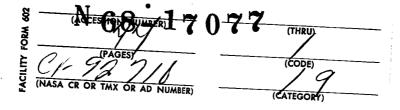VARIATIONAL PROBLEMS INVOLVING

FUNCTIONAL DIFFERENTIAL EQUATIONS

by

H. T. Banks*

Center for Dynamical Systems

Brown University

Providence, Rhode Island

VARIATIONAL PROBLEMS INVOLVING
FUNCTIONAL DIFFERENTIAL EQUATIONS

by

H. T. Banks

Introduction. In recent years an interest in variational problems or optimal control problems involving delayed systems has arisen. In particular, a number of papers have been written on problems involving systems with a time lag in the state variable. More general cases where the system has some type of functional dependency have also been investigated. Existence of optimal solutions for such problems has been discussed in several works [2], [6], [13]. The purpose of this paper is to obtain necessary conditions (in the form of an integrated maximum principle) for problems with quite general nonlinear functional differential systems. These systems will include as special cases many integro-differential systems and time lag (variable or constant) systems.

In this paper integrals will be understood to be Lebesgue or Lebesgue-Stieltjes integrals. Similarly, when speaking of a measurable function, we shall mean a Lebesgue measurable function unless it is specifically stated otherwise. By a solution of a (functional) differential equation will be meant an absolutely continuous (A. C.) function which satisfies the equation almost everywhere with respect to Lebesgue measure. Vector matrix notation will be employed throughout and we shall not distinguish between a vector and its transpose when it is clear what is meant. The notation $|A|$ will denote the Euclidean norm of A in whatever space A lies.

In §1, we shall formulate a general extremal problem for functional differential systems. A theorem concerning necessary conditions for extremals will be given. The proofs involve a generalization of the idea of quasiconvex families due to Gamkrelidze [7]. In §2 we shall apply this theorem for extremals to control problems, obtaining necessary conditions in terms of variations about the optimal solution. An example of a specific functional control problem will be discussed in §3. Finally, in §4, it will be shown that the necessary conditions in terms of variations can be used to derive a general maximum principle involving multipliers or adjoint variables.

§1. Extremal Theory. Let t_0 and α_0 be fixed in R^1 with $-\infty < \alpha_0 < t_0$. Let $I = [\alpha_0, a)$ be a bounded interval containing $[\alpha_0, t_0]$ and put $I' = I \cap \{t : t > t_0\}$. Let $\mathscr{G}$ be a fixed open convex region in R^n (possibly all of R^n). Denote by $\mathscr{C} = C(I, \mathscr{G})$ the space of bounded continuous n-vector functions on I into $\mathscr{G}$ with uniform topology. That is, $\mathscr{C}$ will be considered as a subspace of $C(I, R^n)$. For any set X contained in $\mathscr{G}$, define $AC(I, X)$ to be the subset of $\mathscr{C}$ consisting of all bounded absolutely continuous n-vector functions on I into X. For any non-negative $L_1(I)$ function $K(t)$, we then define

$$AC(I,X)_K = \{x \in AC(I,X) : |\dot{x}(t)| \leq K(t) \text{ a.e. on } I\}.$$

If k is any positive integer, we define

$$P^k = \{\alpha \in R^k : \alpha^i \geq 0 \quad \text{for} \quad i = 1, \ldots, k, \quad \text{and} \quad \sum_1^k \alpha^i = 1\}.$$

Denote by $\mathscr{F}$ a family of n-vector functionals $F(x(\cdot), t)$ where $F : \mathscr{C} \times I' \to R^n$. By the notation $F(x(\cdot), t)$, we shall mean that for each fixed t in I',

$$F(\cdot, t) : C([\alpha_0, t], \mathscr{G}) \to R^n,$$

so that $F(x(\cdot), t)$ may depend on any or all of the values $x(\tau)$, $\alpha_0 \leq \tau \leq t$. If, for each $t \in I'$, F has a Fréchet differential dF with respect to x, then $dF[x(\cdot), t; \cdot]$ is a bounded linear map from $C([\alpha_0, t], R^n)$ into R^n. We shall then write $\|dF[x(\cdot), t; \cdot]\| \leq m(t)$ to mean $|dF[x(\cdot), t; \psi]| \leq m(t) \|\psi\|_t$ for each $\psi \in C([\alpha_0, t], R^n)$ and $t \in I'$, where

$$\|\psi\|_t = \sup\{|\psi(s)| : s \in [\alpha_0, t]\}.$$

Note: In the discussions below, we shall not always distinguish between $\mathscr{C} = C(I, \mathscr{G})$ and $C([\alpha_0, t], \mathscr{G})$. For example, instead of saying that for each t, $F(x(\cdot), t)$ is C^1 w.r.t. x in $C([\alpha_0, t], \mathscr{G})$, we shall say simply that F is C^1 in x on $\mathscr{C}$. It will be clear what is meant. In fact, given any continuous x defined on $[\alpha_0, \tau]$ (contained in I) into $\mathscr{G}$, then x may be considered as an element of $\mathscr{C}$ by the convention $x(t) \equiv x(\tau)$ for $t \geq \tau$. Conversely, any x in $\mathscr{C}$ is also in $C([\alpha_0, t], \mathscr{G})$ for each $t \in I'$. We then make the following definition:

Definition: A family $\mathscr{F}$ is <u>absolutely</u> <u>quasiconvex</u> (A.Q.) if:

 1. Each $F(x(\cdot),t)$ in $\mathscr{F}$ is C^1 in x for fixed $t \in I'$ and measurable on I' for fixed $x \in \mathscr{C}$.

 2. Given any $F \in \mathscr{F}$ and any compact convex X contained in $\mathscr{G}$, there exists an $m \in L_1(I')$ (m depending on X,F) such that

$$|F(x(\cdot),t)| \leqq m(t)$$
$$\|dF[x(\cdot),t; \cdot]\| \leqq m(t)$$

for all $t \in I'$ and $x \in AC(I,X)$, where dF is the Fréchet differential of F w.r.t. x.

 3. For every compact convex X contained in $\mathscr{G}$, non-negative K in $L_1(I)$, finite collection $F_1,\ldots,F_k$ in $\mathscr{F}$, and $\varepsilon > 0$, there exists for each $\alpha \in P^k$ an F_α in $\mathscr{F}$ (depending on X,K, the F_i, and ε) satisfying

$$|F_\alpha(x(\cdot),t)| \leqq \sum_1^k m_i(t)$$
$$\|dF_\alpha[x(\cdot),t; \cdot]\| \leqq \sum_1^k m_i(t)$$

for each $\alpha \in P^k$, $t \in I'$, and $x \in AC(I,X)$, (where the m_i are the $L_1(I')$ functions described in 2. above depending on X and F_i), so that

$$G(x(\cdot),t,\alpha) \equiv \sum_1^k \alpha^i F_i(x(\cdot),t) - F_\alpha(x(\cdot),t)$$

satisfies:

3(i) $\qquad |G(x(\cdot),t,\alpha)| \leq 2\sum_1^k m_i(t)$

$$\|dG[x(\cdot),t,\alpha;\ \cdot]\| \leq 2\sum_1^k m_i(t)$$

for all $x \in AC(I,X)$, $\alpha \in P^k$, and $t \in I'$.

3(ii) $\qquad |\int_{\tau_1}^{\tau_2} G(x(\cdot),t,\alpha)dt| < \varepsilon$

for all $\alpha \in P^k$, $[\tau_1,\tau_2] \subset I'$, and $x \in AC(I,X)_K$.

3(iii) If $\{\alpha_i\}_{i=1}^\infty$ is a sequence in P^k such that $\alpha_i \to \bar{\alpha} \in P^k$, then $\{G(x(\cdot),t,\alpha_i)\}_1^\infty$ converges in measure on I' to $G(x(\cdot),t,\bar{\alpha})$ for each $x \in AC(I,X)_K$.

Next let Φ be the class of A. C. n-vector functions on $[\alpha_0,t_0]$ into $\mathscr{G}$. That is, $\Phi = \{\varphi : \varphi \in AC([\alpha_0,t_0],\mathscr{G})\}$. For $F \in \mathscr{F}$ and $\varphi \in \Phi$, we shall consider solutions to

(1.1)
$$\dot{x}(t) = F(x(\cdot),t) \qquad t > t_0$$
$$x(t) = \varphi(t) \qquad t \in [\alpha_0,t_0].$$

If $z(t)$, $\alpha_0 \leq t \leq \tau_1$, is a solution to (1.1) for $(F,\varphi) \in \mathscr{F} \times \Phi$, we define the $2n+1$ vector $q_z = (z(t_0),z(\tau_1),\tau_1)$. Let Q be the set of all such q_z for solutions to (1.1) for $(F,\varphi) \in \mathscr{F} \times \Phi$.

Let $\mathscr{N}$ be a given C^1 manifold in R^{2n+1} with boundary $\mathscr{M} = \partial\mathscr{N}$. For $q \in \mathscr{M}$, let $\mathscr{N}_T(q)$ be the tangent half-plane to $\mathscr{N}$ at q and let $\mathscr{M}_T(q)$ be the tangent plane to $\mathscr{M}$ at q.

<u>Definition</u>: A solution $z(t)$, $\alpha_0 \leq t \leq \tau$, to (1.1) corresponding to $(F,\varphi) \in \mathscr{F} \times \Phi$ is called an <u>$\mathscr{F},\mathscr{N},\Phi$ extremal</u> if

(i) $q_z \in \mathcal{M}$

(ii) there is a neighborhood V of q_z such that $V \cap Q \cap \mathcal{N} \subset \mathcal{M}$.

Let M be an aribtrary but fixed positive function in $L_1(\alpha_0, t_0)$. Define $\delta\Phi(M)$ to be the set of A.C. n-vector functions $\delta\varphi$ on $[\alpha_0, t_0]$ into R^n satisfying $|\dot{\delta\varphi}(t)| \leq M(t)$ a.e. on $[\alpha_0, t_0]$.

Given an $\mathcal{F}, \mathcal{N}, \Phi$ extremal $\hat{x}(t)$, $\alpha_0 \leq t \leq t_1$, corresponding to $(\hat{F}, \hat{\varphi})$ in $\mathcal{F} \times \Phi$, we shall denote by δF the elements in $[\mathcal{F}] - \hat{F}$ where $[\mathcal{F}]$ is the closed convex hull of $\mathcal{F}$. That is, $\delta F = \sum_1^k \alpha^i F_i - \hat{F}$, where $\alpha \in P^k$, k arbitrary. For δF in $[\mathcal{F}] - \hat{F}$ and $\delta\varphi \in \delta\Phi(M)$, let δx denote the solution to

$$\dot{\delta x}(t) = d\hat{F}[\hat{x}(\cdot), t; \delta x] + \delta F(\hat{x}(\cdot), t) \quad \text{on} \quad [t_0, t_1]$$

(1.2)

$$\delta x(t) = \delta\varphi(t) \quad \text{on} \quad [\alpha_0, t_0].$$

The existence of a solution of (1.2) is guaranteed by previous results of the author (see Theorems 1 and 2 in [1]).

With the above definitions in mind, we then define the set $\hat{\mathcal{K}}$ contained in R^{2n+1} by $\hat{\mathcal{K}} = \{(\delta\varphi(t_0), \delta x(t_1) + \delta t \hat{F}(\hat{x}(\cdot), t_1), \delta t) :$ $\delta\varphi \in \delta\Phi(M)$, $\delta t \in R^1$, δx is the solution to (1.2) for δF in $[\mathcal{F}] - \hat{F}$ and $\delta\varphi \in \delta\Phi(M)\}$.

Theorem 1: Suppose $\mathcal{F}$ is absolutely quasiconvex. Let $\hat{x}(t)$, $\alpha_0 \leq t \leq t_1$, be an $\mathcal{F}, \mathcal{N}, \Phi$ extremal corresponding to $(\hat{F}, \hat{\varphi}) \in$ $\mathcal{F} \times \Phi$. Suppose t_1 is a regular point for $\hat{F}(\hat{x}(\cdot), t)$. Then there exists a non-zero 2n+1 dimensional vector ζ such that

(i) ζ is orthogonal to $\mathscr{M}_T(q_{\hat{x}})$

(ii) $\zeta \cdot w \geqq 0$ for all $2n+1$ vectors w such that $w+q_{\hat{x}} \in \mathscr{N}_T(q_{\hat{x}})$

(iii) $\zeta \cdot p \leqq 0$ for all $p \in \hat{\mathscr{K}}$.

<u>Proof</u>: Let M and $\delta\Phi(M)$ be as defined previously. Let $\hat{x}(t)$, $\alpha_0 \leqq t \leqq t_1$, be the $\mathscr{F}, \mathscr{N}, \Phi$ extremal corresponding to $(\hat{F}, \hat{\varphi}) \in \mathscr{F} \times \Phi$. That is,

$$\dot{\hat{x}}(t) = \hat{F}(\hat{x}(\cdot), t) \qquad [t_0, t_1]$$

(1.3)

$$\hat{x}(t) = \hat{\varphi}(t) \qquad [\alpha_0, t_0] .$$

Let X be a fixed compact convex subset of $\mathscr{G}$ chosen so that each $\hat{x}(t)$, $\alpha_0 \leqq t \leqq t_1$, is an interior point of X. Let $\delta F = \sum_1^k \alpha^i F_i - \hat{F}$ represent an arbitrary element of $[\mathscr{F}] - \hat{F}$. Let $\hat{m}$, m_i, $i = 1, \ldots, k$, be the $L_1(I')$ functions in 2. of the definition of absolute quasi-convexity corresponding to $\hat{F}$, F_i, $i = 1, \ldots, k$, respectively and X. Define an $L_1(I)$ function K by

$$K(t) = \begin{cases} 3(\hat{m}(t) + \sum_1^k m_i(t)) & t \in I' \\ \\ |\dot{\hat{\varphi}}(t)| + M(t) & t \in [\alpha_0, t_0]. \end{cases}$$

Since $\mathscr{F}$ is A.Q., one can use the definition to show that given any $\mathcal{E}$, $0 \leqq \mathcal{E} \leqq 1$, there exists a function $G_{\mathcal{E}}(x(\cdot), t)$ defined on $\mathscr{G} \times I'$ into R^n such that:

(1.4) $$(\hat{F} + \mathcal{E}\delta F + G_{\mathcal{E}}) \in \mathscr{F}$$

$$|G_{\varepsilon}(x(\cdot),t)| \leq 2(\hat{m}(t) + \sum_1^k m_i(t))$$

(1.5)

$$\|dG_{\varepsilon}[x(\cdot),t; \cdot]\| \leq 2(\hat{m}(t) + \sum_1^k m_i(t))$$

for all $x \in AC(I,X)$ and $t \in I'$

(1.6)
$$|\int_{\tau_1}^{\tau_2} G_{\varepsilon}(x(\cdot),t)dt| < \varepsilon^2$$

for every τ_1, τ_2 in I' and $x \in AC(I,X)_K$.

If $z(t)$ is any solution to

$$\dot{z}(t) = \hat{F}(z(\cdot),t) + \varepsilon \delta F(z(\cdot),t) + G_{\varepsilon}(z(\cdot),t) \quad t > t_0$$

$$z(t) = \hat{\varphi}(t) + \varepsilon \delta\varphi(t) \quad \text{on} \quad [\alpha_0, t_0]$$

where $\delta\varphi \in \delta\Phi(M)$, then from the definition of K we get $|\dot{z}(t)| \leq K(t)$

so that the inequality in (1.6) holds for such solutions z over

intervals on which they exist. Note that if δF is not fixed, but

is allowed to range over $[\delta F_1, \ldots, \delta F_\nu]$, then K can be chosen in-

dependently of the particular δF in this set. For then there exists

$F_1, \ldots, F_k$ such that $\delta F_j = \sum_1^k a^{ji} F_i - \hat{F}$, $j = 1, \ldots, \nu$, so that K will

depend on $F_1, F_2, \ldots, F_k, \hat{F}$ but not a particular $\delta F \in [\delta F_1, \ldots, \delta F_\nu]$.

We next consider "perturbations" of the system (1.3). For

arbitrary $\delta t \in R^1$, $\delta\varphi \in \delta\Phi(M)$, $\delta F \in [\mathscr{F}] - \hat{F}$, and $0 \leq \varepsilon \leq 1$, we

consider the system

$$\dot{z}(t,\varepsilon) = \hat{F}(z(\cdot,\varepsilon),t) + \varepsilon \delta F(z(\cdot,\varepsilon),t) + G_{\varepsilon}(z(\cdot,\varepsilon),t) \quad t > t_0$$

(1.7)
$$z(t,\varepsilon) = \hat{\varphi}(t) + \varepsilon \delta\varphi(t) \quad \text{on} \quad [\alpha_0, t_0].$$

Lemmas similar to Lemmas 4.2 and 4.3 in [1] can be proved for this system (the proofs of Lemmas 4.2 and 4.3 are changed only slightly). One then has that for $\varepsilon > 0$ sufficiently small the solution $z(t,\varepsilon)$ to (1.7) exists on $[t_0, t_1 + \varepsilon|\delta t|]$ where it has the form

$$(1.8) \qquad z(t,\varepsilon) = \hat{x}(t) + \varepsilon \delta x(t) + o(\varepsilon)$$

with δx satisfying the linear variational system

$$(1.9) \qquad \begin{aligned} \dot{\delta x}(t) &= d\hat{F}[\hat{x}(\cdot),t;\delta x] + \delta F(\hat{x}(\cdot),t) \quad t \in [t_0, t_1 + \varepsilon|\delta t|] \\ \delta x(t) &= \delta \varphi(t) \quad t \in [\alpha_0, t_0]. \end{aligned}$$

Let $\hat{\mathscr{N}}_T = \mathscr{N}_T(q_{\hat{x}})$, $\hat{\mathscr{M}}_T = \mathscr{M}_T(q_{\hat{x}})$, and $\hat{\mathscr{N}}_T - q_{\hat{x}} = \{w \in R^{2n+1} : w = w^* - q_{\hat{x}},$ where $w^* \in \hat{\mathscr{N}}_T\}$. Then proceeding next exactly as in the proof of Theorem 5 in [1] (the proofs of Lemmas 4.4 and 4.5 in [1] are carried out with only slight modifications) one gets that the convex sets $\hat{\mathscr{K}}$ and $\hat{\mathscr{N}}_T - q_{\hat{x}}$ can be separated by a hyperplane $\mathscr{H}$ through the origin in R^{2n+1}. Choosing ζ to be a non-zero normal to $\mathscr{H}$ such that

$$(1.10) \qquad \zeta \cdot p \leqq 0 \leqq \zeta \cdot w$$

for all $p \in \hat{\mathscr{K}}$ and $w \in \hat{\mathscr{N}}_T - q_{\hat{x}}$, we have that (ii) and (iii) of Theorem 1 hold. It is easy to show that (i) also holds since $\hat{\mathscr{M}}_T - q_{\hat{x}}$ is a plane through the origin which is contained in $\hat{\mathscr{N}}_T - q_{\hat{x}}$.

§2. Applications to Control Problems. Let us denote the closure

of $AC(I,X)_K$ in the uniform topology by $cl\ AC(I,X)_K$. We then can

prove the following lemma.

Lemma 2.1: Let X be any compact convex subset of $\mathscr{G}$ where

$\mathscr{G} \subset R^n$. Let K be a non-negative $L_1(I)$ function, where I is

a finite interval. Then $\mathscr{Y} \equiv cl\ AC(I,X)_K$ is compact convex and

$\mathscr{Y} \subset AC(I,X)$.

Proof: Let $z \in \mathscr{Y}$. Then there is a sequence $\{y_k\}$ in $AC(I,X)_K$

such that y_k converges to z uniformly on I. For any τ_1, τ_2

in I we have

$$|y_k(\tau_1) - y_k(\tau_2)| = |\int_{\tau_1}^{\tau_2} \dot{y}_k(t)dt| \leq \int_{\tau_1}^{\tau_2} K(t)dt$$

for every k. Letting $k \to \infty$ we obtain

$$(2.1) \qquad |z(\tau_1) - z(\tau_2)| \leq \int_{\tau_1}^{\tau_2} K(t)dt \ ,$$

which shows that any z in $\mathscr{Y}$ is A.C. Furthermore, since X

is compact, any such z is a mapping on I into X. Thus,

$\mathscr{Y} \subset AC(I,X)$. From (2.1) we also can conclude that there is some

constant B such that $\|z\| \leq B$ for all z in $\mathscr{Y}$. That is, $\mathscr{Y}$

is a bounded subset of $C(I,R^n)$.

Given any $\mathcal{E} > 0$, let $E_1, \ldots, E_N$ be a partition of I

with $meas(E_i) < \delta$, where δ is that corresponding to $\mathcal{E}$ for the

A.C. of $\int K(t)dt$. Let $\{s_i\}_{i=1}^{N}$ be points chosen so that $s_i \in E_i$. Again using (2.1) we have that $z \in \mathscr{Y}$ implies

$$|z(s)-z(s_i)| \le \left|\int_{s_i}^{s} K(t)dt\right| < \varepsilon$$

for each $s \in E_i$. Thus we get

$$\sup_{z \in \mathscr{Y}} \sup_{s \in E_i} |z(s)-z(s_i)| \le \varepsilon.$$

It then follows that $\mathscr{Y}$ is a conditionally compact subset of $C(I,R^n)$. (See Theorem IV.6.5 in [5].) But $\mathscr{Y}$ is closed, hence compact. The convexity of $\mathscr{Y}$ is easily seen since $\mathscr{Y}$ is the closure of the convex set $AC(I,X)_K$.

We next state a lemma that will be needed to apply the previous theory to control problems. The proof of this lemma will not be given here since it is essentially the same as the proof due to Gamkrelidze of Lemma 4.1 in [7].

Lemma 2.2: Let X be a compact convex subset of $\mathscr{Y}$, K a non-negative $L_1(I)$ function, and $\varepsilon > 0$. Let $F_j(x(\cdot),t)$, $j = 1,\ldots,k$, be mappings from $AC(I,X) \times I'$ into R^n that are measurable in t for fixed x and C^1 in x for fixed t. Assume there exists an $m(t)$ in $L_1(I')$ such that $|F_j(x(\cdot),t)| \le m(t)$, $\|dF_j[x(\cdot),t;\cdot]\| \le m(t)$ for all $x \in AC(I,X)$ and $t \in I'$, $j = 1,\ldots,k$. Let $\mathscr{Y}$, a subset of $AC(I,X)$, be the compact convex set defined by $\mathscr{Y} = cl\ AC(I,X)_K$. Let $p_j(t)$, $j = 1,\ldots,k$, be

given non-negative real-valued measurable functions on I' satisfying $\sum_1^k p_j(t) = 1$ a.e. on I'. Then it is possible to subdivide I' into sufficiently small subintervals E_i, $i = \pm 1, \pm 2, \ldots$, and to assign to each E_i one of the functions $F_1, \ldots, F_k$, which we shall denote by F_{E_i}, so that the function $F(x(\cdot), t)$ defined by

$$F(x(\cdot), t) = F_{E_i}(x(\cdot), t) \qquad \text{for } t \in E_i,$$
$$i = \pm 1, \pm 2, \ldots$$
$$\text{and}$$
$$x \in AC(I, X)$$

satisfies

$$\left| \int_{\tau_1}^{\tau_2} \{ \sum_{j=1}^k p_j(t) F_j(x(\cdot), t) - F(x(\cdot), t) \} dt \right| < \varepsilon$$

for every τ_1, τ_2 in I' and $x \in \mathcal{Y}$.

We are now ready to consider the following optimal control problem:

Minimize $J[\bar{\varphi}, u, \bar{x}, t_1] = \int_{t_0}^{t_1} f^0(\bar{x}(\cdot), u(t), t) dt$ over $\bar{\Phi} \times \Omega \times R^{n-1} \times I'$ subject to

(i) $\dot{\bar{x}}(t) = \bar{f}(\bar{x}(\cdot), u(t), t)$ $t \in [t_0, t_1]$

      $\bar{x}(t) = \bar{\varphi}(t)$ $t \in [\alpha_0, t_0]$

(ii) $(\bar{x}(t_0), \bar{x}(t_1), t_1) \in \mathcal{T}$.

The following assumptions and definitions are made:

$\bar{\mathcal{Y}}$ is a fixed open convex region in R^{n-1}, $\bar{x}$ is an n-1 vector, $f = (f^0, \bar{f}) = (f^0, f^1, \ldots, f^{n-1})$ is an n-vector function

defined on $C(I,\bar{\mathscr{G}}) \times \mathscr{U} \times I'$, where $\mathscr{U} \subset R^r$. Each f^i is assumed C^1 in $\bar{x}$ and Borel measurable in u,t. $\bar{\Phi}$ and Ω are defined by $\bar{\Phi} = AC([\alpha_0, t_0], \bar{\mathscr{G}})$ and $\Omega = \{u : u$ is measurable on I' and $u(t) \in U(t)$ for $t \in I'\}$, where U is a given mapping of I' into subsets of $\mathscr{U}$. $\mathscr{T}$ is a given C^1 manifold in R^{2n-1} of dimension less than $2n-1$ satisfying $\mathscr{T} \subset \bar{\mathscr{G}} \times \bar{\mathscr{G}} \times I'$.

We also assume that given $\bar{X}$ compact, $\bar{X} \subset \bar{\mathscr{G}}$, and $u \in \Omega$, there exists an m in $L_1(I')$ such that

$$|f(\bar{x}(\cdot), u(t), t)| \leq m(t)$$
$$\|df[\bar{x}(\cdot), u(t), t; \cdot]\| \leq m(t)$$

for each $t \in I'$ and $\bar{x} \in AC(I, \bar{X})$, where df is the Fréchet differential of f w.r.t. $\bar{x}$.

Suppose that $(\bar{\phi}^*, u^*, \bar{x}^*, t_1^*)$ is a solution to the above problem. Put

$$x^*(t) \equiv \begin{pmatrix} x^0*(t) \\ \bar{x}^*(t) \end{pmatrix}, \quad \alpha_0 \leq t \leq t_1^*,$$

where

$$x^0*(t) \equiv \int_{t_0}^{t} f^0(\bar{x}^*(\cdot), u^*(\sigma), \sigma) d\sigma \quad t \in [t_0, t_1^*]$$

$$x^0*(t) \equiv 0 \quad t \in [\alpha_0, t_0].$$

Now define $f(x(\cdot), u, t) = f(\bar{x}(\cdot), u, t)$ where $x = (x^0, \bar{x}) \in C(I, \mathscr{G})$ with $\mathscr{G} \equiv R^1 \times \bar{\mathscr{G}}$. (That is, hereafter we shall write f as a function of x even though it does not really depend on x^0.)

Put $\mathscr{F} = \{F(x(\cdot),t) : F(x(\cdot),t) = f(x(\cdot),u(t),t)$ for $u \in \Omega\}$ and $\Phi = \{\varphi = (\varphi^0,\overline{\varphi}) : \overline{\varphi} \in \overline{\Phi}$ and $\varphi^0 \in AC([\alpha_0,t_0],R^1)\}$.

Let γ^0,ξ^0,τ represent scalars and γ,ξ represent n-1 vectors. Then define $\mathscr{N} \subset R^{2n+1}$ to be all $(\gamma^0,\gamma,\xi^0,\xi,\tau)$ with (γ,ξ,τ) near $(\overline{x}*(t_0),\overline{x}*(t_1^*),t_1^*)$ satisfying

$$(\gamma,\xi,\tau) \in \mathscr{F}$$
$$\gamma^0 = 0$$
$$\xi^0 \leq x^0*(t_1^*) \quad .$$

Define $\mathscr{M}$ to be the above set with the last inequality replaced by equality. Then $\mathscr{N}$ is a C^1 manifold with boundary $\mathscr{M}$.

With the above definitions in mind one can prove (the proof is exactly the same as the proof of Lemma 5.1 in [1]) that x* is an $\mathscr{F},\mathscr{N},\Phi$ extremal. Furthermore, the class $\mathscr{F}$ defined above is absolutely quasiconvex. The proof that $\mathscr{F}$ is A.Q. uses Lemma 2.2 and the arguments are similar to those for the proof of a similar result in [2].

Thus, we can apply Theorem 1 to the control problem in the above formulation. Let $\delta\Phi = \delta\Phi(M) = \{\delta\varphi \in AC([\alpha_0,t_0],R^n) : |\dot{\delta\varphi}(t)| \leq M(t)$ a. e. on $[\alpha_0,t_0]\}$ be as defined in §1. For $(\overline{\varphi}*,u*,\overline{x}*,t_1^*)$ optimal, and $\delta\varphi \in \delta\Phi$, α in P^k, $\{u_i\}_1^k$ in Ω, we denote by δx the solution to

$$\dot{\delta x}(t) = df[x*(\cdot),u*(t),t;\delta x] + \sum_1^k \alpha^i f(x*(\cdot),u_i(t),t) - f(x*(\cdot),u*(t),t)$$

(2.2) $\qquad\qquad\qquad\qquad\qquad\qquad$ on $[t_0,t_1^*]$

$$\delta x(t) = \delta\varphi(t) \qquad \text{on } [\alpha_0,t_0]$$

where $x* = (x^0*,\overline{x}*)$. Define $\delta\mathscr{X}$ to be the set of all such solu-tions for $\delta\varphi$ in $\delta\Phi$, $\{u_i\}_1^k$ any finite collection in Ω, and $\alpha \in P^k$, k arbitrary positive integer. Then we have:

Theorem 2: Let $(\overline{\varphi}*, u*, \overline{x}*, t_1^*)$ be optimal. Suppose t_1^* is a regular point for $f(\overline{x}*(\cdot), u*(t), t)$. Then there exists a non-zero $2n+1$ vector $\zeta = (b_0, b_1, a_1) = (b_0^0, b_0^1, \ldots, b_0^{n-1}, b_1^0, b_1^1, \ldots, b_1^{n-1}, a_1) = (b_0^0, \overline{b}_0, b_1^0, \overline{b}_1, a_1)$ such that

(1) The $2n-1$ vector $(\overline{b}_0, \overline{b}_1, a_1)$ is orthogonal to $\mathscr{T}$ at $(\overline{x}*(t_0), \overline{x}*(t_1^*), t_1^*)$

(2) $b_1^0 \leqq 0$

(3) $b_1 \cdot f(\overline{x}*(\cdot), u*(t_1^*), t_1^*) + a_1 = 0$

(4) $b_0 \cdot \delta\varphi(t_0) + b_1 \cdot \delta x(t_1^*) \leqq 0$ for arbitrary $\delta\varphi$ in $\delta\Phi$, δx in $\delta\mathscr{X}$ (δx corresponding to $\delta\varphi$).

Proof: Theorem 2 follows almost immediately from Theorem 1. State-ments (3) and (4) are a direct consequence of (iii) in Theorem 1. Statement (2) follows from (ii) of Theorem 1 since the $2n+1$ vec-tor $w = (0_n, -1, 0_{n-1}, 0)$ is such that $q_{\hat{x}} + w \in \mathscr{N}_T(q_{\hat{x}})$ for the above defined $\mathscr{N}$. Finally, if $(\gamma_T, \xi_T, \tau_T)$ is any tangent vector to $\mathscr{T}$ at $(\overline{x}*(t_0), \overline{x}*(t_1^*), t_1^*)$, then the vector $(0, \gamma_T, 0, \xi_T, \tau_T)$ is tangent to $\mathscr{M}$ (as defined above) at $q_{x*} = (0, \overline{x}*(t_0), x^0*(t_1^*), \overline{x}*(t_1^*), t_1^*)$. Hence condition (1) of Theorem 2 follows from (i) of Theorem 1.

Remarks: First let us note that the results of Theorem 2 imply that $b_1 \neq 0$. For if b_1 were zero, using (3) and (4) would give

$a_1 = 0$ and $b_0 = 0$ and hence $\zeta = (b_0, b_1, a_1) = 0$, a contradiction!

We also point out that if $\overline{\Phi}$ were some given convex class of A.C. initial functions for the control problem, one could still prove a result similar to Theorem 2. In the statement and proof of Theorem 1 then, Φ would be some specified convex class. The set $\delta\Phi(M)$ would be replaced by $[\Phi]-\hat{\varphi}$ and for $\delta\varphi$ in $[\delta\varphi_1, \ldots, \delta\varphi_\ell]$, $K(t)$ would be chosen so that $K(t) \geq |\dot{\hat{\varphi}}(t)| + \sum_1^\nu |\dot{\phi}_i(t)|$, where $\delta\varphi_j = \sum_{i=1}^\nu \alpha^{ji}\varphi_i - \hat{\varphi}$, $j = 1, \ldots, \ell$. In §2, Φ would be defined as before in terms of the given class $\overline{\Phi}$ for the control problem. Then we would have $\delta\Phi = [\Phi]-\varphi*$, where $\varphi* = (0, \overline{\varphi}*)$, in the statement of Theorem 2. It should be noted however that the result about $b_1 \neq 0$ mentioned above does not hold for this problem.

§3. <u>The Control Problem for a Specific Functional System.</u> Let $I, I', \overline{\Phi}, \mathcal{U}, \mathcal{G}, \mathcal{T}$, and Ω be as defined previously. In this section we shall consider a control problem where $f = (f^0, \overline{f})$ has the form

$$(3.1) \qquad f(\overline{x}(\cdot), u(t), t) = \int_{\alpha_0}^t a(s)g(\overline{x}(s), u(t), t)ds$$

where a is a scalar function. Using the usual arguments employed in measure theory and analysis, one can prove the following lemma without difficulty.

<u>Lemma 3.1:</u> Let $a \in L_1(I)$. Let $g = (g^0, \overline{g})$ be defined on

VARIATIONAL PROBLEMS INVOLVING
FUNCTIONAL DIFFERENTIAL EQUATIONS

by

H. T. Banks

Introduction. In recent years an interest in variational problems
or optimal control problems involving delayed systems has arisen.
In particular, a number of papers have been written on problems
involving systems with a time lag in the state variable. More gen-
eral cases where the system has some type of functional dependency
have also been investigated. Existence of optimal solutions for
such problems has been discussed in several works [2], [6], [13].
The purpose of this paper is to obtain necessary conditions (in
the form of an integrated maximum principle) for problems with
quite general nonlinear functional differential systems. These
systems will include as special cases many integro-differential sys-
tems and time lag (variable or constant) systems.

In this paper integrals will be understood to be Lebesgue
or Lebesgue-Stieltjes integrals. Similarly, when speaking of a
measurable function, we shall mean a Lebesgue measurable function
unless it is specifically stated otherwise. By a solution of a
(functional) differential equation will be meant an absolutely con-
tinuous (A. C.) function which satisfies the equation almost every-
where with respect to Lebesgue measure. Vector matrix notation will
be employed throughout and we shall not distinguish between a vec-
tor and its transpose when it is clear what is meant. The notation
$|A|$ will denote the Euclidean norm of A in whatever space A lies.

$\bar{\mathscr{G}} \times \mathscr{U} \times I'$, C^1 in $\bar{x}$ for fixed (u,t) in $\mathscr{U} \times I'$ and Borel measurable in (u,t) for fixed $\bar{x}$ in $\bar{\mathscr{G}}$. Given any compact $\bar{X}$ contained in $\bar{\mathscr{G}}$ and $u \in \Omega$, assume there exists a finite integrable function m on I' such that $|g(\bar{x},u(t),t)| < m(t)$, $|g_{\bar{x}}(\bar{x},u(t),t)| < m(t)$ for each $\bar{x} \in \bar{X}$ and $t \in I'$. Define f by

$$f(\bar{x}(\cdot),u(t),t) = \int_{\alpha_0}^{t} a(s) g(\bar{x}(s),u(t),t) ds$$

for $\bar{x} \in C(I,\bar{\mathscr{G}})$, $u \in \Omega$, and $t \in I'$. Then we have:

(i) f is measurable in t on I' for fixed $\bar{x} \in C(I,\bar{\mathscr{G}})$ and $u \in \Omega$

(ii) For fixed $t \in I'$ and $u \in \Omega$, f is C^1 in $\bar{x}$ on $C(I,\bar{\mathscr{G}})$. That is, for fixed t and u, the Fréchet differential of f w.r.t. $\bar{x}$ exists and is continuous in $\bar{x}$. The differential is given by

$$df[\bar{x}(\cdot),u(t),t;\bar{\psi}] = \int_{\alpha_0}^{t} a(s) g_{\bar{x}}(\bar{x}(s),u(t),t)\bar{\psi}(s) ds$$

(iii) Given $u \in \Omega$ and a compact $\bar{X}$ contained in $\bar{\mathscr{G}}$, there is an $\tilde{m}$ in $L_1(I')$ (namely, $\tilde{m}(t) = m(t)\int_{\alpha_0}^{t}|a(s)| ds$) such that

$$|f(\bar{x}(\cdot),u(t),t)| \leq \tilde{m}(t)$$

$$\|df[\bar{x}(\cdot),u(t),t; \cdot]\| \leq \tilde{m}(t)$$

for all $t \in I'$ and $\bar{x} \in AC(I,\bar{X})$.

As in the previous section, by x we shall mean the n-vector function $x = (x^0,\bar{x})$ where $x^0 \in C(I,R^1)$, $\bar{x} \in C(I,\bar{\mathscr{G}})$ (or $x \in C(I,\mathscr{G})$ with $\mathscr{G} \equiv R^1 \times \bar{\mathscr{G}}$). Then (3.1) may be written

$$f(x(\cdot),u(t),t) = \int_{\alpha_0}^{t} a(s)g(x(s),u(t),t)ds$$

and f has Fréchet differential w.r.t. x given by

$$df[x(\cdot),u(t),t;\psi] = \int_{\alpha_0}^{t} a(s)g_x(x(s),u(t),t)\psi(s)ds$$

where

$$g_x = \begin{pmatrix} 0 \\ \cdot \\ \cdot & & g_{\bar{x}} \\ \cdot \\ 0 \end{pmatrix}$$

and $\psi \in C(I,R^n)$. With this notation adopted, we state and prove the next lemmas.

Lemma 3.2: Let a and $g = (g^0,\bar{g})$ be as described in Lemma 3.1. Let $x \in C(I,\mathscr{G})$ and u measurable on I' into $\mathscr{U}$ be given. Fix t in I'. Then the n × n matrix system

$$\Lambda(t,t) = E$$

(3.2)

$$\frac{\partial\Lambda}{\partial\sigma}(\sigma,t) + \int_{\sigma}^{t}\Lambda(s,t)a(\sigma)g_x(x(\sigma),u(s),s)ds = 0 \quad t_0 \leq \sigma \leq t$$

has a unique solution on $[t_0,t]$.

Proof: For fixed t,x, and u, the above system can be replaced by the equivalent system

(3.3)
$$y(t) = E$$

$$y'(\sigma) + \int_{\sigma}^{t}y(s)a(\sigma)h(\sigma,s)ds = 0 \quad \sigma \in [t_0,t]$$

where $\quad' = \dfrac{d}{d\sigma}$. The function $h(\sigma,s)$ is continuous in σ for each s, measurable in s for each σ, and satisfies $|h(\sigma,s)| \le$ $m(s)$ for each σ, where m is some finite-valued $L_1(I')$ function.

Consider next the system

$$\dot{z}(\tau) - \int_0^\tau z(\eta)\tilde{a}(\tau)\tilde{h}(\tau,\eta)d\eta = 0 \qquad 0 \le \tau \le t-t_0$$

(3.4)

$$z(0) = E$$

where $\quad' = \dfrac{d}{d\tau}$ and $\tilde{a}(\tau) = a(t-\tau)$, $\tilde{h}(\tau,\eta) = h(t-\tau,t-\eta)$. Using the hypotheses on a and h, it is not difficult to show that system (3.4) satisfies the hypotheses of Theorems 1 and 2 in [1]. It follows that system (3.4) has a unique solution z on $[0,t-t_0]$. Then define

(3.5) $$y(\xi) = z(t-\xi) \quad t_0 \le \xi \le t.$$

The function y is A.C. on $[t_0,t]$ and satisfies

$$\frac{d}{d\tau}y(t-\tau) - \int_0^\tau y(t-\eta)\tilde{a}(\tau)\tilde{h}(\tau,\eta)d\eta = 0 \quad 0 \le \tau \le t-t_0$$

$$y(t) = E.$$

Making the change of variables $s = t-\eta$ in the above integral, one obtains

$$\frac{d}{d\tau}y(t-\tau) + \int_t^{t-\tau} y(s)\tilde{a}(\tau)\tilde{h}(\tau,t-s)ds = 0 \quad 0 \le \tau \le t-t_0$$

$$y(t) = E.$$

Letting $\sigma = t-\tau$ in this system gives

$$-y'(\sigma) + \int_t^\sigma y(s)\tilde{a}(t-\sigma)\tilde{h}(t-\sigma, t-s)ds = 0 \quad t_0 \leq \sigma \leq t$$

$$y(t) = E,$$

which is equivalent to (3.3).

Thus the function y defined by (3.5) is a solution to (3.3). By reversing the above procedures, one sees that any solution of (3.3) will transform into a solution of (3.4). Since (3.4) has a unique solution, we have that y defined in (3.5) is the unique solution of (3.3).

Lemma 3.3: Let a and $g = (g^0, \bar{g})$ be as described in Lemma 3.1. Let t_1 in I' be fixed. Assume that x in $C([\alpha_0, t_1], \mathscr{G})$ and u measurable on I' into $\mathscr{U}$ are given. For each t, $t_0 < t \leq t_1$, let $\Lambda(\sigma, t)$ be the matrix solution to

(3.6)
$$\Lambda(t,t) = E$$
$$\frac{\partial \Lambda}{\partial \sigma}(\sigma, t) + \int_\sigma^t \Lambda(s,t)a(\sigma)g_x(x(\sigma), u(s), s)ds = 0 \quad t_0 \leq \sigma \leq t.$$

Let $C(t)$ be an n-vector function in $L_1(t_0, t_1)$ and $\varphi \in AC([\alpha_0, t_0], R^n)$. Then the n-vector solution z to

(3.7)
$$\dot{z}(\sigma) = \int_{\alpha_0}^\sigma a(s)g_x(x(s), u(\sigma), \sigma)z(s)ds + C(\sigma) \quad \sigma \in [t_0, t_1]$$
$$z(\sigma) = \varphi(\sigma) \quad \sigma \in [\alpha_0, t_0]$$

exists on $[t_0, t_1]$ and for $t > t_0$ is given by

(3.8) $z(t) = \Lambda(t_0,t)\varphi(t_0) + \int_{t_0}^{t}\Lambda(s,t)C(s)ds$

$$+ \int_{t_0}^{t}\Lambda(s,t)\{\int_{\alpha_0}^{t_0}a(\sigma)g_x(x(\sigma),u(s),s)\varphi(\sigma)d\sigma\}ds.$$

Proof: Existence follows from Theorems 1 and 2 in [1] and the assumptions on a and g. If $\Lambda(\sigma,t)$ is the solution to (3.6), we multiply the first equation in (3.7) by Λ and integrate over σ from t_0 to t and obtain

$$\int_{t_0}^{t}\Lambda(\sigma,t)\dot{z}(\sigma)d\sigma = \int_{t_0}^{t}\Lambda(\sigma,t)d\sigma\int_{\alpha_0}^{\sigma}a(s)g_x(x(s),u(\sigma),\sigma)z(s)ds$$

$$+ \int_{t_0}^{t}\Lambda(\sigma,t)C(\sigma)d\sigma.$$

Integration by parts $(\Lambda(\sigma,t)$ is A.C. in $\sigma)$ gives

$$z(t) = \Lambda(t_0,t)\varphi(t_0) + \int_{t_0}^{t}\frac{\partial\Lambda}{\partial\sigma}(\sigma,t)z(\sigma)d\sigma$$

$$+ \int_{t_0}^{t}\Lambda(\sigma,t)C(\sigma)d\sigma$$

$$+ \int_{t_0}^{t}\Lambda(\sigma,t)d\sigma\int_{\alpha_0}^{\sigma}a(s)g_x(x(s),u(\sigma),\sigma)z(s)ds$$

$$= \Lambda(t_0,t)\varphi(t_0) + \int_{t_0}^{t}\Lambda(\sigma,t)C(\sigma)d\sigma$$

$$+ \int_{t_0}^{t}\Lambda(\sigma,t)d\sigma\int_{\alpha_0}^{t_0}a(s)g_x(x(s),u(\sigma),\sigma)\varphi(s)ds$$

$$+ \int_{t_0}^{t}\frac{\partial\Lambda}{\partial\sigma}(\sigma,t)z(\sigma)d\sigma + \int_{t_0}^{t}\Lambda(\sigma,t)d\sigma\int_{t_0}^{\sigma}a(s)g_x(x(s),u(\sigma),\sigma)z(s)ds.$$

Considering the last of these integrals (for t fixed) we have

$$\int_{t_0}^{t}(\int_{t_0}^{\sigma}\Lambda(\sigma,t)a(s)g_x(x(s),u(\sigma),\sigma)z(s)ds)d\sigma$$

$$= \int_{t_0}^{t}(\int_{t_0}^{t}h(\sigma,s)ds)d\sigma$$

where

$$h(\sigma,s) \equiv \begin{cases} \Lambda(\sigma,t)a(s)g_x(x(s),u(\sigma),\sigma)z(s) & \text{for } s \leqq \sigma \\[2ex] 0 & \text{for } s > \sigma. \end{cases}$$

It is not hard to see that h is absolutely integrable on $[t_0,t] \times [t_0,t]$. Hence, by Fubini's theorem, we may interchange the order of integration in the above integral. This gives

$$\int_{t_0}^t (\int_{t_0}^t h(\sigma,s)ds)d\sigma = \int_{t_0}^t (\int_{t_0}^t h(\sigma,s)d\sigma)ds$$

$$= \int_{t_0}^t (\int_s^t \Lambda(\sigma,t)a(s)g_x(x(s),u(\sigma),\sigma)z(s)d\sigma)ds$$

$$= \int_{t_0}^t (\int_\sigma^t \Lambda(s,t)a(\sigma)g_x(x(\sigma),u(s),s)z(\sigma)ds)d\sigma.$$

We thus obtain

$$z(t) = \Lambda(t_0,t)\varphi(t_0) + \int_{t_0}^t \Lambda(\sigma,t)C(\sigma)d\sigma$$

$$+ \int_{t_0}^t \Lambda(\sigma,t)d\sigma\int_{\alpha_0}^{t_0}a(s)g_x(x(s),u(\sigma),\sigma)\varphi(s)ds$$

$$+ \int_{t_0}^t \{\frac{\partial\Lambda}{\partial\sigma}(\sigma,t) + \int_\sigma^t \Lambda(s,t)a(\sigma)g_x(x(\sigma),u(s),s)ds\}z(\sigma)d\sigma.$$

But the last term vanishes since Λ satisfies (3.6).

With these preliminary lemmas in mind, we now consider the control problem with $f = (f^0,\bar{f})$ as defined by (3.1) where a and g satisfy the conditions in Lemma 3.1. The results of section 2 (in particular, Theorem 2) are valid for this problem. We shall show that these give a maximum principle in integral form for the control problem under consideration.

Using Lemma 3.3 with $C(t) = \sum_1^k \alpha^i f(x^*(\cdot), u_i(t), t) - f(x^*(\cdot), u^*(t), t)$, $\alpha \in P^k$, $u_i \in \Omega$, one finds that the elements of $\delta\mathfrak{X}$ have the form for $t > t_0$

$$(3.9) \qquad \delta x(t) = \Lambda(t_0, t)\delta\varphi(t_0)$$

$$+ \int_{t_0}^t \Lambda(s,t)\{\sum_1^k \alpha^i f(x^*(\cdot), u_i(s), s) - f(x^*(\cdot), u^*(s), s)\}ds$$

$$+ \int_{t_0}^t \Lambda(s,t)\{\int_{\alpha_0}^{t_0} a(\sigma)g_x(x^*(\sigma), u^*(s), s)\delta\varphi(\sigma)d\sigma\}ds$$

where $\delta\varphi \in \delta\Phi$ and Λ satisfies (3.6) with $x = x^*$ and $u = u^*$.

Define the n-vector function $\psi(\sigma) = (\psi^0(\sigma), \overline{\psi}(\sigma))$ by

$$(3.10) \qquad \psi(\sigma) = b_1 \Lambda(\sigma, t_1^*) \qquad t_0 \leq \sigma \leq t_1^*,$$

where b_1 is as in Theorem 2. Then ψ satisfies

$$\psi(t_1^*) = b_1$$

$$\dot{\psi}(\sigma) + \int_\sigma^{t_1^*} \psi(s)a(\sigma)g_x(x^*(\sigma), u^*(s), s)ds = 0 \qquad t_0 \leq \sigma \leq t_1^*.$$

Note that $\dot{\psi}^0(\sigma) = 0$ since g_x has zeros in its first column. Hence $\psi^0 = b_1^0 \leq 0$ by condition (2) of Theorem 2. Furthermore, not only is $\psi(t_1^*) = b_1$ non-zero, but the continuity of ψ implies $\psi(t)$ is non-zero on some interval $(\beta, t_1^*]$.

Using (3.9) and (3.10), statement (4) of Theorem 2 may be written

$$\{b_0 + \psi(t_0)\} \cdot \delta\varphi(t_0)$$

$$(3.11) \quad + \int_{t_0}^{t_1^*} \psi(s)\{\sum_1^k \alpha^i \dot{f}(x^*(\cdot),u_i(s),s) - f(x^*(\cdot),u^*(s),s)\}ds$$

$$+ \int_{t_0}^{t_1^*} \psi(s)\{\int_{\alpha_0}^{t_0} a(\sigma)g_x(x^*(\sigma),u^*(s),s)\delta\varphi(\sigma)d\sigma\}ds \leqq 0$$

for all $\delta\varphi \in \delta\Phi$, $\alpha = (\alpha^1,\ldots,\alpha^k)$ and $\{u_i\}_1^k$ in P^k and Ω, k arbitrary. Since $\delta\varphi$ and the u_i's are independent of each other, this gives

$$\int_{t_0}^{t_1^*} \psi(s)\int_{\alpha_0}^{s} a(\sigma)g(x^*(\sigma),u(s),s)d\sigma\ ds$$

$$(3.12)$$

$$\leqq \int_{t_0}^{t_1^*} \psi(s)\int_{\alpha_0}^{s} a(\sigma)g(x^*(\sigma),u^*(s),s)d\sigma\ ds$$

for all $u \in \Omega$ and

$$\{b_0 + \psi(t_0)\} \cdot \delta\varphi(t_0)$$

$$(3.13)$$

$$+ \int_{t_0}^{t_1^*} \psi(s)\int_{\alpha_0}^{t_0} a(\sigma)g_x(x^*(\sigma),u^*(s),s)\delta\varphi(\sigma)d\sigma\ ds \leqq 0$$

for arbitrary $\delta\varphi$ in $\delta\Phi$.

Condition (3.12) is a maximum principle in integral form. Condition (3.13) gives

$$(3.14) \quad \int_{t_0}^{t_1^*} \psi(s)ds\int_{\alpha_0}^{t_0} a(\sigma)g_x(x^*(\sigma),u^*(s),s)\delta\varphi(\sigma)d\sigma \leqq 0$$

for arbitrary $\delta\varphi$ in $\delta\Phi$ with $\delta\varphi(t_0) = 0$. Since the integrand in (3.14) is absolutely integrable, we can interchange the order of integration and obtain

$$(3.15) \quad \int_{\alpha_0}^{t_0} a(\sigma)\{\int_{t_0}^{t_1^*} \psi(s)g_x(x^*(\sigma),u^*(s),s)ds\}\delta\varphi(\sigma)d\sigma \leqq 0$$

for arbitrary $\delta\varphi$ in $\delta\Phi$ with $\delta\varphi(t_0) = 0$.

Since $a(\sigma)\int_{t_0}^{t_1^*}\psi(s)g_x(x^*(\sigma),u^*(s),s)ds$ is in $L_1(\alpha_0,t_0)$,

we may apply Lemma 2 of [1] to (3.15). This gives

$$a(\sigma)\int_{t_0}^{t_1^*}\psi(s)g_x(x^*(\sigma),u^*(s),s)ds = 0$$

for almost every σ in $[\alpha_0,t_0]$. Because of the form of g_x, this

may be written

(3.16)
$$\int_{t_0}^{t_1^*}\psi(s)a(\sigma)g_{\overline{x}}(\overline{x}^*(\sigma),u^*(s),s)ds = 0$$

for almost every σ in $[\alpha_0,t_0]$. Combining (3.16) and (3.13) one

gets

$$\{b_0 + \psi(t_0)\}\cdot\delta\varphi(t_0) \leq 0$$

for arbitrary $\delta\varphi$ in $\delta\Phi$. This gives

(3.17)
$$b_0 = -\psi(t_0) \ .$$

From (3) of Theorem 2 we have

(3.18) $a_1 = -\psi(t_1^*)\cdot\int_{\alpha_0}^{t_1^*}a(s)g(\overline{x}^*(s),u^*(t_1^*),t_1^*)ds.$

If we combine (3.10), (3.17), (3.18) and (1) of Theorem 2, we get

that the $2n-1$ vector

$$(-\overline{\psi}(t_0),\overline{\psi}(t_1^*),-\psi(t_1^*)\cdot\int_{\alpha_0}^{t_1^*}a(s)g(\overline{x}^*(s),u^*(t_1^*),t_1^*)ds)$$

is orthogonal to $\mathscr{S}$ at $(\overline{x}^*(t_0),\overline{x}^*(t_1^*),t_1^*)$.

We have thus proved the following theorem.

Theorem 3: Let $(\bar{\varphi}^*, u^*, \bar{x}^*, t_1^*)$ be optimal for the control problem with system equations given by (3.1). Assume that t_1^* is a regular point of $f(\bar{x}^*(\cdot), u^*(t), t)$. Then there exists a non-trivial A. C. n-vector function $\psi(t) = (\psi^0(t), \bar{\psi}(t))$ defined on $[t_0, t_1^*]$ satisfying:

(i) $\psi^0 = \text{constant} \leq 0$, $\psi(t_1^*) \neq 0$

$$\dot{\bar{\psi}}(t) + \int_t^{t_1^*} \psi(s) a(t) g_{\bar{x}}(\bar{x}^*(t), u^*(s), s) ds = 0 \quad t_0 \leq t \leq t_1^*.$$

(ii) $\int_{t_0}^{t_1^*} \psi(s) \int_{\alpha_0}^s a(\sigma) g(\bar{x}^*(\sigma), u(s), s) d\sigma \, ds$

$$\leq \int_{t_0}^{t_1^*} \psi(s) \int_{\alpha_0}^s a(\sigma) g(\bar{x}^*(\sigma), u^*(s), s) d\sigma \, ds$$

for all $u \in \Omega$.

(iii) $\int_{t_0}^{t_1^*} \psi(s) a(\sigma) g_{\bar{x}}(\bar{x}^*(\sigma), u^*(s), s) ds = 0$ for almost every σ in $[\alpha_0, t_0]$.

(iv) The $2n-1$ vector

$$(-\bar{\psi}(t_0), \bar{\psi}(t_1^*), -\psi(t_1^*) \cdot f^*(t_1^*))$$

is orthogonal to $\mathscr{T}$ at $(\bar{x}^*(t_0), \bar{x}^*(t_1^*), t_1^*)$.

§4. **Necessary Conditions for General Functional Systems.** We now consider the control problem with system equations involving a general functional $f = (f^0, \bar{f}) = f(\bar{x}(\cdot), u, t)$. We assume that f

has the general properties assumed in section 2. Let $\overline{x} \in C(I, \overline{\mathscr{G}})$ and $u \in \Omega$ be fixed. Since for each $t \in I'$ $df[\overline{x}(\cdot), u(t), t; \cdot]$ is a bounded linear operator on $C([\alpha_0, t], R^{n-1})$ into R^n, we have by the Riesz theorem that there exists an $n \times (n-1)$ matrix function $\overline{\eta}(t, s)$ such that

$$df[\overline{x}(\cdot), u(t), t; \overline{\psi}] = \int_{\alpha_0}^t d_s \overline{\eta}(t, s) \overline{\psi}(s)$$

for all $\overline{\psi} \in C([\alpha_0, t], R^{n-1})$ and $t \in I'$.

Defining $NBV[\alpha_0, t]$ as all g satisfying (i) g is of bounded variation on $[\alpha_0, t]$ (ii) $g(t) = 0$ (iii) g is continuous from the right, we may then assert the existence of a unique $\overline{\eta}(t, \cdot)$ in $NBV[\alpha_0, t]$ such that the above equation holds.

Let $\bigvee_{s=\alpha_0}^t \overline{\eta}(t, s)$ denote the variation of $\overline{\eta}(t, \cdot)$ on $[\alpha_0, t]$. Then a further consequence of the Riesz theorem is that there exists a constant $D > 0$ such that

$$\bigvee_{s=\alpha_0}^t \overline{\eta}(t, s) \leq D\|df[\overline{x}(\cdot), u(t), t; \cdot]\| .$$

Thus we have

$$\bigvee_{s=\alpha_0}^t \overline{\eta}(t, s) \leq Dm(t) = \widetilde{m}(t)$$

where $m \in L_1(I')$.

In this section we shall adopt the notation previously used in writing

$$f(x(\cdot),u,t) = f(\bar{x}(\cdot),u,t)$$

where $x = (x^0,\bar{x})$ with $x^0 \in C(I,R^1)$ and $\bar{x} \in C(I,\bar{\mathscr{G}})$, or $x \in C(I,\mathscr{G})$ with $\mathscr{G} \equiv R^1 \times \bar{\mathscr{G}}$. Then we may write

$$(4.1) \qquad df[x(\cdot),u(t),t;\psi] = \int_{\alpha_0}^t d_s\eta(t,s)\psi(s)$$

for $\psi \in C(I,R^n)$, where $\eta(t,s)$ is the $n \times n$ matrix function

$$\begin{pmatrix} 0 & & \\ \vdots & \bar{\eta}(t,s) & \\ 0 & & \end{pmatrix}.$$

The integral in (4.1) is the Lebesgue-Stieltjes integral. In the following discussions we shall assume that $\eta(t,s)$ has been extended as follows:

$$\eta(t,s) = \eta(t,t) = 0 \quad \text{for} \quad s > t$$
$$\eta(t,s) = \eta(t,\alpha_0) \qquad \text{for} \quad s < \alpha_0 .$$

This may be done without changing (4.1). Furthermore, in what follows we shall be considering $df[x(\cdot),u(t),t; \cdot]$ for x and u fixed unless otherwise indicated. Hence we shall write $df[t; \cdot]$ for $df[x(\cdot),u(t),t; \cdot]$ with the understanding that we are considering df for some fixed x and u.

Using (4.1) and the fact that $df[t;\psi]$ is measurable in t on I' for each $\psi \in C(I,R^n)$, it is not hard to prove that $\eta(t,s)$ is measurable in t for each fixed s. Furthermore, for

each σ in $[\alpha_0, t]$,

$$|\eta(t,\sigma)| = |\eta(t,t) - \eta(t,\sigma)|$$

$$\leq |\eta(t,t) - \eta(t,\sigma)| + |\eta(t,\sigma) - \eta(t,\alpha_0)|$$

$$\leq \bigvee_{s=\alpha_0}^{t} \eta(t,s) \leq \widetilde{m}(t).$$

That is, one can prove the following lemma.

Lemma 4.1: For each fixed s in I, the function $\eta(\cdot, s)$ is measurable on I'. Furthermore, $|\eta(t,\sigma)| \leq \widetilde{m}(t)$ uniformly in σ on $[\alpha_0, t]$ (and hence on I), where $\widetilde{m}$ is the $L_1(I')$ function satisfying $\bigvee_{s=\alpha_0}^{t} \eta(t,s) \leq \widetilde{m}(t)$.

We next prove a theorem concerning the existence of solutions to an integral equation. The multiplier (or adjoint) equation for our control problem will be a special case of this equation.

Theorem 4: Let the $n \times n$ matrix $N(\sigma, t)$ be measurable in σ on $\mathscr{I}'$ satisfying $|N(\sigma, t)| \leq r(\sigma)$ for every t in $[0, T]$, where $[0, T] \subset \mathscr{I}'$ and $r \in L_1(\mathscr{I}')$. Let $N(\sigma, t)$ be $BV[0, T]$ as a function of t, satisfying $\bigvee_{t=0}^{T} N(\sigma, t) \leq r(\sigma)$ for each σ in $\mathscr{I}'$. Let $F(z, t)$ be defined on $R^n \times [0, T]$ into R^n, continuous in z for each t, of bounded variation in t with $\bigvee_{t=0}^{T} F(z, t) \leq h(z)$, where h is a bounded measurable function on R^n. Furthermore, suppose there exists a bounded measurable $\gamma(t)$ on $[0, T]$ such that $|F(z, t)| \leq |z| \gamma(t)$. Let ξ be a constant

n-vector. Then the system

(4.2)
$$z(0) = \xi$$
$$z(t) + \int_0^t N(\sigma,t)F(z(\sigma),t)d\sigma = \xi \quad t \in (0,T]$$

has a solution that is in BV[0,T].

Proof: For $k = 1,2,\ldots,$ we define the sequence of functions

$$z_k(t) = \begin{cases} \xi & t \in [0,\ T/k] \\ \\ \xi - \int_0^{t-T/k} N(\sigma,t)F(z_k(\sigma),t)d\sigma & t \in (T/k,T]. \end{cases}$$

Then using the hypotheses on F and N, it is not difficult to show that $\{z_k\}$ is uniformly bounded on [0,T] and, in fact, $\bigvee_{t=0}^T z_k(t)$ is uniformly bounded. It follows from a well-known theorem of Helly that $\{z_k\}$ has a convergent subsequence, which we again call $\{z_k\}$, such that $z_k(t) \to z(t)$ for every t in [0,T] and the limit function z is in BV[0,T].

We show that z is a solution to (4.2). Clearly $z(0) = \xi$. Hence fix t in (0,T]. Then for k sufficiently large, we have t in (T/k,T] and thus

(4.3) $z_k(t) = \xi - \int_0^{t-T/k} N(\sigma,t)F(z_k(\sigma),t)d\sigma.$

The integral may be written

$$\int_0^{t-T/k} N(\sigma,t) F(z_k(\sigma),t) d\sigma$$

$$= \int_0^t N(\sigma,t) F(z_k(\sigma),t) d\sigma - \int_{t-T/k}^t N(\sigma,t) F(z_k(\sigma),t) d\sigma.$$

The second of these integrals approaches zero as $k \to \infty$ since

$$\left| \int_{t-T/k}^t N(\sigma,t) F(z_k(\sigma),t) d\sigma \right| \leq \int_{t-T/k}^t r(\sigma) Q \gamma(t) d\sigma$$

where Q is a bound for $\{z_k\}$. From the continuity of F in z and the fact that $|N(\sigma,t) F(z_k(\sigma),t)| \leq r(\sigma) Q \gamma(t)$, we have

$$\int_0^t N(\sigma,t) F(z_k(\sigma),t) d\sigma \to \int_0^t N(\sigma,t) F(z(\sigma),t) d\sigma$$

as $k \to \infty$, by dominated convergence. Thus, letting $k \to \infty$ in (4.3) gives

$$z(t) = \xi - \int_0^t N(\sigma,t) F(z(\sigma),t) d\sigma.$$

Corollary 4.1: In addition to the assumptions in Theorem 4, assume there exists a bounded measurable function $p(t)$ on $[0,T]$ such that $|F(z_1,t) - F(z_2,t)| \leq |z_1 - z_2| p(t)$ for all z_1, z_2 in R^n. Then (4.2) has a unique solution on $[0,T]$.

This corollary follows from Theorem 4 and the use of standard arguments from the theory of ordinary differential equations.

Corollary 4.2: Let $\eta(t,s)$ be the $n \times n$ matrix (see (4.1)) corresponding to df as discussed earlier, so that η has the properties given in Lemma 4.1. Then (for $t \in I'$) the system

$$Y(t,t) = E$$

(4.4)

$$Y(\sigma,t) + \int_\sigma^t Y(\beta,t)\eta(\beta,\sigma)d\beta = E \quad t_0 \leqq \sigma < t$$

has a unique solution $Y(\sigma,t)$ on $[t_0,t]$ which is of bounded variation in σ. Furthermore, $Y(\sigma,t)$ is continuous in σ at $\sigma = t$.

Proof: It is sufficient to consider a vector system on a fixed interval $[t_0,t_1]$, say

$$y(t_1) = e_1 = (1,0,\ldots,0)$$

$$y(\sigma) + \int_\sigma^{t_1} y(\beta)\eta(\beta,\sigma)d\beta = e_1 \quad t_0 \leqq \sigma < t_1$$

or, in terms of column vectors,

$$y(t_1) = e_1$$

(4.5)

$$y(\sigma) + \int_\sigma^{t_1} \eta^T(\beta,\sigma)y(\beta)d\beta = e_1 \quad t_0 \leqq t < t_1.$$

Define $N(\sigma,t) = \eta^T(t_1-\sigma,t_1-t)$ and $F(z,t) = z$. By a simple transformation we shall see that system (4.5) is equivalent to (4.2) with N and F as defined above and $\xi = e_1$.

Let z be the solution of (4.2) on $[0,t_1-t_0]$ for this N,F, and $\xi=e_1$. Define $y(s) = z(t_1-s)$ for s in $[t_0,t_1]$. Then $y(t_1) = z(0) = \xi = e_1$. Also, for t in $(0,t_1-t_0]$ we have

$$z(t) + \int_0^t N(\sigma,t)z(\sigma)d\sigma = e_1.$$

Let $s = t_1 - t$ in this equation. Then

$$z(t_1-s) + \int_0^{t_1-s} N(\sigma, t_1-s) z(\sigma) d\sigma = e_1.$$

Make the substitution $\beta = t_1 - \sigma$ in the integral above. We obtain

$$z(t_1-s) - \int_{t_1}^{s} N(t_1-\beta, t_1-s) z(t_1-\beta) d\beta = e_1,$$

or, since $y(s) = z(t_1-s)$ and $N(t_1-\beta, t_1-s) = \eta^T(\beta, s)$,

$$y(s) + \int_s^{t_1} \eta^T(\beta, s) y(\beta) d\beta = e_1$$

for s in $[t_0, t_1)$.

Since the above transformations are reversible, we have that $y(\sigma) = z(t_1-\sigma)$ is the unique solution to (4.5).

To show that $Y(\sigma, t)$ is continuous at $\sigma = t$, we note that $Y(\sigma, t)$ is bounded in σ on $[t_0, t]$, say $|Y(\sigma, t)| \leq M$. Then

$$|Y(t,t) - Y(\sigma, t)| = \left| \int_\sigma^t Y(\beta, t) \eta(\beta, \sigma) d\beta \right|$$

$$\leq \int_\sigma^t M |\eta(\beta, \sigma)| d\beta$$

$$\leq \int_\sigma^t M \tilde{m}(\beta) d\beta.$$

It follows that $|Y(t,t) - Y(\sigma, t)| \to 0$ as $\sigma \to t^-$.

Note that in addition to continuity at $\sigma = t$, one also has that $Y(\sigma, t)$ is continuous from the right in σ on $[t_0, t]$

since $\eta(\beta,\cdot)$ is right continuous for each β.

We are now able to prove a theorem concerning the representation of solutions in terms of adjoint or multiplier variables. That is, we shall give a type of variation of constants formula for functional differential equations.

Theorem 5: Let η be as given in (4.1) (x and u fixed). Let the n-vector functions φ and C be given with $\varphi \in AC([\alpha_0,t_0],R^n)$ and $C \in L_1(t_0,t_1)$. For each t in $(t_0,t_1]$, let $\Gamma(\sigma,t)$ be the matrix solution to

$$\Gamma(t,t) = E$$

(4.6)

$$\Gamma(\sigma,t) + \int_\sigma^t \Gamma(\beta,t)\eta(\beta,\sigma)d\beta = E \quad t_0 \leq \sigma < t.$$

Then the n-vector solution z to

(4.7)
$$\dot{z}(\sigma) = \int_{\alpha_0}^\sigma d_s\eta(\sigma,s)z(s) + C(\sigma) \quad t_0 \leq \sigma \leq t_1$$

$$z(\sigma) = \varphi(\sigma) \quad \alpha_0 \leq \sigma \leq t_0$$

is given for $t > t_0$ by

$$z(t) = \Gamma(t_0,t)\varphi(t_0) + \int_{t_0}^t \Gamma(\beta,t)d\beta \int_{\alpha_0}^{t_0} d_s\eta(\beta,s)\varphi(s)$$

(4.8)
$$+ \int_{t_0}^t \Gamma(\beta,t)C(\beta)d\beta.$$

Proof: As has already been pointed out, the existence of a solution to (4.7) is guaranteed by Theorems 1 and 2 in [1]. For

$t > t_0$ we have

$$z(t) = \Gamma(t_0,t)z(t_0) + \int_{t_0}^{t} d_\beta \{\Gamma(\beta,t)z(\beta)\}$$

$$= \Gamma(t_0,t)\varphi(t_0) + \int_{t_0}^{t} d_\beta \Gamma(\beta,t)z(\beta) + \int_{t_0}^{t} \Gamma(\beta,t)\dot{z}(\beta)d\beta$$

$$= \Gamma(t_0,t)\varphi(t_0) + \int_{t_0}^{t} d_\beta \Gamma(\beta,t)z(\beta)$$

$$+ \int_{t_0}^{t} \Gamma(\beta,t)\{\int_{\alpha_0}^{\beta} d_s\eta(\beta,s)z(s) + C(\beta)\}d\beta$$

$$= \Gamma(t_0,t)\varphi(t_0) + \int_{t_0}^{t} \Gamma(\beta,t)d\beta\int_{\alpha_0}^{t_0} d_s\eta(\beta,s)\varphi(s)$$

$$+ \int_{t_0}^{t} \Gamma(\beta,t)C(\beta)d\beta + \int_{t_0}^{t} d_\beta \Gamma(\beta,t)z(\beta)$$

$$+ \int_{t_0}^{t} \Gamma(\beta,t)d\beta\int_{t_0}^{\beta} d_s\eta(\beta,s)z(s).$$

Consider $\int_{t_0}^{t} \Gamma(\beta,t)d\beta\int_{t_0}^{\beta} d_s\eta(\beta,s)z(s)$. Since $\eta(\beta,s) = 0$ for $s \geq \beta$, this integral may be written $\int_{t_0}^{t} \Gamma(\beta,t)d\beta\int_{t_0}^{t} d_s\eta(\beta,s)z(s)$. Under the conditions which Γ,η,z satisfy, it is possible to use a slight modification of an unsymmetric Fubini type theorem of Cameron and Martin [3] to interchange the order of integration in this integral. We obtain

$$\int_{t_0}^{t} \{d_s\int_{t_0}^{t} \Gamma(\beta,t)\eta(\beta,s)d\beta\}z(s)$$

$$= \int_{t_0}^{t} \{d_s\int_{s}^{t}\Gamma(\beta,t)\eta(\beta,s)d\beta\}z(s)$$

since $\eta(\beta,s) = 0$ for $\beta \leq s$. We then have

$$z(t) = \Gamma(t_0, t)\varphi(t_0) + \int_{t_0}^t \Gamma(\beta, t)d\beta \int_{\alpha_0}^{t_0} d_s \eta(\beta, s)\varphi(s)$$

$$+ \int_{t_0}^t \Gamma(\beta, t)c(\beta)d\beta$$

$$+ \int_{t_0}^t d_s \{\Gamma(s, t) + \int_s^t \Gamma(\beta, t)\eta(\beta, s)d\beta\}z(s).$$

But the last integral vanishes since Γ satisfies (4.6). This gives the desired representation.

Let us return now to the general control problem formulated in §2. For $(\bar{\varphi}*, u*, \bar{x}*, t_1^*)$ optimal, we shall apply the results of the present section (with x and u fixed, $x = x* = (x^0*, \bar{x}*)$ and $u = u*$) to obtain necessary conditions from Theorem 2.

In (4.1), let $\eta = \eta*$ correspond to $x*, u*$. That is,

$$df[x*(\cdot), u*(t), t; \psi] = \int_{\alpha_0}^t d_s \eta*(t, s)\psi(s)$$

where

$$\eta*(t, s) = \begin{pmatrix} 0 \\ \vdots \\ \vdots & \bar{\eta}*(t, s) \\ \vdots \\ 0 \end{pmatrix}.$$

Using the notation of §2 and Theorem 5, we get that the elements of $\delta\mathscr{X}$ have the form for $t > t_0$

(4.9) $\delta x(t) = \Gamma(t_0, t)\delta\varphi(t_0)$

$$+ \int_{t_0}^t \Gamma(\beta, t)d\beta \int_{\alpha_0}^{t_0} d_s \eta*(\beta, s)\delta\varphi(s)$$

$$+ \int_{t_0}^t \Gamma(\beta, t)\{\sum_1^k \alpha^i f(x*(\cdot), u_i(\beta), \beta) - f(x*(\cdot), u*(\beta), \beta)\}d\beta$$

where $\delta\varphi \in \delta\Phi$, $\alpha \in P^k$, $u_i \in \Omega$, and Γ satisfies (4.6) with $\eta = \eta^*$.

As in §3, we define multipliers λ by

$$(4.10) \qquad \lambda(\sigma) = b_1 \Gamma(\sigma, t_1^*) \quad t_0 \le \sigma \le t_1^*$$

where $b_1 \ne 0$ is as described in Theorem 2. Then, since Γ satisfies (4.6) with $\eta = \eta^*$, we have that λ satisfies

$$\lambda(t_1^*) = b_1$$

$$\lambda(\sigma) + \int_\sigma^{t_1^*} \lambda(\beta) \eta^*(\beta,\sigma) d\beta = b_1 \quad t_0 \le \sigma < t_1^*.$$

Furthermore, λ is continuous at t_1^* by Corollary 4.2. Hence, there is an interval $(\beta, t_1^*]$ on which λ does not vanish. We also have that

$$\lambda^0(\sigma) = \lambda^0(t_1^*) = b_1^0 \le 0$$

since the first column in η^* is zero.

Using (4.9), (4.10), and (4) of Theorem 2 gives

$$\{b_0 + \lambda(t_0)\} \cdot \delta\varphi(t_0)$$

$$+ \int_{t_0}^{t_1^*} \lambda(\beta) d\beta \int_{\alpha_0}^{t_0} d_s \eta^*(\beta,s) \delta\varphi(s)$$

$$+ \int_{t_0}^{t_1^*} \lambda(\beta) \{\sum_1^k \alpha^i f(x^*(\cdot), u_i(\beta), \beta) - f(x^*(\cdot), u^*(\beta), \beta)\} d\beta$$

$$\le 0$$

for arbitrary $\delta\varphi \in \delta\Phi$, $\alpha \in P^k$, $\{u_i\}_1^k$ in Ω, k arbitrary. Since

$\delta\varphi$ and α, $\{u_i\}$ are independent, this may be written

(4.11) $\qquad \int_{t_0}^{t_1^*} \lambda(\beta)\{f(x^*(\cdot),u(\beta),\beta)-f(x^*(\cdot),u^*(\beta),\beta)\}d\beta \leqq 0$

for arbitrary $u \in \Omega$, and

$$\{b_0 + \lambda(t_0)\}\cdot\delta\varphi(t_0)$$

(4.12)

$$+ \int_{t_0}^{t_1^*}\lambda(\beta)d\beta\int_{\alpha_0}^{t_0}d_s\eta^*(\beta,s)\delta\varphi(s) \leqq 0$$

for arbitrary $\delta\varphi \in \delta\Phi$.

Since $-\delta\varphi$ is in $\delta\Phi$ whenever $\delta\varphi$ is, (4.12) may be written

$$\{b_0 + \lambda(t_0)\}\cdot\delta\varphi(t_0)$$

(4.13)

$$+ \int_{t_0}^{t_1^*}\lambda(\beta)d\beta\int_{\alpha_0}^{t_0}d_s\eta^*(\beta,s)\delta\varphi(s) = 0$$

for arbitrary $\delta\varphi \in \delta\Phi$.

Interchanging the order of integration (which again is possible by an unsymmetric Fubini theorem), we can write (4.13) as

$$\{b_0 + \lambda(t_0)\}\cdot\delta\varphi(t_0)$$

(4.14)

$$+ \int_{\alpha_0}^{t_0}\{d_s\int_{t_0}^{t_1^*}\lambda(\beta)\eta^*(\beta,s)d\beta\}\delta\varphi(s) = 0$$

for arbitrary $\delta\varphi \in \delta\Phi$.

Defining the n-vector function H by

(4.15) $\qquad H(s) = \int_{t_0}^{t_1^*} \lambda(\beta)\eta^*(\beta,s)d\beta \quad \alpha_0 \leqq s \leqq t_0$

and taking $\delta\varphi$ with j^{th} component equal 1, all other components zero in (4.14) yields

$$b_0^j + \lambda^j(t_0) + H^j(t_0) - H^j(\alpha_0) = 0.$$

Hence

(4.16) $\qquad b_0 + \lambda(t_0) + H(t_0) - H(\alpha_0) = 0$

or

(4.17) $\qquad b_0 = -\lambda(t_0) + \int_{t_0}^{t_1^*} \lambda(\beta)\{\eta^*(\beta,\alpha_0)-\eta^*(\beta,t_0)\}d\beta.$

Combining the results of the above discussions with Theorem 2 gives:

Theorem 6: Let $(\overline{\varphi}^*,u^*,\overline{x}^*,t_1^*)$ be optimal for the control problem with the general functional system equations. Suppose that t_1^* is a regular point of $f(\overline{x}^*(\cdot),u^*(t),t)$. Then there exists a non-trivial n-vector function $\lambda(t) = (\lambda^0(t),\overline{\lambda}(t))$ of bounded variation on $[t_0,t_1^*]$, continuous at t_1^*, satisfying:

(i) $\lambda^0(t) = $ constant $\leqq 0, \quad \lambda(t_1^*) \neq 0$

$$\overline{\lambda}(t) + \int_t^{t_1^*} \lambda(\beta)\overline{\eta}^*(\beta,t)d\beta = \overline{\lambda}(t_1^*) \quad t_0 \leqq t < t_1^*.$$

(ii) $\int_{t_0}^{t_1^*} \lambda(t)f(x^*(\cdot),u^*(t),t)dt \geqq \int_{t_0}^{t_1^*} \lambda(t)f(x^*(\cdot),u(t),t)dt$

for all $u \in \Omega$.

(iii) The 2n-1 vector

$$(-\overline{\lambda}(t_0) + \int_{t_0}^{t_1^*}\lambda(\beta)\{\overline{\eta}^*(\beta,\alpha_0)-\overline{\eta}^*(\beta,t_0)\}d\beta, \ \overline{\lambda}(t_1^*), -\lambda(t_1^*)\cdot f^*(t_1^*))$$

is orthogonal to $\mathscr{T}$ at $(\overline{x}^*(t_0),\overline{x}^*(t_1^*),t_1^*)$.

Let us make a few observations about this theorem. Return now to equations (4.13) and (4.14). We consider two cases:

Case 1: Suppose the matrix function $\eta^*(t,s)$ is such that $d_s\eta^*(t,s) = v^*(t,s)ds$ on $[\alpha_0,t_0]$. That is, $\eta^*(t,s)$ is A.C. in s. Then one can use (4.13) to show (using Lemma 2 of [1] and arguments similar to those in [1])

$$b_0 + \lambda(t_0) = 0$$

and

$$\int_{t_0}^{t_1^*}\lambda(\beta)v^*(\beta,s)d\beta = 0 \quad \text{a.e. } s \text{ in } [\alpha_0,t_0].$$

Then (iii) of Theorem 6 would give $(-\overline{\lambda}(t_0),\overline{\lambda}(t_1^*),-\lambda(t_1^*)\cdot f^*(t_1^*))$ orthogonal to $\mathscr{T}$.

Case 1 was exactly the situation we had in §3 of this paper. There we had

$$d_s\eta^*(t,s) = a(s)g_x(x^*(s),u^*(t),t)ds$$

or

$$\eta^*(t,s) = -\int_s^t a(\sigma)g_x(x^*(\sigma),u^*(t),t)d\sigma$$

Case 2: Suppose the function H defined by (4.15) is such that $dH(s) = h(s)ds$ on $[\alpha_0, t_0]$. That is, H is A.C. Then (4.14) may be written

$$\{b_0 + \lambda(t_0)\} \cdot \delta\varphi(t_0)$$

$$+ \int_{\alpha_0}^{t_0} h(s) \delta\varphi(s) ds = 0$$

for arbitrary $\delta\varphi \in \delta\Phi$. One can then show that this implies $h(s) = 0$ a.e. on $[\alpha_0, t_0]$, and, hence

$$b_0 + \lambda(t_0) = H(\alpha_0) - H(t_0) = 0.$$

Again (iii) of Theorem 6 would take the form stated in Case 1 above.

Case 2 includes the case when the functional is a functional involving only lags (see the discussion below).

Let us point out that it is not difficult to show Case 1 implies Case 2. However, the converse is not true as will be seen in the case of lag problems.

We mention briefly a few special types of functionals f that are included in the above formulation. First, let us consider the case where the functional dependence is in terms of lags. This type of system has been considered in detail in [1]. We just remark here that the results of Theorem 6 agree with those previous results. For simplicity let f depend on a single constant lag. That is, take $f = \tilde{f}(x(t), x(t-\theta), u(t), t)$ where $\tilde{f} = \tilde{f}(x, y, u, t)$ is a mapping of $R^n \times R^n \times \mathscr{U} \times I'$ into R^n. Then

$\eta^*(t,s)$ has the form

$$\eta^*(t,s) = \begin{cases} 0 & s \geq t \\ -\widetilde{f}^*_x(t) & t-\theta \leq s < t \\ -\widetilde{f}^*_x(t)-\widetilde{f}^*_y(t) & s < t-\theta, \end{cases}$$

where $\widetilde{f}^*(t) = \widetilde{f}(x^*(t),x^*(t-\theta),u^*(t),t)$. Note that in this case $\eta^*(t,s)$ is not A.C. in s. However (see 4.15)

$$H(s) = -\int_{s+\theta}^{t^*_1}\lambda(\beta)\widetilde{f}^*_y(\beta)d\beta - \int_{t_0}^{t^*_1}\lambda(\beta)\widetilde{f}^*_x(\beta)d\beta,$$

so that Case 2 above holds while Case 1 does not.

For this type of problem, the multipliers are usually given as A.C. functions satisfying a set of advanced differential-difference equations. These advanced equations are just the differentiated form of the equations for λ given in (i) of Theorem 6. In the case of lags, one can show that the multipliers λ of Theorem 6 are actually A.C. and satisfy the equations in (i) in differentiated form. Thus the results of Theorem 6 agree with the known results whenever we deal with a system with lags.

Remarks similar to those above hold for the special system discussed in §3 of this paper, so that Theorem 3 is just a very special case of Theorem 6.

Finally, for systems given by

$$\dot{x}(t) = f(x_t,u(t),t)$$

where x_t denotes the values $x(t+\sigma)$, $-\tau \leq \sigma \leq 0$, we have that $\eta*(t,s)$ satisfies

$$\eta*(t,s) = \eta*(t,t-\tau) \quad \text{for} \quad s \leq t-\tau.$$

Then (4.1) becomes

$$df[x_t^*, u*(t), t; \psi] = \int_{t-\tau}^{t} d_s \eta*(t,s) \psi(s).$$

These systems are included in the general case above and hence Theorem 6 holds, where $\eta*$ has the additional property noted above.

A final observation concerning a pointwise maximum principle for the general case should be made. In the case that $f(x(\cdot), u, t)$ is continuous in all arguments and the mapping $U(t)$ is such that $U(t) = \mathfrak{U}$ for $t \in I'$, where $\mathfrak{U}$ is a fixed subset of R^r, then one can show that (ii) of Theorem 6 implies a Pontryagin type maximum principle. That is,

$$\lambda(t) \cdot f(x*(\cdot), u*(t), t) = \sup\{\lambda(t) \cdot f(x*(\cdot), u, t) : u \in \mathfrak{U}\}$$

holds almost everywhere on $[t_0, t_1^*]$.

References

[1] H. T. Banks, "Necessary Conditions for Control Problems with Variable Time Lags", J. SIAM Control 6 (1968)

[2] H. T. Banks, "Optimal Control Problems with Delays", Doctoral thesis, Division of Mathematical Sciences, Purdue University, Lafayette, Indiana, 1967.

[3] R. H. Cameron and W. T. Martin, "An Unsymmetric Fubini Theorem", Bull. Amer. Math. Soc. 47 (1941), 121-125.

[4] D. H. Chyung and E. B. Lee, "Linear Optimal Systems With Time Delay", J. SIAM Control 4 (1966), 548-575.

[5] N. Dunford and J. T. Schwartz, "Linear Operators", Vol. I, Interscience, New York, 1958.

[6] Avner Friedman, "Optimal Control for Hereditary Processes", Arch. Rational Mech. Anal. 15 (1964), 396-416.

[7] R. V Gamkrelidze, "On Some Extremal Problems in the Theory of Differential Equations with Applications to the Theory of Optimal Control", J. SIAM Control 3 (1965), 106-128.

[8] L. M. Graves, "The Theory of Functions of Real Variables", 2nd ed., McGraw-Hill, New York, 1956.

[9] A. Halanay, "Differential Equations: Stability, Oscillations, Time Lags", Academic Press, New York, 1966.

[10] T. H. Hildebrandt, "On Integrals Related to and Extensions of the Lebesgue Integral", Bull. Amer. Soc. 24 (1917), 177-202.

[11] S. Lang, "Introduction to Differentiable Manifolds", Interscience, New York, 1962.

[12] I. P. Natanson, "Theory of Functions of a Real Variable", Ungar, New York, 1955.

[13] M. N. Oguztoreli, "Time-Lag Control Systems", Academic Press, New York, 1966.

[14] V. Volterra, "Theory of Functionals", Blackie and Son Limited, London, 1930.