

where $x_t(s) := x(t+s) \in \mathbb{R}^n$, $s \in [-r, 0]$ and $t_1 - r > t_0$, $0 \leq r < \infty$, $g : \mathbb{R}^n \times \mathbb{R}^m \times T \to \mathbb{R}$ $f : C^n[-r, 0] \times \mathbb{R}^m \times T \to \mathbb{R}^n$, $h : \mathbb{R}^n \times T_1 \to \mathbb{R}^k$, $\varphi_0 \in C^n[-r, 0]$ and $\Omega(t) \subset \Omega_0 \subset \mathbb{R}^m$, Ω_0 compact, are given and $(t \in T)$ means for Lebesgue almost all $t \in T$. r denotes the length of the time delay. The *state* at time t of the hereditary system (0.2) is given by the function segment x_t. Hence, (0.4) is a condition for the final state x_{t_1}. This infinite-dimensional or "function space" end condition appears appropriate if the behavior of the system after t_1 is of any interest. We call h the output function of the system (0.2). The function space end condition (0.4)—opposed to a finite dimensional one $h(x(t_1), t_1) = 0$—presents particular difficulties which were dealt with in a series of papers: Banks/Jacobs [2], Banks/Kent [3], Barbu [5], Barbu/Precupanu [6, Chap. 4, § 3], Bien [9], Bien/Chyung [10], Buehler [12], Colonius [13], [14], [16], Colonius/Hinrichsen [17], [18], Das [19], Jacobs [25], Jacobs/Kao [26], Kamenskii [27], [28], Kent [33], [34], Kurcyusz [35], Olbrot [39], Utthoff [45] (see also the surveys given in Banks [1], Banks/Manitius [4], Kamenskii/Skubachevskii [32]). Apparently, the first formulation of such a problem—in the context of the calculus of variations—was given by Elsgolts (see Zverkin et al. [54]).

It is appropriate to discuss the relation of the present paper to Banks/Kent [3], Barbu [5], Bien/Chyung [10], Olbrot [39] and Colonius [16] also dealing with constraints like $\widetilde{(0.5)}$.

Banks and Kent split the end condition into two conflicting inequality constraints and use methods by Neustadt in order to prove that the maximum principle is a necessary and, in case of normality and convexity, also sufficient optimality condition.

However, due to their problem formulation, the Lagrange multipliers corresponding to the two conflicting inequality constraints may eliminate each other. Hence, in the necessity part, nontriviality cannot be guaranteed.

Barbu [5] uses the methods of convex analysis and the existence theory of differential equations associated with nonlinear monotone operators in Hilbert space. Problems with fixed final states and pointwise control constraints are formally included [5, Problem (2.4)–(2.6) and § 5]. However, the required assumption (local reachability in $W^{2,n}$-norm) cannot be satisfied for bounded sets of admissible control values (see the discussion in Colonius [16, § 4]).

Colonius [16] gives a special approach to linear time invariant single delay systems with fixed final states. In this case stronger results than in the present paper are obtainable.

Bien and Chyung [10] transform the pure phase constraint into a mixed control/phase variable constraint using a classical device (cf. Pontryagin, et al. [41]). For a trajectory x, condition (0.4) is—under sufficient smoothness conditions— equivalent to

$$(0.6) \qquad 0 = h(x(t_1 - r), t_1 - r),$$

$$(0.7) \qquad 0 = \frac{d}{dt} h(x(t), t) = \frac{\partial}{\partial x} h(x(t), t)\dot{x}(t) + \frac{\partial}{\partial t} h(x(t), t)$$

$$= \frac{\partial}{\partial x} h(x(t), t) f(x_t, u(t), t) + \frac{\partial}{\partial t} h(x(t), t) \quad \text{a.a. } t \in T_1.$$

They generalize the theory of Makowski and Neustadt [37] to hereditary systems with a single constant delay (Olbrot [39, Remark (6.3c)] proposes a similar procedure having reduced the retarded system to an unretarded one). The obtained maximum principle has only a local (in the sense of Girsanov [20]) form on the final interval T_1. That is, the maximum condition has a differentiated form. It has to be assumed that the optimal solution satisfies a certain a priori condition. This regularity condition also has a local form and involves derivatives along the optimal solution as well as cone approximations to the set of admissible control values. As the authors remark, it appears very difficult to assure the validity of the regularity condition *before* the computation of the optimal solution. Furthermore, the regularity condition requires implicitly that the number m of independent control variables is not less than the dimension k of the output space. This restrictive condition appears also in the optimal control of nonlinear systems with *energy constrained* controls (see Kurcyusz [35]).

The problems connected with the end condition led Olbrot [39] to another problem formulation. He required instead of (0.4) (for the fixed final state problem, where $h(x(t), t) = x(t) - \varphi_1(t - t_1)$, $\varphi_1 : [-r, 0] \to \mathbb{R}^n$) that

$$\|x_{t_1} - \varphi_1\| \leq \varepsilon \quad \text{for a constant } \varepsilon > 0.$$

Here the norm may be taken in various Banach spaces. Then he proved necessary optimality conditions for these much simpler problems. In an engineering interpretation, the number ε specifies the accuracy required in reaching the final state, and the norm to measure the distance between the desired and the reached final state can be chosen on the basis of technological requirements. However, it is not clear what happens for $\varepsilon \to 0$; in particular, the problem might become ill behaved for small ε.

It appears more satisfactory to require that the end condition be fulfilled with *arbitrary accuracy*, i.e., we minimize over sequences of controls and corresponding

trajectories satisfying approximately the end condition (0.4). This can be achieved by a "relaxation" of the problem (see Young [52], Warga [48]). Following the approach by Warga [48], we consider the set $\mathscr{S}^{\#}$ of relaxed controls and in the "preproblem" (0.1)–(0.4), $(\widetilde{0.5})$ replace the condition $(\widetilde{0.5})$ by

$$(0.5) \qquad\qquad\qquad v \in \mathscr{S}^{\#}$$

and insert v in (0.1), (0.2) instead of u. We denote (0.1)–(0.5) as Problem 1 and study necessary optimality conditions for an optimal solution (x^0, v^0) of this problem, which is well defined under the assumptions stated in § 1. This extends the approach in Colonius [13], [14], where the fixed final state problem was treated, to the general function space end condition (0.4).

The relaxation of the problem will allow us to weaken the regularity assumption needed in Bien/Chyung [10]. In particular, the condition $m \geqq k$ on the dimensions of the control and output spaces, respectively, is no longer necessary.

We obtain a global, pointwise maximum principle, provided that the optimal trajectory satisfies a certain regularity condition referring to the infinite dimensional part (0.7) of the end condition. This regularity condition has a similar form to those given—in the theory of ordinary differential systems—by Warga [46], [47], Schwarzkopf [42] for inequality and by Schwarzkopf [43], [44] for equality constraints on control and phase variables.

Two equivalent forms of the maximum principle are stated corresponding to the formulations (0.4) and (0.6), (0.7) of the end condition. We exploit the results in Colonius [15] (this issue, pp. 675–694) in order to show that the maximum principle is—in a certain sense—generically valid.

This paper is built up as follows. In § 1, the assumptions are formulated. Section 2 establishes the existence of Lagrange multipliers $(l_0, l_1, l_2) \in \mathbb{R}_+ \times \mathbb{R}^k \times (L_\infty^k(T_1))^*$, provided that the infinite dimensional part of the attainable set of the linearized system has a nonempty interior. In § 3, the regularity assumption on the optimal trajectory is used to regularize l_2. That is, l_2 can be identified with an element of $L_\infty^k(T_1)$. The global pointwise maximum principle is obtained as a straightforward consequence. Section 4 discusses the range of validity of the maximum principle.

We retain the notation and conventions of Colonius [15].

1. Assumptions. The following assumptions will be imposed throughout this paper:

(1.1) The functions $f: C^n[-r, 0] \times \mathbb{R}^m \times T \to \mathbb{R}^n$ and $g: \mathbb{R}^n \times \mathbb{R}^m \times T \to \mathbb{R}$ are continuous in $(\varphi, \omega) \in C^n[-r, 0] \times \mathbb{R}^m$ and $(y, \omega) \in \mathbb{R}^n \times \mathbb{R}^m$, respectively, and measurable in $t \in T$.

(1.2) There are $p, q: \mathbb{R}_+ \times T \to \mathbb{R}_+$ such that for all $\varphi \in C^n[-r, 0]$, $y \in \mathbb{R}^m$ and $\omega \in \Omega_0$,

$$|f(\varphi, \omega, t)| \leqq p(\|\varphi\|_\infty, t) \qquad (t \in T),$$

$$|g(y, \omega, t)| \leqq q(|y|, t) \qquad (t \in T),$$

and $p(s, \cdot) \in L_\infty^1(T)$, $q(s, \cdot) \in L_2^1(T)$ and $p(\cdot, t)$, $q(\cdot, t)$ are monotonically increasing for all $(s, t) \in \mathbb{R}_+ \times T$.

(1.3) The functions f and g are continuously Fréchet differentiable in the first argument; the corresponding derivatives $D_1 f(\varphi, \omega, t)$ and $D_1 g(y, \omega, t)$ are continuous in $(\varphi, \omega\, t)$ and (y, ω), respectively; $h = (h_1, \cdots, h_k): \mathbb{R}^n \times T_1 \to \mathbb{R}^k$ is continuously Fréchet differentiable and the derivative is continuously Fréchet differentiable with

respect to $y \in \mathbb{R}^n$;

$$|D_1 f(\varphi, \omega, t)| \leqq p(\|\varphi\|_\infty, t) \qquad (t \in T),$$

$$|D_1 g(y, \omega, t)| \leqq q(|y|, t) \qquad (t \in T),$$

where p, q are as in (1.2).

(1.4) For each relaxed control $v \in \mathscr{S}^*$, there is a unique trajectory $x \in C^n[t_0 - r, t_1]$ satisfying (0.2) and (0.3).

The conditions (1.1)–(1.3) may be required only in a neighborhood of the optimal solution similarly as in Berkovitz [8], Bates [7]. The assumptions on f coincide with Colonius [15, (1.1)–(1.4)]. Hence, [15, Lemmas 1.1, 1.2, 1.3 and 1.5] are valid here. The continuity (instead of integrability) assumption on $D_1 f(\varphi, \omega, t)$ in t is stronger than desirable. However, it is needed for [15, Lemma 1.2] used in the proof of Lemma 2.1 below. By [15, Thm. 2.1], the relaxed problem admits the interpretation given in the introduction. Furthermore, existence of an optimal solution is guaranteed:

THEOREM 1.1. *If the trajectories x, satisfying (0.2) and (0.3), are uniformly bounded, there exists an optimal solution (x^0, v^0) of Problem 1.*

The proof of this theorem follows from compactness of $\mathscr{S}^*$ and Colonius [15, Lemma 1.3].

In the rest of this paper we assume that (x^0, v^0) is an optimal solution of Problem 1. Furthermore, for the sake of simplicity, we let $\varphi_0(0) = 0$.

2. The abstract maximum principle. In this section, we reformulate Problem 1 in an abstract setting of operators on Banach spaces. Then we obtain the existence of Lagrange multipliers $(l_0, l_1, l_2) \in \mathbb{R}_+ \times \mathbb{R}^k \times (L_\infty^k(T_1))^*$ provided that the output set of the linearized system has a nonempty interior. This yields two different forms of the abstract maximum principle reflecting the two formulations (0.4) and (0.6), (0.7) of the end condition. These two formulations are equivalent under the assumptions stated in § 1.

For the ease of notation, we introduce

$$\mathbf{C}^n(T) := \{x \in C^n(T): x(t_0) = 0\},$$

$$\mathbf{W}^{n,\infty}(T) := \{x \in W^{n,\infty}(T): x(t_0) = 0\}.$$

Observe that the natural embedding of $\mathbf{W}^{n,\infty}(T)$ into $\mathbf{C}^n(T)$ is compact. $\mathbf{W}^{n,\infty}(T)$ is in a natural way isomorphic to $L_\infty^n(T)$. We extend each element x of $\mathbf{W}^{n,\infty}(T)$ (resp. $\mathbf{C}^n(T)$) to a continuous function $x: [t_0 - r, t_1] \to \mathbb{R}^n$ by $x_{t_0} := \varphi_0$.

Define

$$G: \mathbf{C}^n(T) \times \mathscr{N} \to \mathbb{R}$$

$$\text{by } G(x, v) := \int_T g(x(t), v(t), t) \, dt;$$

$$F: \mathbf{C}^n(T) \times \mathscr{N} \to \mathbf{W}^{n,\infty}(T)$$

$$\text{by } F(x, v)(t) := \int_{t_0}^t f(x_s, v(s), s) \, ds, \qquad t \in T;$$

$$B: \mathbf{C}^n(T) \to \mathbb{R}^k$$

$$\text{by } B(x) := h(x(t_1 - r), t_1 - r);$$

$$C:\mathbf{C}^n(T)\times\mathcal{N}\to L_\infty^k(T_1)$$

by $C(x,v)(t):=\dfrac{\partial}{\partial x}h(x(t),t)f(x_t,v(t),t)+\dfrac{\partial}{\partial t}h(x(t),t),\qquad t\in T_1;$

$$K:\mathbf{W}^{n,\infty}(T)\to L_\infty^k(T_1)$$

by $(Kx)(t):=\dfrac{\partial}{\partial x}h(x(t),t)\dot{x}(t)+\dfrac{\partial}{\partial t}h(x(t),t),\qquad t\in T_1.$

Then (B,K) and (B,C) may be considered as maps with values in

$$W^{k,\infty}(T_1)=\mathbb{R}^k\times L_\infty^k(T_1).$$

Define the solution map $S:\mathcal{S}^\#\to W^{n,\infty}(T)$ by $S(v)=x$, where x is the response of system (0.2) with initial condition (0.3) to the relaxed control v.

Problem 1 is equivalent to the abstract one, Problem 2.

Problem 2.

$$\underset{v\in\mathcal{S}^\#}{\text{Minimize }} G(S(v),v)$$

subject to

$$(B,K)(S(v))=0.$$

Observe that

(2.1)
$$(B,K)(S(v))=(B(S(v)),K(S(v)))=(B(S(v)),C(S(v),v))$$
$$=(B,C)(S(v),v).$$

Remark 2.1. The problem formulation with the operator (B,K) will yield the form of the maximum principle proposed by Banks/Kent [3], the formulation with (B,C) will yield the form given by Bien/Chyung [10].

In the following, we assure differentiability of the involved operators and give concrete representations for the derivatives. First, by the Riesz representation theorem, there is a $(n\times n)$-matrix function η defined on $T\times[t_0-r,t_1]$ such that

(2.2)
$$D_1f(x_s^0,v^0(s),s)x_s=\int_{t_0-r}^s d_t\eta(s,t)x(t)\qquad(s\in T)$$

for all $x\in C^n[t_0-r,t_1]$; here $\eta(\cdot,t)$ is measurable and $\eta(s,\cdot)$ is of bounded variation for $(s,t)\in T\times[t_0-r,t_1]$. The integral is meant in the Riemann–Stieltjes sense.

For all $s\in T$, we require that $\eta(s,\cdot)$ is normalized, i.e., left continuous on (t_0-r,t_1) and $\eta(s,t)=\eta(s,s)=0$, $t_0\le s\le t\le t_1$. This determines $\eta(s,\cdot)$ uniquely. One can show (using Bourbaki [11, §8, exc. 6]) that η is even measurable on the rectangle $T\times[t_0-r,t_1]$. Define $h(x,t)=0$ for $x\in\mathbb{R}^n$, $t\in T\backslash T_1$, and let for $t\in T$:

(2.3)
$$H_x(t):=\left(\frac{\partial}{\partial x_j}h_i(x^0(t),t)\right)_{\substack{i=1,\cdots,k,\\j=1,\cdots,n,}}$$

(2.4)
$$H_{xx_j}(t):=\left(\frac{\partial^2}{\partial x_l\partial x_j}h_i(x^0(t),t)\right)_{\substack{i=1,\cdots,k,\\l=1,\cdots,n,}}$$

(2.5)
$$H_{xt}(t):=\left(\frac{\partial^2}{\partial x_j\partial t}h_i(x^0(t),t)\right)_{\substack{i=1,\cdots,k,\\j=1,\cdots,n.}}$$

LEMMA 2.1. (i) *The maps F and G are linear in $v \in \mathcal{N}$, C is affine in $v \in \mathcal{N}$.*

(ii) *The restrictions of F, G, (B, C) and (B, K) to $\mathbf{W}^{n,\infty}(T) \times \mathcal{S}^*$ are weakly* continuous, S is weakly* continuous.*

(iii) *The maps B, C, F, G, K and S are continuously Fréchet differentiable.*

(iv) *The derivatives have the following form:*

$$[D_1 F(x_0, v_0)x](t) = \int_{t_0}^{t} D_1 f(x_s^0, v^0(s), s)x_s \, ds, \qquad t \in T,$$

$$D_1 G(x^0, v^0)x = \int_T \frac{\partial}{\partial x} g(x^0(t), v^0(t), t)x(t) \, dt,$$

$$DB(x^0)x = H_x(t_1 - r)x(t_1 - r),$$

$$[D_1 C(x^0, v^0)x](t) = \sum_{j=1}^{n} f_j(x_t^0, v^0(t), t)H_{xx_j}(t)x(t)$$

$$+ H_x(t)D_1 f(x_t^0, v^0(t), t)x_t + H_{xt}(t)x(t), \qquad t \in T_1, \text{ where } x_{t_0} := 0,$$

$$[DK(x^0)x](t) = \sum_{j=1}^{n} \dot{x}_j^0(t)H_{xx_j}(t)x(t) + H_x(t)\dot{x}(t) + H_{xt}(t)x(t), \qquad t \in T_1.$$

For $v \in \mathcal{S}^*$, the function $x(v) := DS(v^0)(v - v^0)$ satisfies

$$\dot{x}(t) = D_1 f(x_t^0, v^0(t), t)x_t + f(x_t^0, v(t) - v^0(t), t) \qquad (t \in T),$$

(2.6)

$$x_{t_0} = 0.$$

Proof. Assertion (i) follows directly from the definitions. Assertion (ii) follows from Colonius [15, Lemma 1.1]. The operators B, C, F, G, K are continuously Fréchet differentiable by [15, Lemma 1.2] and the chain rule. This yields also the form of the derivatives as indicated in (iv). Finally, the existence and the form of the derivative $DS(v^0)$ follows from [15, Lemma 1.5]. $\square$

Consider the linearized system (2.6) in its abstract form

(2.7) $$x = D_1 F(x^0, v^0)x + F(x^0, v - v^0), \qquad v \in \mathcal{S}^*,$$

with output

(2.8) $$(DB(x^0), DK(x^0))(x) \in W^{k,\infty}(T_1).$$

Then the chain rule and (2.1) imply

(2.9) $$DK(x^0)x(v) = D_1 C(x^0, v^0)x(v) + C(x^0, v) - C(x^0, v^0), \qquad v \in \mathcal{S}^*.$$

Define the set $\mathcal{A}$ of attainable output values by

$$\mathcal{A} := \{\varphi \in W^{k,\infty}(T_1): \text{there is a } v \in \mathcal{S}^* \text{ such that } \varphi = (DB(x^0), DK(x^0))x(v)\},$$

and let

$$\pi_\infty \mathcal{A} := \{z \in L_\infty^k(T_1): \text{there is } \varphi \in \mathcal{A} \text{ with } z = \dot{\varphi}\}.$$

The set $\mathcal{A}$ will always be considered in $W^{k,\infty}$-norm and the set $\pi_\infty \mathcal{A}$ in L_∞^k-norm. We have

(2.10)

$$\pi_\infty \mathcal{A} = \{z \in L_\infty^k(T_1): \text{there is a } v \in \mathcal{S}^* \text{ such that } z = D_1 C(x^0, v^0)x(v) + C(x^0, v)\},$$

since

$$0 = \frac{d}{dt} h(x^0(t), t) = \frac{\partial}{\partial x} h(x^0(t), t) f(x_t^0, v^0(t), t) + \frac{\partial}{\partial t} h(x^0(t), t)$$

$$= [C(x^0, v^0)](t) \qquad (t \in T_1).$$

THEOREM 2.1. *Let* $v^0 \in \mathscr{S}^*$, $x^0 = S(v^0)$ *be an optimal solution of Problem 2 and assume that* int $\pi_\infty \mathscr{A} \neq \varnothing$. *Then there are nontrivial Lagrange multipliers* $(l_0, l) := (l_0, l_1, l_2) \in \mathbb{R}_+ \times \mathbb{R}^k \times (L_\infty^k(T_1))^*$ *such that*

(2.11)
$$l_0 D_1 G(x^0, v^0) x(v) + l_0 G(x^0, v - v^0) + l_1 \cdot DB(x^0) x(v)$$
$$+ l_2 \circ DK(x^0) x(v) \geqq 0 \quad \text{for all } v \in \mathscr{S}^*.$$

If $0 \in$ int $\pi_\infty \mathscr{A}$, *then* $(l_0, l_1) \neq (0, 0)$, *and if* $0 \in$ int $\mathscr{A}$, *then* $l_0 \neq 0$.

Proof. If $0 \in \partial \pi_\infty \mathscr{A}$, there is a support functional l_2 at 0 to $\pi_\infty \mathscr{A}$, and (2.11) is satisfied with $l_0 = 0$, $l_1 = 0$. Now let $0 \in$ int $\pi_\infty \mathscr{A}$. We verify the assumptions in Colonius [16, Thm. 1.3].

By Lemma 2.1, $v \mapsto G(S(v), v)$ and $v \mapsto (B, K) S(v)$ are continuously Fréchet differentiable and weakly* continuous. By Warga [48, Thm. IV.3.11], $\mathscr{S}^*$ is weakly* compact and convex. The space $L_\infty^k(T_1)$ can weakly* continuously be embedded into the Hilbert space $L_2^k(T_1)$. Hence there are Lagrange multipliers $(l_0, l_1, l_2) \in \mathbb{R}_+ \times \mathbb{R}^k \times (L_\infty^k(T_1))^*$ satisfying (2.11) and $(l_0, l_1) \neq (0, 0)$. By the same theorem, $l_0 \neq 0$ if $0 \in$ int $\mathscr{A}$. $\square$

Remark 2.2. In the case where $0 \in$ int $\mathscr{A}$, Theorem 2.1 follows easily from the multiplier theorem by Zowe/Kurcyusz [53, Thm. 2.1], which does not presuppose properties with respect to weak* topology.

Remark 2.3. By weak* continuity and compactness of $\mathscr{S}^*$, $\pi_\infty \mathscr{A}$ is a weakly* closed convex set in $L_\infty^k(T_1)$. Hence, Phelps [40, Thm. 1] shows that the set M of points in $\pi_\infty \mathscr{A}$ admitting a weakly* continuous support functional $l_2 \neq 0$ is norm dense in the norm boundary of $\pi_\infty \mathscr{A}$. If $0 \in M$, then $(0, 0, l_2) \in \mathbb{R}_+ \times \mathbb{R}^k \times L_1^k(T_1) \subset \mathbb{R}_+ \times \mathbb{R}^k \times (L_\infty^k(T_1))^*$ are Lagrange multipliers satisfying (2.11). This shows that also in the case where int $\pi_\infty \mathscr{A} = \varnothing$ there are "many" points in $\pi_\infty \mathscr{A}$ admitting nontrivial Lagrange multipliers, even with $l_2 \in L_1^k(T_1)$. The same argument applies where int $\mathscr{A} = \varnothing$. Then one obtains Lagrange multipliers $(0, l) \in \mathbb{R}_+ \times W^{k,1}(T_1)$.

COROLLARY 2.1. *Let the assumptions of Theorem 2.1 be satisfied. Then*
(i) *the adjoint variable* $y^K \in (\mathbf{W}^{n,\infty}(T))^*$ *defined by*

$$y^K = D_1 F(x^0, v^0)^* y^K + l_0 D_1 G(x^0, v^0) + DB(x^0)^* l_1 + DK(x^0)^* l_2$$

satisfies

$$l_0 G(x^0, v - v^0) + y^K \circ F(x^0, v - v^0) \geqq 0 \quad \text{for all } v \in \mathscr{S}^*.$$

(ii) *The adjoint variable* $y^C \in (\mathbf{C}^n(T))^*$ *defined by*
$$y^C = D_1 F(x^0, v^0)^* y^C + l_0 D_1 G(x^0, v^0) + DB(x^0)^* l_1 + D_1 C(x^0, v^0)^* l_2$$

satisfies

$$l_0 G(x^0, v - v^0) + y^C \circ F(x^0, v - v^0) + l_2 \circ C(x^0, v) \geqq 0 \quad \text{for all } v \in \mathscr{S}^*.$$

(iii) *The difference* $y^K - y^C \in (\mathbf{W}^{n,\infty}(T))^*$ *satisfies*
$$y^K - y^C = D_1 F(x^0, v^0)^* (y^K - y^C) + [DK(x^0)^* - D_1 C(x^0, v^0)^*] l_2.$$

Proof. Observe that both adjoint equations are uniquely solvable since $\mathrm{Id}_{C(T)^*} - D_1 F(x^0, v^0)^*$ and $Id_{W^{n,\infty}(T)^*} - D_1 F(x^0, v^0)^*$ are isomorphisms, where Id denotes the identity map on the respective spaces. By definition of y^K, (2.11) and (2.7), we find

$$l_0 G(x^0, v - v^0) + y^K \cdot F(x^0, v - v^0)$$

$$= l_0 G(x^0, v - v^0) + [l_0 D_1 G(x^0, v^0) + l_1 \cdot DB(x^0) + l_2 \cdot DK(x^0)]$$

$$\cdot (\mathrm{Id} - D_1 F(x^0, v^0))^{-1} F(x^0, v - v^0)$$

$$= l_0 G(x^0, v - v^0) + l_0 D_1 G(x^0, v^0) x(v) + l_1 \cdot DB(x^0) x(v) + l_2 \cdot DK(x^0) x(v)$$

$$\geqq 0 \quad \text{for all } v \in \mathscr{S}^*.$$

Assertions (ii) and (iii) follow similarly, taking into account (2.9).   □

3. The global pointwise maximum principle. The abstract Theorem 2.1 is only a first step. The optimality conditions involve the multiplier $l_2 \in (L_\infty^k(T_1))^*$ which may not be identifiable with a real function. In order to regularize l_2, we make use of the following notion.

DEFINITION 3.1. A trajectory x^0 satisfying (0.2), (0.3) is called *regular* (with respect to (0.4)) if and only if there exists a neighborhood V of $0 \in \mathbb{R}^k$ such that

$$(3.1) \qquad V \subset \frac{\partial}{\partial x} h(x^0(t), t)[\mathrm{co}\, f(x_t^0, \Omega(t), t) - \dot{x}^0(t)] \qquad (t \in T_1).$$

Observe that 0 is always in the set at the right-hand side. By Colonius [15, Lemma 1.4] x^0 is regular if and only if there is a neighborhood V_∞ of $0 \in L_\infty^k(T_1)$ such that

$$V_\infty \subset \{C(x^0, v) : v \in \mathscr{S}^*\}.$$

The proof of the global pointwise maximum principle is prepared by the following two lemmas, which contain the crux of the proof.

LEMMA 3.1. *If the optimal trajectory x^0 is regular, then $0 \in \mathrm{int}\, \pi_\infty \mathscr{A}$.*

Proof. First we construct an inverse of the map $v \mapsto C(x^0, v)$ from $\mathscr{S}^*$ into $L_\infty^k(T_1)$. Take V in (3.1) as a k-simplex with vertices $e_1, e_2, \cdots, e_{k+1}$. Let $e_0 := 0$. Since $0 \in \mathrm{int}\, V$, we can divide V into $k+1$ subsimplices V^j, $j = 1, \cdots, k+1$ with vertices $e_0, \cdots, e_{j-1}, e_{j+1}, \cdots, e_{k+1}$. Then there are $v^i \in \mathscr{S}^*$ such that

$$e_i = C(x^0, v^i), \qquad i = 1, \cdots, k+1,$$

and we know that

$$e_0 = 0 = C(x^0, v^0).$$

We may assume that for $i = 1, \cdots, k+1$,

$$v^i(t) = v^0(t), \qquad t \in [t_0, t_1 - r].$$

By definition of C,

$$e_i = H_x(t) f(x_t^0, v^i(t) - v^0(t), t) \qquad (t \in T_1).$$

For $t \in T$ we define continuous piecewise affine maps

$$A_t : V \to \mathrm{rpm}(\Omega_0)$$

in the following way:

$$A_t(e^i) := v^i(t), \qquad i = 0, 1, \cdots, k+1.$$

Each $z \in V$ lies in some V^j. Consider the barycentric coordinates α_i with respect to V^j, i.e.,

$$z = \sum_{\substack{i=0 \\ i \neq j}}^{k+1} \alpha_i e_i,$$

where $\alpha_i \geqq 0$, $\sum_{i=0, i \neq j}^{k+1} \alpha_i = 1$, and define

$$A_t z := \sum_{\substack{i=0 \\ i \neq j}}^{k+1} \alpha_i A_t(e_i).$$

We have to show that this is well defined. Clearly, $A_t z \in \mathrm{rpm}(\Omega_0)$. Suppose that $z \in V^j \cap V^l$. Then only those coordinates α_i of z in V^j(resp. V^l) are nonzero for which $e_i \in V^j \cap V^l$. These coordinates have the same value in V^j and V^l. Now let

$$V_\infty := \{z \in L_\infty^k(T_1): z(t) \in V \, (t \in T_1)\}$$

and define $A: V_\infty \to \mathscr{S}^*$ by

$$(Az)(t) := A_t z(t) \qquad (t \in T).$$

Then $Az \in \mathscr{S}^*$, since the coordinate functions of z may be chosen measurable and the support of $(Az)(t)$ is contained in $\Omega(t)$. We have

(3.2) $$C(x^0, Az) = z$$

for all $z \in V_\infty$, since C and A satisfy this equality pointwise.

A is extended to a map $A: L_\infty^k(T_1) \to \mathscr{N}$ by an affine continuation of A_t in the $k+1$ sectors of $\mathbb{R}^n$ corresponding to V^j, $j = 0, 1, \cdots, k$. Then A satisfies a Lipschitz condition for a constant $L, 0$ (see Warga [48, p. 268] for the definition of $\|\cdot\|_{\mathscr{N}}$):

(3.3) $$\|Az_1 - Az_2\|_{\mathscr{N}} \leqq L \|z_1 - z_2\|_\infty.$$

Now consider the equation

(3.4) $$x = D_1 F(x^0, v^0)x + F(x^0, A(z - D_1 C(x^0, v^0)x)) - F(x^0, v^0).$$

Since for fixed $z \in L_\infty^k(T_1)$ the value of the right-hand side depends only on $x_t \in C^n[-r, 0]$, this equation can be written as

$$\dot{x}(t) = f^z(x_t, t) \qquad (t \in T), \quad x_{t_0} = 0.$$

Using (3.3) and the assumptions (1.1)–(1.3) one finds that f^z is continuous in the first argument and measurable in the second argument; furthermore, f^z satisfies a global Lipschitz condition uniformly with respect to z in the first argument. Hence for each $z \in L_\infty^k(T_1)$ there exists a unique solution of this functional differential equation (cf. Hale [22, Thm. 2.3 and p. 55], which by definition is equivalent to (3.4).

Take $z = 0$. Then $x = 0$ solves (3.4) since

$$F(x^0, A(z - D_1 C(x^0, v^0)x)) = F(x^0, A(0)) = F(x^0, v^0).$$

The uniform global Lipschitz condition and Gronwall's inequality imply that x depends continuously on the right-hand side of (3.4). Hence there is $\delta > 0$ such that for all z with $\|z\|_\infty < \delta$,

$$z - D_1 C(x^0, v^0)x \in V_\infty.$$

Thus by construction of A there is $v \in \mathscr{S}^*$ with

(3.5)
$$v = A(z - D_1 C(x^0, v^0)x).$$

Then

(3.6)
$$x = D_1 F(x^0, v^0)x + F(x^0, v - v^0),$$

and by (3.2)

(3.7)
$$C(x^0, v) = z - D_1 C(x^0, v^0)x.$$

Since this holds for all z with $\|z\|_\infty < \delta$, we find by (2.10) that $0 \in \mathrm{int}\ \pi_\infty \mathscr{A}$. □

LEMMA 3.2. *If the optimal trajectory x^0 is regular and (l_0, l_1, l_2) are Lagrange multipliers satisfying (2.11), then l_2 can be identified with a function $\rho \in L_\infty^k(T_1) \subset (L_\infty^k(T_1))^*$.*

Proof. Consider the subspace S of simple functions in $L_\infty^k(T_1)$. S is dense in L_∞ (see, e.g., Hewitt/Stromberg [23, Thm. 11.35]). We shall prove that $l_2|S$ is continuous with respect to L_1-norm on S. Then $l_2|S$ can be extended to a continuous linear functional on $L_1^k(T_1)$ which by duality of L_1 and L_∞ can be identified with an element ρ of $L_\infty^k(T_1)$. Then l_2 and the functional defined by ρ coincide on S, hence, on $L_\infty^k(T_1)$.

The general element $s \in S$ has the form

$$s(t) = \sum_{i=1}^{l} \sum_{j=1}^{k} s_{ij} \chi_{E_i}(t) a_j \qquad (t \in T_1),$$

where $s_{ij} \in \mathbb{R}$, $\{a_j\}$ is a base of $\mathbb{R}^n$ and $\{E_i\}$ is a measurable decomposition of T_1.

Since x^0 is regular, we may assume that $\pm \chi_{E_i}(t) a_j \in V$, where V satisfies (3.1). Then there are $v_{ij}^\pm \in \mathscr{S}^*$ such that

$$\pm \chi_{E_i} a_j = C(x^0, v_{ij}^\pm)$$

and

(3.8)
$$\int_T \|v_{ij}^\pm(t) - v^0(t)\| \, dt < 2\lambda(E_i).$$

Decompose s into its positive and negative parts:

$$s = \sum_{i=1}^{l} \sum_{j=1}^{k} (s_{ij}^+ - s_{ij}^-)\chi_{E_i} a_j,$$

where $s_{ij}^\pm := \max(0, \pm s_{ij})$.

Apply the maximum condition in Corollary 2.1(ii) $2kl$ times in order to obtain

$$l_2(s) = l_2\left(\sum_{i,j} (s_{ij}^+ - s_{ij}^-)\chi_{E_i} a_j\right)$$

$$= \sum_{i,j} \{s_{ij}^+ l_2 \circ C(x^0, v_{ij}^+) + s_{ij}^- l_2 \circ C(x^0, v_{ij}^-)\}$$

$$\geqq -\sum_{i,j} \{s_{ij}^+ [l_0 G(x^0, v_{ij}^+ - v^0) + y^C \circ F(x^0, v_{ij}^+ - v^0)]$$

$$+ s_{ij}^- [l_0 G(x^0, v_{ij}^- - v^0) + y^C \circ F(x^0, v_{ij}^- - v^0)]\}$$

$$\geqq -c_0 \sum_{i,j} (s_{ij}^+ + s_{ij}^-)\lambda(E_i)$$

$$= -c_0 \|s\|_{L_1}$$

for a constant $c_0 > 0$; here we used (3.8) and the assumptions (1.1) and (1.2).

Hence, for $\|s\|_{L_1} \to 0$,

$$\lim \inf l_2(s) \geq 0;$$

the same argument for $-s$ proves that $l_2(s) \to 0$ for $\|s\|_{L_1} \to 0$. Thus, l_2 is continuous on S in L_1-norm and the lemma is proven. □

The following maximum principle has the form proposed by Banks/Kent [3] for a more general class of systems (including neutral equations) with fixed final state.

THEOREM 3.1 (BK-form of the maximum principle). *Let x^0, v^0 be an optimal solution of Problem 1 and assume that x^0 is a regular trajectory.*

Then there exist Lagrange multipliers $(l_0, l_1, \rho) \in \mathbb{R}_+ \times \mathbb{R}^k \times L_\infty^k(T_1)$ such that $(l_0, l_1) \neq (0, 0)$ and the adjoint variable $\psi^K \in L_\infty^k(T_1)$ defined by

(i) $\psi^K(t) = -\displaystyle\int_t^{t_1} \eta(s, t)^* \psi^K(s)\, ds + l_0 \int_t^{t_1} \frac{\partial}{\partial x} g(x^0(s), v^0(s), s)\, ds$

$$+ H_x(t_1 - r)^* l_1 + \int_{T_1} \left[\sum_{j=1}^n \dot{x}_j^0(s) H_{xx_j}(s)^* + H_{xt}(s)^* \right] \rho(s)\, ds$$

for $t \in [t_0, t_1 - r]$,

$$\psi^K(t) = -\int_t^{t_1} \eta(s, t)^* \psi^K(s)\, ds + l_0 \int_t^{t_1} \frac{\partial}{\partial x} g(x^0(s), v^0(s), s)\, ds$$

$$+ H_x(t)^* \rho(t) + \int_t^{t_1} \left[\sum_{j=1}^n \dot{x}_j^0(s) H_{xx_j}(s)^* + H_{xt}(s)^* \right] \rho(s)\, ds$$

for $t \in (t_1 - r, t_1]$, satisfies

(ii) $l_0 g(x^0(t), v^0(t), t) + \psi^K(t) f(x_t^0, v^0(t), t) \geq l_0 g(x^0(t), \omega, t) + \psi^K(t) f(x_t^0, \omega, t)$

for all $\omega \in \Omega(t)$, a.a. $t \in T$.

Proof. In view of Corollary 2.1(i), Lemma 3.1 and Lemma 3.2, we have to compute adjoint operators. By partial integration, $DK(x^0)^* l_2 \in (\mathbf{W}^{n,\infty}(T))^*$ can be identified with the following element of $W^{n,\infty}(T)$:

$$DK(x^0)^* l_2 = \begin{cases} \displaystyle\int_{T_1} \left[\sum_{j=1}^n \dot{x}_j^0(t) H_{xx_j}(t)^* + H_{xt}(t)^* \right] \rho(t)\, dt, & t \in [t_0, t_1 - r], \\[4mm] \displaystyle\int_t^{t_1} \left[\sum_{j=1}^n \dot{x}_j^0(s) H_{xx_j}(s)^* + H_{xt}(s)^* \right] \rho(s)\, ds + H_x(t)^* \rho(t), & t \in (t_1 - r, t_1]. \end{cases}$$

Similarly,

$$[l_0 D_1 G(x^0, v^0)](t) = -l_0 \int_t^{t_1} \frac{\partial}{\partial x} g(x^0(s), v^0(s), s)\, ds, \qquad t \in T,$$

$$[DB(x^0)^* l_1](t) = \begin{cases} H_x(t_1 - r)^* l_1, & t \in [t_0, t_1 - r], \\ 0, & t \in (t_1 - r, t_1]. \end{cases}$$

Then y^K can be identified with an element $\psi^K \in L_\infty^k(T)$ and

$$[D_1 F(x^0, v^0)^* y^K](t) = -\int_t^{t_1} \eta(s, t)^* \psi^K(s)\, ds, \qquad t \in T.$$

This yields the adjoint equation (i). Furthermore,

$$y^K \circ F(x^0, v - v^0) = \int_T \psi^K(t) f(x_t^0, v(t) - v^0(t), t)\, dt,$$

and the maximum condition in integral form follows. This implies (ii) by standard arguments (see, e.g., Warga [48, Thm. VI.2.3]). □

Remark 3.1. Since $\eta(s, \cdot)$ is of bounded variation, one can identify $\psi^K|[t_0, t_1 - r]$ with a function of bounded variation. For special systems, e.g., systems with a single constant delay, $\psi^K|[t_0, t_1 - r]$ is even absolutely continuous (see Banks/Kent [3, p. 583]).

Remark 3.2. One can give the following nontriviality condition in terms of the adjoint variable ψ^K:

$$(0, 0, 0) \neq (l_0, \psi^K(t_1 - r), \psi^K|T_1) \in \mathbb{R}_+ \times \mathbb{R}^n \times L_\infty^n(T_1).$$

Assume that $(l_0, \psi^K(t_1 - r), \psi^K|T_1)$ is trivial. Regularity implies that the multiplication operator

$$\tilde{H}: L_\infty^n(T_1) \to L_\infty^k(T_1), \qquad (\tilde{H}z)(t) := H_x(t)z(t) \qquad (t \in T_1)$$

is surjective. Hence Kurcyusz/Olbrot [36, Lemmas 3 and 4] imply that

$$\operatorname{rank} H_x(t)^* = \operatorname{rank} H_x(t) = k$$

for *all* $t \in T_1$ and that the generalized inverse $[H_x(t)^*]^+$ of $H_x(t)^*$ is bounded on T_1.

Using the adjoint equation (i), we find that ρ satisfies the homogeneous Volterra equation

$$\rho(t) = -\int_t^{t_1} [H_x(t)^*]^+ \left[\sum_{j=1}^n \dot{x}_j^0(s)H_{xx_j}(s)^* + H_{xt}(s)^* \right] \rho(s) \, ds \qquad (t \in T_1).$$

By unique solvability it follows that $\rho(t) = 0$, $t \in T_1$. Then

$$0 = \psi^K(t_1 - r) = -H_x(t_1 - r)^* l_1$$

implies that $l_1 = 0$. This contradicts the nontriviality condition $(l_0, l_1) \neq (0, 0)$.

Remark 3.3. Consider the maximum principle for the case of a fixed final state. Then $l_0 = 0$ implies $\psi^K(t) = 0$ $(t \in T_1)$. This follows from regularity, since the maximum condition has the form

$$\psi^K(t)f(x_t^0, v^0(t), t) \leq \psi^K(t)y$$

for all $y \in \operatorname{co} f(x_t^0, \Omega(t), t)$.

Remark 3.4. If there exists $\varepsilon > 0$ such that

$$\dot{x}^0(t) \in \operatorname{int} \operatorname{co} f(x_t^0, \Omega(t), t) \qquad (t \in [t_1 - r - \varepsilon, t_1]),$$

then $l_0 \neq 0$.

Suppose $l_0 = 0$. Then as in Remark 3.3 it follows that $\psi^K(t) = 0$ $(t \in [t_1 - r - \varepsilon, t_1])$. This used in the adjoint equation (i) shows that $l_1 = 0$, contradicting the nontriviality condition. Observe that we do not require the existence of a *uniform* neighborhood of $\dot{x}^0(t)$ contained in $\operatorname{co} f(x_t^0, \Omega(t), t)$ as it is required for stable reachability in Colonius [15, Theorem 2.2].

The following theorem gives a second form of the maximum principle proposed by Bien/Chyung [10]. Recall that the dual space of $\mathbf{C}^n(T)$ can be identified with $\mathbf{NBV}^n(T)$, the space of normalized functions of bounded variation on T with values in $\mathbb{R}^n$ being right continuous in t_0 (cf. the definition (2.2) of η).

THEOREM 3.2 (BC-form of the maximum principle). *Let* x^0, v^0 *be an optimal solution of Problem* 1 *and assume that* x^0 *is regular.*

Then there exist Lagrange multipliers $(l_0, l_1, \rho) \in \mathbb{R}_+ \times \mathbb{R}^k \times L_\infty^k(T)$ *such that* $(l_0, l_1) \neq$ $(0, 0)$, $\rho \,|\, [t_0, t_1 - r] = 0$, *and the adjoint variable* $\psi^C \in \mathbf{NBV}^n(T)$ *defined by*

$$(\text{i}) \quad \psi^C(t) = -\int_t^{t_1} \eta(s, t)^* (\psi^C(s) + H_x(s)^* \rho(s))\, ds$$

$$-l_0 \int_t^{t_1} \frac{\partial}{\partial x} g(x^0(s), v^0(s), s)\, ds$$

$$-\int_t^{t_1} \Big[\sum_{j=1}^n f_j(x_s^0, v^0(s), s) H_{xx_j}(s)^* + H_{xt}(s)^* \Big] \rho(s)\, ds$$

$$-\begin{cases} H_x(t_1 - r)^* l_1 & t \in [t_0, t_1 - r], \\ 0, & t \in (t_1 - r, t_1], \end{cases}$$

is right continuous in t_0 *and satisfies*

$$(\text{ii}) \quad -l_0 g(x^0(t), v^0(t), t) + (\psi^C(t) - H_x(t)^* \rho(t)) f(x_t^0, v^0(t), t)$$

$$\geqq -l_0 g(x^0(t), \omega, t) + (\psi^C(t) - H_x(t)^* \rho(t)) f(x_t^0, \omega, t)$$

for all $\omega \in \Omega(t)$, *a.a.* $t \in T$.

Proof. Theorem 3.2 follows from Corollary 2.1(ii) in the same way as Theorem 3.1 follows from Corollary 2.1(i). Observe that $D_1 C(x^0, v^0) l_2$ can be identified with an element of $\mathbf{NBV}^n(T)$:

$$[D_1 C(x^0, v^0)^* l_2](t) = -\int_t^{t_1} \eta(t, s)^* H_x(s)^* \rho(s)\, ds$$

$$-\int_t^{t_1} \Big[\sum_{j=1}^n f_j(x_s^0, v^0(s), s) H_{xx_j}(s)^* + H_{xt}(s)^* \Big] \rho(s)\, ds.$$

y^C can be identified with $\psi^C \in \mathbf{NBV}^n(T)$. Furthermore,

$$y^C \circ F(x^0, v - v^0) = \int_T \int_{t_0}^t f(x_s^0, v(s) - v^0(s), s)\, ds\, d\psi^C(t)$$

$$= -\int_T f(x_t^0, v(t) - v^0(t), t) \psi^C(t)\, dt$$

and

$$l_2 \circ C(x^0, v) = \int_T \rho(t) H_x(t) f(x_t^0, v(t) - v^0(t), t)\, dt.$$

This yields the maximum condition (ii). $\square$

Remark 3.5. On subintervals of T, the adjoint equation and the maximum condition have a simpler form.

Since $\rho(t) = 0$ on $[t_0, t_1 - r]$, the maximum condition does not involve ρ on this interval. Furthermore,

$$\int_t^{t_1} \eta(s, t)^* H_x(s)^* \rho(s)\, ds = \int_{T_1} \eta(s, t)^* H_x(s)^* \rho(s)\, ds$$

for $t \in [t_0, t_1 - r]$ and for $t \in [t_0, t_1 - 2r]$ and $s \in [t_1 - r, t_1]$,

$$\eta(s, t) = \eta(s, s - r);$$

this follows from the definition (2.2) of η. In special cases, e.g., systems with a single

constant delay, the adjoint variable ψ^C is even absolutely continuous on T with the possible exception of a jump at $t_1 - r$.

Remark 3.6. Again the nontriviality condition can be written in terms of the adjoint variable

$$(0, 0) \neq (l_0, \psi^C \mid T_1) \in \mathbb{R}_+ \times \mathbf{NBV}^n(T_1).$$

Suppose that $(l_0, \psi^C \mid T_1)$ is trivial. Then $\rho(t) = 0$ $(t \in T_1)$, since regularity is assumed and the maximum condition on T_1 has the form

$$\rho(t)H_x(t)f(x_t^0, v^0(t), t) \leqq \rho(t)y$$

for all $y \in H_x(t)$ co $f(x_t^0, \Omega(t), t)$. Hence,

$$0 = \psi^C(t_1 - r) = -H_x(t_1 - r)^* l_1.$$

Since by regularity, rank $H_x(t_1 - r)^* = k$, it follows that $l_1 = 0$ contradicting the condition $(l_0, l_1) \neq (0, 0)$.

Remark 3.7. Compare the two forms of the maximum principle: The adjoint variables ψ^K and ψ^C are related by an integral equation which in its abstract form is given in Corollary 2.1(iii).

The BK-form appears natural because the end condition is originally a pure phase equality constraint. Hence, the Lagrange multiplier corresponding to the end condition should appear only in the adjoint equation, not in the maximum condition as is the case in the BC-form.

Remark 3.8. Bien/Chyung [10] require, instead of (0.6), that $h(x(t_1), t_1) = 0$. Then the adjoint variable has a corresponding jump in t_1, while (0.6) induces a jump of ψ^C in $t_1 - r$. Bien and Chyung redefine their adjoint variable in t_1 such that the Lagrange multiplier corresponding to the finite dimensional part of the end condition does not appear explicitly in the adjoint equation (see Bien/Chyung [10, Thm. 3.1(ii) and (iii)]). Apart from this minor variation, their adjoint equation coincides with the adjoint equation above specialized to the case of a single constant delay. However, the maximum condition given above has a global form on the whole interval T, not only on $[t_0, t_1 - r]$ as in Bien/Chyung [10].

Remark 3.9. Sufficiency of the maximum principle has been analyzed by Banks/Kent [3] and Bien/Chyung [10]. Restricting their analysis to the fixed final state problem, Banks and Kent establish sufficiency of the BK-form under the usual normality and convexity assumptions. Bien and Chyung show sufficiency of the BC-form under similar assumptions for general function space and conditions. Taking into account the equivalence of the two forms, their results extend those of Banks and Kent.

Remark 3.10. We have treated the phase equality constraint (0.4) by using its equivalent formulation as a mixed control phase variable constraint. Now it is clear from the proof of Theorem 3.2 that we can deal with *any* constraint of the form

$$b(x_t, v(t), t) = 0 \qquad (t \in T_1)$$

where $b: C^n[-r, 0] \times \mathbb{R}^m \times T_1 \to \mathbb{R}^k$ satisfies the same assumptions as f in (1.1)–(1.3). In particular, for $r = 0$ the results of Makowski/Neustadt [37] are extended to relaxed— instead of ordinary—optimal solutions of problems with ordinary differential equations. The maximum principle obtained in this way is equivalent to that in Schwarzkopf [43] (where $r = 0$, $T_1 = T$). An advantage of our approach is that we can make use of results in general optimization theory (cf. Theorem 2.1)). Furthermore,

a relation between regularity and structural properties of the system is obtained (Lemma 3.1).

Remark 3.11. The regularization procedure performed in the proof of Lemma 3.2 can easily be extended in order to deal with inequality constraints of the form

$$b(x_t, v(t), t) \leqq 0 \qquad (t \in T_1),$$

where b is as in the remark above.

Suppose that there exist Lagrange multipliers $(l_0, l_1, l_2) \in \mathbb{R}_+ \times \mathbb{R}^k \times (L_\infty^k(T_1))^*$ satisfying the analogue of (2.11) and additionally

$$\langle l_2, z \rangle \geqq 0 \quad \text{for all } z \in L_\infty^k(T_1) \quad \text{with } z \geqq 0.$$

Then we impose the following regularity condition: There exists a neighborhood V of $0 \in \mathbb{R}^k$ such that

$$V \cap \mathbb{R}_-^k \subset \text{co } b(x_t^0, \Omega(t), t) \cap \mathbb{R}_-^k \qquad (t \in T_1),$$

where $\mathbb{R}_-^k := \{x = (x_1, \cdots, x_k) \in \mathbb{R}^k : x_i \leqq 0\}$.

In order to prove that l_2 can be identified with an element of $L_\infty^k(T_1)$, consider first nonpositive simpler functions s. Here everything goes through due to the above regularity assumption. Now suppose s is a *general* simple function. Decompose s into its positive and negative parts:

$$s = s^+ - s^-, \quad \text{where } s^+, s^- \geqq 0.$$

Due to the positivity of l_2, we find

$$\langle l_2, s \rangle = \langle l_2, s^+ \rangle + \langle l_2, -s^- \rangle \geqq \langle l_2, -s^- \rangle$$

and

$$\langle l_2, -s \rangle \geqq \langle l_2, -s^+ \rangle.$$

Since $\langle l_2, -s^- \rangle$ and $\langle l_2, -s^+ \rangle$ converge to 0 for $\|s\|_{L_1} \to 0$, also $\langle l_2, s \rangle$ converges to 0. Hence, l_2 can be identified with an element of $L_\infty^k(T_1)$.

Remark 3.12. The maximum principle reduces the optimal control problem (0.1)–(0.5) to an "infinite defect boundary value problem" (as defined by Kamenskii/Myshkis [31]) consisting of a retarded and an advanced equation coupled by the maximum condition. Solution of such a system is very difficult (cf. also Grimm/Schmitt [21], Kamenskii [27], Hutson [24], Kamenskii/Kamenskii/Myshkis [30]).

However, the existence of Lagrange Multipliers is also important in order to prove the convergence of computational procedures (cf. Wierzbicki/Hatko [49], Wierzbicki/Kurcyusz [50], who use shifted penalty methods in order to compute solutions of problems with function space end condition, and Williamson/Polak [51]).

Remark 3.13. Conversely, one obtains results on infinite defect boundary value problems by considering a corresponding optimal control problem. Existence of an optimal solution and validity of the maximum principle imply that a certain boundary value problem has a solution (see Kamenskii [29]).

Remark 3.14. Colonius [16] gave an example of an optimal control problem where the assumptions of Theorem 2.1 are met while the optimal trajectory is not regular. It is shown that the maximum principle is *not* satisfied (for a certain performance index), i.e., there are no nontrivial Lagrange multipliers $(l_0, l) \in \mathbb{R}_+ \times W^{n,\infty}[-r, 0]$. This shows that the regularity assumption is crucial for the validity of the maximum principle.

4. Discussion. The proof of the maximum principle relies on two main assumptions: (i) the existence of Lagrange multipliers $(l_0, l_1) \in \mathbb{R}_+ \times (W^{k,\infty}(T_1))^*$ can be established if $\pi_\infty \mathscr{A}$ has a nonempty interior in $L_\infty^k(T_1)$, and (ii) l can be identified with an element of $W^{k,\infty}(T_1)$ if the optimal trajectory is regular. We shall discuss, for the fixed final state problem, how restrictive these assumptions are. For simplicity, we assume that $\Omega(t) = \Omega_0$ on T and that the function f defining the right-hand side of the system equation is independent of t and an element of the following Banach space $\mathscr{F}$:

$$\mathscr{F} := \left\{ f: C^n[-r, 0] \times \mathbb{R}^m \to \mathbb{R}^n : \|f\| \right.$$
$$\left. := \max \left\{ \sup_{\varphi \in C^n, \omega \in \mathbb{R}^m} |f(\varphi, \omega)|, \sup_{\varphi \in C^n, \omega \in \mathbb{R}^m} \|D_1 f(\varphi, \omega)\| \right\} < \infty \right\}.$$

We have for $f \in \mathscr{F}$:

$$\pi_\infty \mathscr{A} = \{z \in L_\infty^n[-r, 0]: \text{ there is } v \in \mathscr{S}^\# \text{ such that } z = (\dot{x})_{t_1} \text{ for } x \text{ satisfying } x_{t_0} = 0$$

$$\text{and } \dot{x}(t) = D_1 f(x_t^0, v^0(t))x_t + f(x_t^0, v(t) - v^0(t))(t \in T)\}.$$

The linearized system satisfies the assumptions in Colonius [15, Corollary 4.1]. Hence, int $\pi_\infty \mathscr{A} \neq \varnothing$ if and only if

(4.1)   int $\{z \in L_\infty^n(T_1): z(t) \in \text{co } f(x_t^0, \Omega_0)(t \in T_1)\} \neq \varnothing.$

By the same arguments as in [15, Remark 4.4] one can see that the set of elements $f \in \mathscr{F}$ satisfying condition (4.1) for all $x^0 \in C^n[t_0 - r, t_1]$ is open and dense in $\mathscr{F}$ provided that Ω_0 contains at least $n + 1$ points. In this sense the condition int $\pi_\infty \mathscr{A} \neq \varnothing$ is generically satisfied for hereditary differential systems defined by $f \in \mathscr{F}$.

Linear systems Λ (see Colonius [15, § 2]) are not included in the class of systems defined by $\mathscr{F}$. However, on the basis of [15, Remark 4.4], a similar genericity statement can easily be proven.

Observe, however, that the situation is quite different if we restrict ourselves to the class of functions f where $\omega \in \mathbb{R}^m$ appears *affinely*. Then the condition $n \leq m$ on the dimensions of the state space (=output space) and the control space is necessary for int $\pi_\infty \mathscr{A} \neq \varnothing$.

Now consider the second assumption concerning regularity. Colonius [15, Examples 4.1, 4.2] specifies classes of systems, where *all* trajectories reaching a certain final function are regular. In general the situation is much more complicated, and we have to look at the linearized system.

Colonius [15, Prop. 3.1] states that a trajectory x^0 is regular if and only if the zero trajectory of the corresponding linearized system is regular. Since $0 \in \mathscr{A}$, we have either that $0 \in \partial \mathscr{A}$ or $0 \in \text{int } \mathscr{A}$. The first situation is a degenerate one: if $0 \in \partial \mathscr{A}$ and int $\mathscr{A} \neq \varnothing$ there are Lagrange multipliers $(0, l) \in \mathbb{R}_+ \times (W^{n,\infty}[-r, 0])^*$ and the optimality condition (2.11) is independent of the performance index (see Remark 2.3 for the case where int $\mathscr{A} = \varnothing$).

Now suppose $0 \in \text{int } \mathscr{A}$. Then by [15, Corollaries 4.1, 4.2], 0 is regularly reachable, and by [15, Thm. 4.2], the set of regular trajectories is open and dense in the set of all trajectories reaching 0. Thus, irregularity, in particular, irregularity of the zero trajectory, is "exceptional" (compare, e.g., the discussion by Maurin [38, p. 29]).

In this sense, the maximum principle is generically valid and its use as a necessary optimality condition appears to be justified.

REFERENCES

[1] H. T. BANKS, *Control of functional differential equations with function space boundary conditions*, Delay and Functional Differential Equations and their Applications, K. Schmitt, ed., Academic Press, New York, 1972, pp. 1–16.
[2] H. T. BANKS AND M. Q. JACOBS, *An attainable sets approach to optimal control of functional differential equations with function space boundary conditions*, J. Differential Equations, 13 (1973), pp. 127–149.
[3] H. T. BANKS AND G. A. KENT, *Control of functional differential equations of retarded and neutral type with target sets in function space*, SIAM J. Control, 10 (1972), pp. 567–593.
[4] H. T. BANKS AND A. MANITIUS, *Application of abstract variational theory to hereditary systems—A survey*, IEEE Trans. Automat. Control, AC-19 (1974), pp. 524–533.
[5] V. BARBU, *Convex control problems for linear differential systems of retarded type*, Ricerche Math., XXVI (1977), pp. 3–26.
[6] V. BARBU AND TH. PRECUPANU, *Convexity and Optimization in Banach Spaces*, Editura Academiei, Bucuresti/Sijthoff and Noordhoff, Groningen, 1978.
[7] G. R. BATES, *Hereditary optimal control problems*, Ph.D. thesis, Purdue University, W. Lafayette, IN, 1977.
[8] L. D. BERKOVITZ, *A penalty function proof of the maximum principle*, Appl. Math. Optim., 2 (1975/76), pp. 291–303.
[9] Z. BIEN, *Optimal control of delay systems*, Ph.D. thesis, University of Iowa, Ames, Iowa, 1975.
[10] Z. BIEN AND D. H. CHYUNG, *Optimal control of delay systems with a final function condition*, Internat. J. Control, 32 (1980), pp. 539–560.
[11] N. BOURBAKI, *Intégration*, Livre V, Chap. V (2^{me}éd.), Hermann, Paris, 1968.
[12] H. H. BUEHLER, *Application of Neustadt's theory of extremals to an optimal control problem with a functional differential equation and a functional inequality constraint*, Appl. Math. Optim., 2 (1975/76), pp. 34–74.
[13] F. COLONIUS, *Necessary optimality conditions for nonlinear hereditary differential systems with function space end constraints*, Functional Differential Systems and Related Topics, M. Kisielewicz, ed., The Higher College of Engineering in Zielona Gora, Zielona Gora, Poland, 1980, pp. 62–71.
[14] ———, *Regularization of Lagrange multipliers for time delay systems with fixed final state*, in Optimization and Optimal Control, A. Auslender, W. Oettli and J. Stoer, eds., Springer-Verlag, Berlin–Heidelberg–New York, 1981, pp. 163–177.
[15] ———, *Stable and regular reachability for relaxed hereditary differential systems*, this Journal, this issue, pp. 675–694.
[16] ———, *A penalty function proof of a Lagrange multiplier theorem with application to linear delay systems*, Appl. Math. Optim., 7 (1981), pp. 309–334.
[17] F. COLONIUS AND D. HINRICHSEN, *Optimal control of functional differential systems*, this Journal, 16 (1978), pp. 861–879.
[18] ———, *Optimal control of hereditary differential systems*, in Recent Theoretical Developments in Control, M. J. Gregson, ed., Academic Press, London, 1978, pp. 215–240.
[19] P. C. DAS, *Application of Dubovitskii/Milyutin formalism to optimal settling problems with constraints*, Optimization and Optimal Control (Oberwolfach 1974), Springer-Verlag, Berlin–Heidelberg–New York, 1975.
[20] I. V. GIRSANOV, *Lectures on Mathematical Theory of Extremum Problems*, Springer-Verlag, Berlin–Heidelberg–New York, 1967.
[21] L. J. GRIMM AND K. SCHMITT, *Boundary value problems for differential equations with deviating arguments*, Aequationes Math., 4 (1970), pp. 176–190.
[22] J. HALE, *Theory of Functional Differential Equations*, second ed., Springer-Verlag, Berlin–Heidelberg–New York, 1977.
[23] E. HEWITT AND K. STROMBERG, *Real and Abstract Analysis*, Springer-Verlag, Berlin–Heidelberg–New York, 1969.
[24] V. HUTSON, *A note on a boundary value problem for linear differential difference equations of mixed type*, J. Math. Anal. Appl., 61 (1977), pp. 416–425.
[25] M. Q. JACOBS, *An optimization problem for an nth order scalar neutral functional differential equation with functional side conditions*, in Delay and Functional Differential Equations and Their Applications, K. Schmitt, ed., Academic Press, New York, 1972, pp. 345–372.
[26] M. Q. JACOBS AND T. J. KAO, *An optimum settling problem for time lag systems*, J. Math. Anal. Appl., 40 (1972), pp. 687–707.
[27] G. A. KAMENSKII, *Variational and boundary value problems with deviating argument*, Differential Equations, 6 (1970), pp. 1026–1032.

[28] ———, *On a conditional extremum of a functional with deviating argument*, Soviet Math. Dokl., 18 (1977), pp. 921–924.

[29] ———, *A variational method for solving boundary value problems for certain linear differential equations with deviating arguments*, Differential Equations, 13 (1977), pp. 820–825.

[30] A. G. KAMENSKII, G. A. KAMENSKII AND A. D. MYSHKIS, *On the convergence of the finite difference method of numerically solving boundary value problems for linear differential-difference equations*, Soviet Math. Dokl., 18 (1977), pp. 321–324.

[31] G. A. KAMENSKII AND A. D. MYSHKIS, *Boundary value problems with infinite defect*, Differential Equations, 7 (1971), pp. 1612–1618.

[32] G. A. KAMENSKII AND A. L. SKUBACHEVSKII, *Extremals of functionals with deviating argument*, Moskovskii Aviazionnii Institut, Moscow, 1979 (in Russian).

[33] G. A. KENT, *Optimal control of functional differential equations of neutral type*, Doctoral thesis, Brown Univ., Providence, RI, 1971.

[34] ———, *A maximum principle for optimal control problems with neutral functional differential systems*, Bull. Amer. Math. Soc., 77 (1971), pp. 565–570.

[35] S. KURCYUSZ, *A local maximum principle for operator constraints and its application to systems with time lag*, Control and Cyber., 2 (1973), pp. 99–125.

[36] S. KURCYUSZ AND A. W. OLBROT, *On the closure in $W^{1,q}$ of the attainable subspace of linear time lag systems*, J. Differential Equations, 24 (1977), pp. 29–50.

[37] K. MAKOWSKI AND L. W. NEUSTADT, *Optimal control problems with mixed control phase variable equality and inequality constraints*, SIAM J. Control, 12 (1974), pp. 184–228.

[38] K. MAURIN, *Analysis, Part II*, PWN-Polish Scientific Publishers, Warszawa, and D. Reidel Publishing Company, London, 1980.

[39] A. W. OLBROT, *Control of retarded systems with function space constraints: necessary optimality conditions*, Control and Cyber., 5 (1976), pp. 5–31.

[40] R. R. PHELPS, *Weak* support points of convex sets in E^**, Israel J. Math., 2 (1964), pp. 177–182.

[41] L. S. PONTRYAGIN, V. G. BOLTJANSKII, R. V. GAMKRELIDZE AND E. E. MISCHENKO, *The Mathematical Theory of Optimal Control Processes*, John Wiley, New York, 1962.

[42] A. B. SCHWARZKOPF, *Optimal controls for problems with a restricted state space*, SIAM J. Control, 10 (1972), pp. 487–511.

[43] ———, *Relaxed control problems with state equality constraints*, SIAM J. Control, 13 (1975), pp. 677–694.

[44] ———, *Optimal controls with equality state constraints*, J. Opt. Theory Appl., 19 (1976), pp. 455–468.

[45] J. UTTHOFF, *Optimale Kontrolle neutraler Funktional-differentialgleichungen*, Diplomarbeit, Freie Universität Berlin, Berlin, 1979.

[46] J. WARGA, *Minimizing variational curves restricted to a preassigned set*, Trans. Amer. Math. Soc., 12 (1964), pp. 449–455.

[47] ———, *Unilateral variational problems defined by integral equations*, Michigan Math. J., 12 (1965), pp. 449–480.

[48] ———, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, 1972.

[49] A. P. WIERZBICKI AND A. HATKO, *Computational methods in Hilbert space for optimal control problems with delays*, 5th IFIP Conference on optimization Techniques, R. Conti, A. Ruberti, eds., Springer-Verlag, Berlin–Heidelberg–New York, 1973.

[50] A. P. WIERZBICKI AND S. KURCYUSZ, *Projection on a cone, penalty functionals and duality theory for problems with inequality constraints in a Hilbert space*, this Journal, 15 (1977), pp. 25–56.

[51] L. J. WILLIAMSON AND E. POLAK, *Relaxed controls and the convergence of optimal control algorithms*, this Journal, 14 (1976), pp. 737–756.

[52] L. C. YOUNG, *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, 1969.

[53] J. ZOWE AND S. KURCYUSZ, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Opt., 5 (1979), pp. 49–62.

[54] A. M. ZVERKIN, G. A. KAMENSKII, S. B. NORKIN AND L. E. ELSGOLTS, *Differential equations with a perturbed argument*, Russian Math. Surveys, 17 (1962), pp. 61–146.