THE OPTIMIZATION OF TRAJECTORIES

OF

LINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS

by

H. T. Banks and Marc Q. Jacobs
Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island

This research was supported in part by the National Aeronautics and Space Administration under Grant No. NGL-40-002-015 and in part by the Air Force Office of Scientific Research under Grant No. AF-AFOSR 67-0693A.

ABSTRACT

Our aim in this paper is to examine a number of fundamental questions in the theory of optimal control of processes monitored by certain general systems of linear functional differential equations with finite memories. In our model the controls may appear in a very general nonlinear functional manner which permits us to consider retardations of a rather general character in the control variables. In particular, we prove a maximal principle for such systems. We consider existence questions in the class of admissible Borel measurable (resp. piecewise continuous, almost piecewise continuous) initial functions and controls. We also show that certain solutions of an uncontrolled linear functional differential equation are piecewise analytic or quasi piecewise analytic.

THE OPTIMIZATION OF TRAJECTORIES

OF

LINEAR FUNCTIONAL DIFFERENTIAL EQUATIONS

by

H. T. Banks and Marc Q. Jacobs

§1. Introduction. The linear functional differential equation describing the controlled systems studied in this paper is given in equation (2.3) below. Many authors have studied control systems with delays in the state variables and there are several extensive bibliographies available in these areas [B-5, E-1, O-2, H-1, Z-1]. Recently, models for systems with delays in the control parameters have been proposed and some results for these systems have been obtained [B-3, B-4, D-1, H-3, H-11, K-1, K-2, L-1, L-2]. Such models occur naturally in the study of gas-pressurized bipropellant rocket systems [D-1], in population models [B-3, L-5], and in some complex economic models currently under study.

In section 2, we set down the notation, definitions, and standing hypotheses that will be required throughout. In section 3, we prove (see Theorem 3.1) that the collection of points in R^n, which can be attained at time t from admissible Borel measurable initial functions and controls, is compact and depends continuously (with respect to the Hausdorff metric) on t. The assumptions required for this theorem are in effect no more than is usually required just to prove the existence of solutions to the linear functional differential equation (2.3) (see [B-1, B-4]). Since the Lebesgue-Stieltjes measures, which will appear below in the variation of parameters formula (2.7), can be atomic, we cannot conclude that the above mentioned fixed-time cross sections of the attainable set are convex. However, if we add rather mild assumptions (Properties (Δ_1) and (Δ_2) in section 3), then we do obtain the convexity of the fixed-time cross sections of the attainable set (see Theorem 3.2). We then adapt an argument of Lee and Markus [L-3] for ordinary control problems to obtain the statements of the maximal principle in section 4 (Theorem 4.1 and Remark 4.1). Theorems 3.1 and 3.2 can be regarded as extensions of some well known results by Neustadt [N-2] and Olech [O-1]. Several very special cases of these two theorems have appeared in the literature [C-5, L-1, L-2, O-2]. The actual statement of the maximal principal is confined to the time optimal control problem, although this is not an essential feature (cf. the remarks preceding Lemma 4.1). This maximal principle complements recent work of Banks [B-3] and Kharatishvili [K-1, K-2], and in effect contains some of Lee's work [L-1, L-2] as special cases, although Lee has considered a somewhat different class of cost functionals. Also our work in essence includes the necessary conditions determined by Halanay in [H-2]. Even in the cases where our work overlaps with that of the above authors, our methods of proof differ in that we have made extensive use of a number of fairly recent developments in the theory of measurable multifunctions [A-2, C-2, C-3, H-10, J-1, J-2, K-3, O-1] to greatly simplify the arguments.

In section 5 we turn to a study of analyticity properties of "fundamental matrix" solutions to certain systems of functional differential equations. Many authors (see the references in [B-5, E-1, Z-1]) have studied various aspects of the analyticity of the solutions of very special types of functional differential equations, although none of these results appear to include those presented in section 5. Theorem 5.1 is a rather straightforward application of known results in ordinary differential equations. However, Theorem 5.2 which is extremely believable, seems to require a proof involving a substantially more intricate form of analysis than is needed to prove its simple counterpart in the theory of ordinary differential equations. It should be noted that the conclusion of Theorem 5.1 guarantees a type of piecewise analyticity of the "fundamental matrix", whereas Theorem 5.2 gives only what we have termed quasi piecewise analyticity. One might expect that if the coefficient matrices in system (5.2) are analytic, and if one starts with an analytic initial function, then the solution of the functional differential equation will also be analytic. Indeed, several authors have attempted to prove such results (for example, see [O-2], [P-1]), but very simple examples reveal that such general theorems are not true (see Remark 5.1).

Finally, in section 6 we apply the aforementioned piecewise analyticity (resp. quasi piecewise analyticity) properties to show that under certain circumstances the admissible initial functions and the admissible controls may be delimited to an appropriate class of piecewise continuous (resp. almost piecewise continuous) functions

functions and the attainable set will be the same as if one were using Borel measurable admissible initial functions and controls. These results are simply analogs of those obtained by Halkin [H-5] for ordinary linear control problems using the work on subintegrals by Halkin and Hendricks [H-6]. Halkin's paper extends earlier work in [G-1, H-4, L-4].

§2. Notation, Definitions, and General Hypotheses. If X and Y are nonempty sets, then a multifunction $\Omega\colon Y \to Y$ is simply a subset of $X \times Y$ with domain equal to X; equivalently Ω is a mapping of X into the nonempty subsets of Y. If Y is a topological space and $\Omega(x)$ is compact for each $x \in X$, then we say $\Omega\colon X \to Y$ is a compact multifunction. If $\mathcal{A}$ is a σ-algebra of subsets of X and if Y is a topological space, then we say a multifunction $\Omega\colon X \to Y$ is $\mathcal{A}$-measurable if

$$\Omega^{-}F \equiv \{x \in X \mid \Omega(x) \cap F \neq \emptyset\}$$

belongs to $\mathcal{A}$ for each closed $F \subset Y$. If X is a topological space and $\mathcal{A}$ is the collection of Borel sets in X, then we shall write Borel measurable instead of $\mathcal{A}$-measurable. If (Y,d) is a metric space, then diam $(\Omega(x))$, $x \in X$ denotes the diameter of $\Omega(x)$, i.e.,

$$\operatorname{diam} \Omega(x) = \sup \{d(y_1,y_2) \mid y_1,y_2 \in \Omega(x)\}.$$

The real vector space of all real $p \times q$ matrices will be denoted by $\mathcal{L}_{pq}$ for any pair of positive integers p and q. It is assumed that a definite norm, $|\cdot|$, is given on any of the finite dimensional vector spaces which come into our discussion. Let $[a,b]$ be a compact interval in R, and let $H\colon I \to \mathcal{L}_{pq}$ be a function of bounded variation. We shall use μ_H to denote the Lebesgue-Stieltjes measure on $[a,b]$ determined by H (see [D-3, pg. 358 ff.]). In constructing such measures from H, H <u>will always be taken to be left continuous on</u> (a,b). We observe that if $t \to T_H(t)$, $t \in [a,b]$ denotes the scalar function defined by

$$T_H(t) = \operatorname{Var}_{s\in[a,t]} H(s), \qquad t \in [a,b],$$

and if $|\mu_H|$ denotes the variation of the Lebesgue-Stieltjes measure μ_H, then one has [D-3, pg. 362],

$$(2.1) \qquad |\mu_H| = \mu_{T_H}.$$

For conciseness we frequently use $|H|(t)$ for $T_H(t)$ (this should not be confused with $|H(t)|$ which is the norm of the matrix $H(t)$). If $g\colon [a,b] \to R^p$ is μ_H-integrable, then $\int_a^b g(t)dH(t)$ denotes the integral of g over $[a,b]$ with respect to the measure μ_H. We use $\mathcal{L}_1([a,b], \mu_H, R^p)$ to denote the collection of all μ_H-integrable functions $g\colon [a,b] \to R^p$.

If $\Omega\colon [a,b] \to R^p$ is a multifunction, then $\int^b \Omega(t)dH(t)$ is used to denote the set (possibly empty)

$$\{\int_a^b g(t)dH(t) \mid g \in \mathcal{L}_1([a,b], \mu_H, R^p), \quad g(t) \in \Omega(t), \ a \leq t \leq b\}$$

(cf. [A-2, C-2, C-3, D-2, H-7, O-1, J-3]).

We shall deal frequently with mappings $f\colon X \times Y \to Z$ where X,Y,Z are sets. It will be convenient to use $f(x,\cdot)$, where x is a fixed element of X, to denote the mapping $y \to f(x,y)$, $y \in Y$. The mapping $f(\cdot,y)\colon X \to Z$, for y a fixed element of Y is similarly defined.

Throughout the paper we make the following standing hypotheses: 1°) F and G are two Lebesgue measurable mappings from $R \times R$ into $\mathcal{L}_{nn}$, 2°) $F(t,s) = 0$ for $s \geq 0$, 3°) $F(t,s) = F(t,-\tau)$ for $s \leq -\tau$ where τ is a given positive constant, 4°) $G(t,s) = 0$ for $s \geq t$, 5°) $G(t,s) = G(t,-\tau)$ for $s \leq -\tau$, 6°) for each fixed $t \in R$ the functions $G(t,\cdot)$ and $F(t,\cdot)$ are of bounded variation on R, and in addition 7°) there is a Lebesgue measurable function $\beta\colon R \to R$ which is Lebesgue summable on every finite interval and which satisfies

$$(2.2) \qquad |G(t,s)|,\ |F(t,s)| \leq \beta(t), \quad t,s \in R,$$

$$\mathrm{Var}_{s\in R} F(t,s) = \mathrm{Var}_{s\in[-\tau,0]} F(t,s) \leq \beta(t), \quad t \in R,$$

$$\mathrm{Var}_{s\in R} G(t,s) = \mathrm{Var}_{s\in[-\tau,t]} G(t,s) \leq \beta(t), \quad t \in R.$$

Let $h\colon R^m \times R \to R^n$ be a given function such that for each $t \in R$ the function $u \to h(u,t)$, $u \in R^m$, is continuous, and for each $u \in R^m$ the function $t \to h(u,t)$, $t \in R$, is Borel measurable. We shall consider control systems which can be described by systems of real functional differential equations (FDE's) of the form

$$(2.3) \qquad \dot{x}(t) = \int_{-\tau}^{0} x(t+s)d_sF(t,s) + \int_{-\tau}^{\tau} h(u(s),s)d_sG(t,s),$$

where both integrals in (2.3) are understood in the Lebesgue-Stieltjes sense with the symbol d_s being used to emphasize that the measures are constructed from the functions $F(t,\cdot)$ and $G(t,\cdot)$.

Let $U: [-\tau,\infty) \to R^m$ and $\Phi: [-\tau,0] \to R^n$ be given Borel measurable, compact multifunctions. It will be assumed that there is a positive constant M such that

$$(2.4) \qquad \begin{aligned} &\operatorname{diam} U(t),\ \operatorname{diam} h(U(t),t) \leqq M, \quad t \geqq -\tau \\ &\operatorname{diam} \Phi(t) \leqq M, \quad -\tau \leqq t \leqq 0. \end{aligned}$$

A triple $\{\varphi,u,t_1\}$ is called <u>admissible</u> if $\varphi: [-\tau,0] \to R^n$ and $u: [-\tau,t_1] \to R^m$, $t_1 \geqq 0$ are Borel measurable functions satisfying:

$$(2.5) \qquad \begin{aligned} &\varphi(t) \in \Phi(t), \quad -\tau \leqq t \leqq 0, \\ &u(t) \in U(t), \quad -\tau \leqq t \leqq t_1. \end{aligned}$$

The selection theorem of Kuratowski and Ryll-Nardzewski [K-3] assures the existence of admissible triples.

<u>Remark 2.1.</u> It is noted that if $u: [a,b] \to R^m$ is a Borel measurable function, then the function $t \to h(u(t),t)$, $t \in [a,b]$ is also Borel measurable. This follows easily from the fact that there is a sequence of Borel functions, $u_n: [a,b] \to R^m$, whose range is a

countable set, and which satisfy $\lim u_n(t) = u(t)$ for each $t \in [a,b]$. It follows now from the assumptions on h that $t \to h(u_n(t),t)$, $t \in [a,b]$ are each Borel measurable functions and $\lim h(u_n(t),t) = h(u(t),t)$, $t \in [a,b]$. Consequently, $t \to h(u(t),t)$, $t \in [a,b]$ is Borel measurable.

For any admissible triple $\{\varphi,u,t_1\}$ there is a unique absolutely continuous function (or response) $t \to x(t,\varphi,u)$, $0 \leq t \leq t_1$ satisfying (2.3) almost everywhere on $[0,t_1]$ and the initial condition

$$(2.6) \qquad x(t,\varphi,u) = \varphi(t), \qquad -\tau \leq t \leq 0.$$

According to the variation of parameters formula [B-1], this response is given by

$$(2.7) \quad x(t,\varphi,u) = \varphi(0)Y(0,t) + \int_{-\tau}^{0} \varphi(s)d_s\{\int_{0}^{\tau} F(\alpha,s-\alpha)Y(\alpha,t)d\alpha\}$$
$$+ \int_{0}^{t} \{\int_{-\tau}^{\alpha} h(u(s),s)d_s G(\alpha,s)\}Y(\alpha,t)d\alpha,$$

where for fixed $t \geq 0$ the function $s \to Y(s,t)$, $0 \leq s \leq t$ is an $n \times n$ matrix solution of

$$(2.8) \quad Y(s,t) + \int_{s}^{t} F(\alpha,s-\alpha)Y(\alpha,t)d\alpha = E, \qquad 0 \leq s \leq t,$$

which is of bounded variation and which satisfies $Y(t,t) = E$, the $n \times n$ identity matrix, and $Y(s,t) \equiv 0$ for $s > t$.

A point $x \in R^n$ is <u>attainable</u> if there is an admissible triple $\{\varphi, u, t_1\}$ such that $x(t_1, \varphi, u) = x$. The <u>attainable set</u> $\mathscr{A}(\Phi, U)$ (or simply $\mathscr{A}$ when Φ and U are understood) is defined by the equation

$$\mathscr{A}(\Phi, U) \equiv \{x \in R^n \mid x \text{ is attainable}\}.$$

The fixed time cross sections of $\mathscr{A}(\Phi, U)$ at $t \geq 0$ are denoted by $\mathscr{A}_t(\Phi, U)$ (or simply by $\mathscr{A}_t$ when Φ and U are understood) and are defined by the equation

$$\mathscr{A}_t(\Phi, U) \equiv \{x \in R^n \mid \text{there exist } \{\varphi, u, t\} \text{ admissible such that } x(t, \varphi, u) = x\}.$$

§3. Properties of the Attainable Set without Convexity Assumptions. We begin with some simple lemmas and observations.

Lemma 3.1. Let the standing hypotheses of section 2 be satisfied. Then $|Y(s,t)| \leq |E| \exp \int_s^t \beta(\xi)d\xi$, $0 \leq s \leq t$.

Proof. This is an easy consequence of (2.8) and the boundary conditions.

Remark 3.1. If $\mathscr{I}$ is a compact interval and $H: \mathscr{I} \to \mathscr{L}_{pq}$ is of bounded variation, then H has the well known decomposition into a sum of a singular function, an absolutely continuous function, and a saltus (jump) function. We note also that if $H = A + N$ where A is the saltus function and N is continuous, then $\operatorname{Var} H = \operatorname{Var} A + \operatorname{Var} N$. It is also observed that if H is continuous, then $t \to T_H(t)$, $t \in \mathscr{I}$ is also continuous. Consequently from (2.1) it can be shown that $|\mu_H|$ is nonatomic whenever H is continuous.

The next lemma is in essence contained in the papers of Liapunov [L-6], Blackwell [B-6], and Olech [O-1]. There are, however, some technical differences so we include a proof for the sake of completeness.

Lemma 3.2. Let $\mathscr{I}$ be a compact interval and let $H: \mathscr{I} \to \mathscr{L}_{pq}$ of bounded variation on $\mathscr{I}$. Let $\Omega: \mathscr{I} \to R^p$ be a μ_H-measurable compact multifunction. Let $\rho \in \mathscr{L}_1(\mathscr{I}, |\mu_H|, R)$ be such that diam $\Omega(t) \leq \rho(t)$, $t \in \mathscr{I}$. Then $\int_{\mathscr{I}} \Omega(t)dH(t)$ is compact.

Proof. First we observe that by the Lebesgue-Nikodym theorem (for example see [D-3, pg. 263]) there is a $|\mu_H|$-integrable function $B: \mathscr{I} \to \mathscr{L}_{pq}$ such that

$$\int_{\mathscr{I}} g(t)dH(t) = \int_{\mathscr{I}} g(t)B(t)d|H|(t), \quad g \in \mathscr{L}_1(\mathscr{I}, \mu_H, R^p).$$

We write $T_H = \alpha + \nu$ where α is the saltus function of T_H and ν is continuous. It is an easy matter to prove that the multifunction, $t \to \Omega(t)B(t)$, $t \in \mathscr{I}$ is measurable where $\Omega(t)B(t) = \{x \in R^q | x = yB(t)$ for some $y \in \Omega(t)\}$. Moreover, diam $\Omega(t)B(t) \leq \rho(t)|B(t)|$, $t \in \mathscr{I}$. One can also verify the identities:

$$(3.1) \qquad \int_{\mathscr{I}} \Omega(t)dH(t) = \int_{\mathscr{I}} \Omega(t)B(t)d|H|(t) = \int_{\mathscr{I}} \Omega(t)B(t)d\alpha(t) + \int_{\mathscr{I}} \Omega(t)B(t)d\nu(t);$$

the proof of the first equality is facilitated by versions of Filippov's selection lemma [C-3, J-1], and the second equality follows from the definition of α and ν. Now μ_α is purely atomic and μ_ν is nonatomic so the conclusion of the lemma follows from (3.1) and a remark of Olech's [O-1, pg. 100] (see [C-3] also for the nonatomic case).

Lemma 3.3. Let H and $\mathscr{I}$ be as in Lemma 3.2. Let $\Omega: \mathscr{I} \to R^p$

be a multifunction with a Borel measurable selection; that is there is a Borel function $g^*: \mathscr{I} \to R^p$ such that $g^*(t) \in \Omega(t)$ for each $t \in \mathscr{I}$. Then $\int_{\mathscr{I}} \Omega(t)dH(t)$ coincides with the set $\mathscr{B}(\Omega,H) \equiv \{\int_{\mathscr{I}} g(t)dH(t) \mid$ the function $g: \mathscr{I} \to R^p$ is Borel measurable and $g(t) \in \Omega(t),\ t \in \mathscr{I}\}$

<u>Proof</u>. Clearly $\mathscr{B}(\Omega,H) \subset \int_{\mathscr{I}} \Omega(t)dH(t)$. Conversely suppose $g \in \mathscr{L}_1(\mathscr{I},\mu_H,R^p)$, $g(t) \in \Omega(t)$, $t \in \mathscr{I}$. Then there is a Borel set $E_o \subset \mathscr{I}$ with $\mu_H(E_o) = 0$ and there is a Borel function $\bar{g}: \mathscr{I} \to R^p$ such that $\bar{g} = g$ on $\mathscr{I} \setminus E_o$ [R-1, pg. 225]. Using χ_S for the characteristic function of a set S we see that $\tilde{g} \equiv \bar{g}\cdot\chi_{\mathscr{I}\setminus E_o} + g^*\cdot\chi_{E_o}$ is a Borel function satisfying $\tilde{g}(t) \in \Omega(t)$, $t \in \mathscr{I}$, and $\tilde{g} = g$ a.e. $[\mu_H]$. Hence $\int_{\mathscr{I}} \tilde{g}(t)dH(t) = \int_{\mathscr{I}} g(t)dH(t)$, and so $\int_{\mathscr{I}} g(t)dH(t) \in \mathscr{B}(\Omega,H)$. This completes the proof.

In preparation for the next lemma let us introduce some additional notation. $\tilde{F},\tilde{G}: [0,\infty) \times R \to \mathscr{L}_{nn}$ are mappings defined by the following two relations,

$$\tilde{F}(t,s) \equiv \int_0^\tau F(\alpha,s-\alpha)Y(\alpha,t)d\alpha,$$

$$\tilde{G}(t,s) \equiv \int_0^t G(\alpha,s)Y(\alpha,t)d\alpha, \quad t \geq 0,\ s \in R,$$

where $F,G,$ and Y are the functions defined in section 2 which appear in equations (2.3) and (2.7). We define a function $\mathscr{A}: R^m \times R \to R^m \times R^n$ by the equation

$$\mathscr{A}(u,t) \equiv (u,h(u,t)),\ t \in R,\ u \in R^m,$$

where h is the function introduced in section 2 (see equation (2.3)).
A function $\Gamma\colon [0,\infty) \times R \to \mathscr{L}_{(m+n)n}$ is defined by the equation

$$\Gamma(t,s) = \begin{bmatrix} 0_{mn} \\ \tilde{G}(t,s) \end{bmatrix},\quad t \geq 0,\quad s \in R,$$

where 0_{mn} denotes an $m \times n$ matrix all of whose entries are zero.
A multifunction $L\colon R \to R^m \times R^n$ is defined by the condition

$$L(t) = \mathscr{A}(U(t),t),\ t \geq -\tau.$$

<u>Remark 3.2.</u> The sets $L(t)$, $t \geq -\tau$ are evidently compact. Let $b \geq 0$ be given. If μ is any Lebesgue-Stieltjes measure on $[-\tau,b]$, then the multifunction $U|[-\tau,b]$ is μ-measurable. This follows from the assumption that U is Borel measurable. Using Lusin's theorem for μ-measurable multifunctions [C-3, J-1] and extensions of Scorza-Dragoni's theorem [C-3, J-2] it can be proved that the multifunction $t \to L(t) = \mathscr{A}(U(t),t)$, $t \in [-\tau,b]$ is μ-measurable. Hence $L|[-\tau,b]$ is μ-measurable for every Lebesgue-Stieltjes measure μ on $[-\tau,b]$. We note also that because U is Borel measurable and compact, there is a Borel measurable function $u^*\colon [-\tau,\infty) \to R^m$ such that $u^*(t) \in U(t)$, $t \geq -\tau$ [K-3]. Remark 2.1 shows then that L has a Borel measurable selection. Recall now

that Φ was also assumed to be Borel measurable. Hence in evaluating either of the integrals

$$\int_{-\tau}^{0} \Phi(s)d_s\tilde{F}(t,s) \qquad \text{or} \qquad \int_{-\tau}^{t} L(s)d_s\Gamma(t,s)$$

the conclusion of Lemma 3.3 may be applied.

<u>Lemma 3.4</u>. Let the standing hypotheses of section 2 be satisfied. For $t \geqq 0$ define $\mathscr{R}_t(\Phi,U)$ to be the set

$$\Phi(0)Y(0,t) + \int_{-\tau}^{0} \Phi(s)d_s\tilde{F}(t,s) + \int_{-\tau}^{t} L(s)d_s\Gamma(t,s).$$

Then we have the identity: $\mathscr{A}_t(\Phi,U) = \mathscr{R}_t(\Phi,U)$, $t \geqq 0$.

<u>Proof</u>. Examining the third summand on the right hand side of (2.7) we use the unsymmetric Fubini theorem [C-1] and the assumptions on G in section 2 to write

$$(3.2) \qquad \int_{0}^{t} \{\int_{-\tau}^{\alpha} h(u(s),s)d_sG(\alpha,s)\}Y(\alpha,t)d\alpha =$$

$$\int_{0}^{t} \{\int_{-\tau}^{t} h(u(s),s)d_sG(\alpha,s)\}Y(\alpha,t)d\alpha =$$

$$\int_{-\tau}^{t} h(u(s),s)d_s\{\int_{0}^{t} G(\alpha,s)Y(\alpha,t)d\alpha\}.$$

We have the identity*

$$(3.3)\qquad \int_{-\tau}^{t} h(u(s),s)d_s\tilde{G}(t,s) = \int_{-\tau}^{t} \mathscr{A}(u(s),s)d_s\Gamma(t,s).$$

Consequently from (3.2), (3.3), and (2.7) we have

$$\mathscr{A}_t(\Phi,U) \subset \mathscr{R}_t(\Phi,U).$$

Since $F(t,\cdot)$ is left continuous on $(-\tau,0)$ we have that if $-\tau \leq s_n < 0$, and $s_n \to 0$ as $n \to \infty$, then $\lim F(\alpha, s_n-\alpha) = F(\alpha,-\alpha)$, $0 < \alpha \leq \tau$. Therefore from the Lebesgue dominated convergence theorem and the definition of $\tilde{F}$ we get that $\{0\}$ is not an atom of $\mu_{\tilde{F}(t,\cdot)}$. Hence

$$\int_{-\tau}^{0} \varphi(s)d_s\tilde{F}(t,s) = \int_{-\tau}^{0} \bar{\varphi}(s)d_s\tilde{F}(t,s)$$

if $\varphi(s) = \bar{\varphi}(s)$ except at $s = 0$. From this remark, the variation of parameters formula (2.7), and Remark 3.2 one can show the reverse inclusion $\mathscr{R}_t(\Phi,U) \subset \mathscr{A}_t(\Phi,U)$. If the detailed proof of this inclusion is carried out, then the meaning of the comment in the preceding footnote becomes clear. This completes the proof of the representation formula of the lemma.

---

*Our reason for introducing the auxiliary function $\mathscr{A}$ and Γ is to avoid certain questions concerning the existence of Borel measurable selections. Halkin used a similar device in [H-5].

Theorem 3.1. Let the standing hypotheses of section 2 be satisfied. Then

(i) The sets $\mathscr{A}_t(\Phi,U)$, $t \geq 0$ are compact;

(ii) The mapping $t \to \mathscr{A}_t(\Phi,U)$, $t \geq 0$ taking its values in the compact nonempty subsets of R^n is continuous with respect to the Hausdorff metric [A-1];

(iii) For any $\bar{t} \geq 0$ the set $\bigcup_{t\in[0,\bar{t}]} \mathscr{A}_t(\Phi,U)$ is compact.

Proof of (i). This is an immediate consequence of the representatio formula in Lemma 3.4 and Lemma 3.2.

Proof of (iii). This is readily deduced from (ii).

Proof of (ii). Let S denote the closed unit ball in R^n with center at the origin. We must prove that given $t_1 \geq 0$ and $\epsilon > 0$ there is a $\delta > 0$ such that

$$(3.4) \qquad \mathscr{A}_{t_1} + \epsilon S \supset \mathscr{A}_t \quad \text{and} \quad \mathscr{A}_t + \epsilon S \supset \mathscr{A}_{t_1}, \quad |t-t_1| \leq \delta,\ t \geq 0.$$

The relations in (3.4) can be verified by considering two cases: Case (a). $t \geq t_1$; Case (b). $0 \leq t < t_1$. Consider case (a) first. Suppose $x_t \in \mathscr{A}_t$; then there is an admissible triple $\{\varphi_t,u_t,t\}$ such that $x_t = x(t,\varphi_t,u_t)$. Define the function $\bar{u}$: $[-\tau,t_1] \to R^m$ to be the restriction of u_t to $[-\tau,t_1]$. Using the variation of parameters formula (2.7), Lemma 3.1, inequalities (2.2) and (2.4), hypothesis 4°) of section 2, and some standard manipulations with Lebesgue-Stieltjes integrals we obtain the estimate

$$(3.5)\quad |x(t,\varphi_t,u_t) - x(t_1,\varphi_t,\bar{u})| \leq M|Y(0,t) - Y(0,t_1)| + M\int_0^\tau \beta(\alpha)|Y(\alpha,t) - Y(\alpha,t_1)|d\alpha + M|E|\int_{t_1}^t \beta(\alpha)[\exp\int_\alpha^t \beta(\xi)d\xi]d\alpha + M\int_0^{t_1} \beta(\alpha)|Y(\alpha,t) - Y(\alpha,t_1)|d\alpha.$$

We now give a similar estimate for Case (b). By the Kuratowski-Ryll-Nardzewski selection theorem [K-3] there is a Borel function $u^*\colon [-\tau,\infty) \to R^m$ such that $u^*(t) \in \Omega(t)$, $t \geq -\tau$. We note that $u_1 \equiv u_t \circ \chi_{[-\tau,t]} + u^* \circ \chi_{(t,t_1]}$ is a Borel function and $\{\varphi_t, u_1, t_1\}$ is admissible. For reasons similar to those adduced to support (3.5) we get the inequality

$$(3.5')\quad |x(t,\varphi_t,u_t) - x(t_1,\varphi_t,u_1)| \leq M|Y(0,t) - Y(0,t_1)| + M\int_0^\tau \beta(\alpha)|Y(\alpha,t) - Y(\alpha,t_1)|d\alpha + M|E|\int_t^{t_1} \beta(\alpha)[\exp\int_\alpha^{t_1} \beta(\xi)d\xi]d\alpha + M\int_0^t \beta(\alpha)|Y(\alpha,t) - Y(\alpha,t_1)|d\alpha.$$

From the continuity of $Y(\alpha,\cdot)$, the Lebesgue dominated convergence theorem, and inequalities (3.5) and (3.5') there results

(3.6) Given $t_1 \geq 0$ and $\epsilon > 0$ there is a $\delta > 0$ depending only on t_1 and ϵ such that $|t-t_1| \leq \delta$, $t \geq 0$ imply $|x(t,\varphi_t,u_t) - x(t_1,\varphi_t,\bar{u})| \leq \epsilon$ and $|x(t,\varphi_t,u_t) - x(t_1,\varphi_t,u_1)| \leq \epsilon$.

Statement (3.6) implies

$$\mathscr{A}_{t_1} + \epsilon S \supset \mathscr{A}_t, \quad |t-t_1| \leq \delta, \quad t \geq 0.$$

The other inclusion relationship in (3.4) is proved by a symmetric argument which is omitted.

Remark 3.3. Let $t \to \mathscr{T}(t)$, $t \geq 0$ be a compact multifunction which is continuous with respect to the Hausdorff metric. If we impose a terminal condition of the form

$$x(t_1,\varphi,u) \in \mathscr{T}(t_1), \tag{3.7}$$

then by the usual device [N-2] Theorem 3.1 yields an existence theorem for the time optimal control problem. If we consider only admissible controls whose domain $[-\tau,t_1]$ lies in some fixed interval $[-\tau,\bar{t}]$, and if there is a terminal constraint (3.7) or, indeed, if the right end is free, then Theorem 3.1 can be used to give an existence theorem for the problem of minimizing $P(x(t_1,\varphi,u))$ on the class of admissible triples $\{\varphi,u,t_1\}$ such that (3.7) is satisfied*, or for the problem of minimizing $P(x(t_1,\varphi,u))$ on the

*Actually for these existence statements it is not necessary to assume that the multifunction $\mathscr{T}$ is continuous or even compact. It suffices to have the multifunction $\mathscr{T}$ closed (i.e., $\mathscr{T}(t)$ is closed for $t \geq 0$) and upper semicontinuous in the Kuratowski sense (see for example [K-4, C-4, J-1]). We keep the stronger hypothesis of continuity because it is needed in proving necessary conditions for a minimum.

class of admissible triples $\{\varphi, u, t_1\}$, where P is a real valued continuous function on R^n.

In order to deduce necessary conditions for the optimization problems mentioned in Remark 3.3 it is desirable to have that the sets $\mathscr{A}_t(\Phi,U)$ are convex. This cannot be deduced under the general circumstances of Theorem 3.1 because the Lebesgue-Stieltjes measures involved in the representation formula, $\mathscr{A}_t(\Phi,U) = \mathscr{R}_t(\Phi,U)$, of Lemma 3.4 can be atomic. It is noted that any function on an interval $[a,b]$ into $\mathscr{L}_{pq}$ which is of bounded variation has only a denumerable number of discontinuities. We say that F has <u>property (Δ_1)</u> if for each $t \in R$ it is possible to index the points $-\theta_i(t)$, $i = 1,2,\ldots$, in the interior of $[-\tau,0]$ at which $F(t,\cdot)$ is discontinuous, in such a way that continuous functions $t \to \theta_i(t)$, $t \in R$, are defined and $t \to t - \theta_i(t)$, $t \in R$ is strictly increasing $i = 1,2,\ldots$. We say that G has <u>property (Δ_2)</u> if for each $t \in R$ it is possible to index the points $\zeta_i(t)$, $i = 1,2,\ldots$ in the interior of $[-\tau,t]$, at which $G(t,\cdot)$ is discontinuous in such a way that continuous strictly increasing functions $t \to \zeta_i(t)$, $t \in R$, $i = 1,2,\ldots$, are defined.

<u>Theorem 3.2</u>. If in addition to the standing hypotheses of section 2 we assume that $\Phi(0)$ is convex, F and G are Borel measurable, F has property (Δ_1), and G has property (Δ_2), then conclusions (i), (ii), and (iii) of Theorem 3.1 are still valid and $\mathscr{A}_t(\Phi,\Omega)$, $t \geqq 0$ are convex.

Before proceeding with the proof we give another lemma that will be useful in the proof.

Lemma 3.5. Let $\rho\colon R \to R$ be a continuous strictly increasing function. Let $f\colon [a,b] \to \mathcal{L}_{pq}$ be a Lebesgue summable function. We define three functions $W_1, W_2, W_3\colon R \to \mathcal{L}_{pq}$ by the equations

$$W_1(s) = \int_a^b f(\xi)I(s-\rho(\xi))d\xi, \qquad s \in R,$$

$$W_2(s) = \int_a^b f(\xi)J(s-\rho(\xi))d\xi, \qquad s \in R$$

$$W_3(s) = \int_a^b f(\xi)J(-s+\rho(\xi))d\xi, \qquad s \in R$$

where $I, J\colon R \to R$ are the step functions defined by the relations

$$I(x) = \begin{cases} 0 & x \leqq 0 \\ 1 & x > 0 \end{cases}$$

$$J(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geqq 0 \end{cases}$$

Then W_i is continuous, $i = 1,2,3$.

Proof. First we remark that $\rho\colon [a,b] \to [\rho(a), \rho(b)]$ has a continuous inverse, $\rho^{-1}\colon [\rho(a), \rho(b)] \to [a,b]$ which is also strictly increasing. Some elementary calculations yield the following formulas:

$$W_1(s) = \begin{cases} 0 & \text{for} \quad s \leqq \rho(a) \\ \int_a^b f(\xi)d\xi & \text{for} \quad s > \rho(b) \\ \int_a^{\rho^{-1}(s)} f(\xi)d\xi & \text{for} \quad \rho(a) < s \leqq \rho(b), \end{cases}$$

$$W_2(s) = \begin{cases} 0 & \text{for} \quad s < \rho(a) \\ \int_a^b f(\xi)d\xi & \text{for} \quad s \geqq \rho(b) \\ \int_a^{\rho^{-1}(s)} f(\xi)d\xi & \text{for} \quad \rho(a) \leqq s < \rho(b), \end{cases}$$

and

$$W_3(s) = \begin{cases} 0 & \text{for} \quad s > \rho(b) \\ \int_o^b f(\xi)d\xi & \text{for} \quad s \leqq \rho(a) \\ \int_{\rho^{-1}(s)}^b f(\xi)d\xi & \text{for} \quad \rho(a) < s \leqq \rho(b). \end{cases}$$

The continuity of the functions W_i, $i = 1,2,3$ is an immediate consequence of these formulas and the continuity of ρ^{-1} on $[\rho(a),\rho(b)]$.

Proof of Theorem 3.2. We write $F(t,s) = A_F(t,s) + N_F(t,s)$ and $G(t,s) = A_G(t,s) + N_G(t,s)$ where $A_F(t,\cdot)$ is the saltus function for $F(t,\cdot)$, $A_G(t,\cdot)$ is the saltus function for $G(t,\cdot)$, and both $N_F(t,\cdot)$ and $N_G(t,\cdot)$ are continuous. Denote the jump of $F(t,\cdot)$ at $-\theta_i(t)$ by $B_i(t)$ and the jump of $G(t,\cdot)$ at $\zeta_i(t)$ by $C_i(t)$, $i = 1,2,\ldots$. The jump of $F(t,\cdot)$ at $-\tau$ is denoted by $B_{-1}(t)$ and the jump of $F(t,\cdot)$ at 0 is denoted by $B_o(t)$. The jump of $G(t,\cdot)$ at $-\tau$ is denoted by $C_{-1}(t)$ and the jump at t is denoted by $C_o(t)$. From Remark 3.1 and inequalities (2.2) it follows that

$$(3.8) \qquad \sum_{i=1}^{\infty} |B_i(t)|, \quad \sum_{i=1}^{\infty} |C_i(t)| \leq \beta(t), \quad t \in R.$$

Since F and G are Borel measurable, the functions B_i, C_i, $i = -1,0,1,2,\ldots$ are all Lebesgue measurable. For example let us show B_i is Lebesgue measurable, $i \geq 1$. Define $s_n(t) = 1/n - \theta_i(t)$, $n = 1,2,3,\ldots$, then $s_n(t) > -\theta_i(t)$ and $\lim s_n(t) = -\theta_i(t)$. Since $F(t,\cdot)$ is left continuous on $(-\tau,0)$ we have

$$B_i(t) = \lim F(t,s_n(t)) - F(t,-\theta_i(t)).$$

Since F is Borel measurable, the functions $t \to F(t,s_n(t))$ and $t \to F(t,-\theta_i(t))$ are Borel measurable (a fortiori Lebesgue measurable). Hence B_i is a Borel function and thus Lebesgue measurable. The proof of the measurability of the other functions is similar. Define

$B(t) \equiv \sum_{i=1}^{\infty} B_i(t)$ and $C(t) \equiv \sum_{i=1}^{\infty} C_i(t)$ (both series converge by (3.8)).

The saltus functions A_F and A_G can be written in the form

$$(3.9)\qquad A_F(t,s) = B_{-1}(t)J(-s-\tau) + (B_o(t) - B(t))J(s) + \sum_{i=1}^{\infty} B_i(t)I(s+\theta_i(t)),$$

and

$$A_G(t,s) = C_{-1}(t)J(-s-\tau) + (C_o(t) - C(t))J(s-t) + \sum_{i=1}^{\infty} C_i(t)I(s-\zeta_i(t)).$$

We have

$$(3.10)\qquad \tilde{F}(t,s) = \int_o^{\tau} A_F(\alpha,s-\alpha)Y(\alpha,t)d\alpha + \int_o^{\tau} N_F(\alpha,s-\alpha)Y(\alpha,t)d\alpha$$

$$\tilde{G}(t,s) = \int_o^{t} A_G(\alpha,s)Y(\alpha,t)d\alpha + \int_o^{t} N_G(\alpha,s)Y(\alpha,t)d\alpha,$$

and the second terms on the right hand side of both equations depend continuously on s by the Lebesgue dominated convergence theorem. Using (3.8), (3.9) and the dominated convergence theorem we get

$$(3.11)\qquad \int_o^{\tau} A_F(\alpha,s-\alpha)Y(\alpha,t)d\alpha = \int_o^{\tau} B_{-1}(\alpha)Y(\alpha,t)J(-s+\alpha-\tau)d\alpha +$$

$$\int_o^{\tau} (B_o(\alpha) - B(\alpha))Y(\alpha,t)J(s-\alpha)d\alpha + \sum_{i=1}^{\infty} \int_o^{\tau} B_i(\alpha)Y(\alpha,t)I(s-\alpha+\theta_i(\alpha))d\alpha.$$

According to property (Δ_1) and Lemma 3.5 each term in the series (3.11) is continuous in s. We also have

$$(3.12) \quad \left|\int_0^{\tau} B_i(\alpha)Y(\alpha,t)I(s-\alpha+\theta_i(\alpha))d\alpha\right| \leq K|E|\int_0^{\tau} |B_i(\alpha)|d\alpha, \quad i = 1,2,\ldots,$$

where $K = \exp \int_0^t \beta(\xi)d\xi$. Moreover, the series $\sum_{i=1}^{\infty} \int_0^{\tau} |B_i(\alpha)|d\alpha$ converges by (3.8). Hence by the Weierstrass M-test and (3.12) the series in (3.11) converges uniformly for $s \in [-\tau,0]$ ($t \geqq 0$ is fixed). Therefore the function $s \to \int_0^{\tau} A_F(\alpha, s-\alpha)Y(\alpha,t)d\alpha$, $-\tau \leqq s \leqq 0$ is continuous, and we conclude that $\tilde{F}(t,\cdot)$ is continuous on $[-\tau,0]$ for each fixed $t \geqq 0$. By an entirely parallel argument it can be shown that $\tilde{G}(t,\cdot)$ (also $\Gamma(t,\cdot)$) is continuous for each fixed $t \geqq 0$. Using the Lebesgue-Nikodym theorem [D-3, pg. 263] it is determined that there exist integrable Borel functions $V_{\tilde{F}}: [-\tau,0] \to \mathcal{L}_{nn}$ and $V_{\Gamma}: [-\tau,t] \to \mathcal{L}_{(m+n)n}$ such that

$$\int_{-\tau}^{0} \varphi(s)d_s\tilde{F}(t,s) = \int_{-\tau}^{0} \varphi(s)V_{\tilde{F}}(s)d|\mu_{\tilde{F}(t,\cdot)}|$$

for $\varphi \in \mathcal{L}_1([-\tau,0], \mu_{\tilde{F}(t,\cdot)}, R^n)$ and

$$\int_{-\tau}^{t} g(s)d_s\Gamma(t,s) = \int_{-\tau}^{t} g(s)V_{\Gamma}(s)d|\mu_{\Gamma(t,\cdot)}|,$$

for $g \in \mathcal{L}_1([-\tau,t], \mu_{\Gamma(t,\cdot)}, R^n)$. From the representation formula in Lemma 3.4 and an extension of Filippov's selection principle [C-3, J-1] we obtain

$$(3.13)\quad \mathscr{A}_t(\Phi,U) = \Phi(0)Y(0,t) + \int_{-\tau}^{0} \Phi(s)V_{\tilde{F}}(s)d|\mu_{\tilde{F}(t,\cdot)}| + \int_{-\tau}^{t} L(s)V_{\Gamma}(s)d|\mu_{\Gamma(t,\cdot)}|.$$

By Remark 3.1 $|\mu_{\tilde{F}(t,\cdot)}|$ and $|\mu_{\Gamma(t,\cdot)}|$ are nonatomic, and we conclude that $\mathscr{A}_t(\Phi,U)$ is convex (see [O-1, C-3]).

We shall use co (B) to denote the convex hull of a set $B \subset R^p$.

Corollary 3.1. Let Ψ denote the multifunction, $t \to \text{co}\,(\Phi(t))$, $-\tau \leq t \leq 0$. Let $U^*\colon [-\tau,\infty) \to R^m$ be a Borel measurable compact multifunction such that $\text{diam}\,(U^*(t)) \leq M$ and $\text{co}\,(h(U(t),t)) = \text{co}\,(h(U^*(t),t))$ for $t \geq -\tau$. Let the hypotheses of Theorem 3.2 be satisfied. Then

$$\mathscr{A}_t(\Phi,U) = \mathscr{A}_t(\Psi,U^*), \qquad t \geq 0.$$

Proof. The identity is easily verified by using equation (3.13), the linearity of $V_{\tilde{F}}(s)$ and $V_{\Gamma}(s)$ and Theorem 7.1 in [C-3].

As a particular case of Corollary 3.1 we obtain:

Corollary 3.2. Let the hypotheses of Corollary 3.1 be satisfied. In addition suppose $\mathring{\Psi}\,(t)$ is the set of extreme points of $\Psi(t)$, and $U^{\#}\colon [-\tau,\infty) \to R^n$ is a multifunction such that $\text{diam}\, U^{\#}(t) \leq M$

and the set of extreme points of co $(h(U(t),t))$ is equal to $h(U^{\#}(t),t)$, $t \geq -\tau$. If the multifunctions $\ddot{\Psi}$ and $U^{\#}$ are compact and Borel measurable, then

$$\mathscr{A}_t(\Phi,U) = \mathscr{A}_t(\ddot{\Psi},U^{\#}), \qquad t \geq 0.$$

Remark 3.4. In Theorem 3.2 it was assumed that $\Phi(0)$ is convex. If this should happen not to be the case, then one can always select Φ_o, a compact convex subset of $\Phi(0)$ (for example Φ_o could be a singleton point set), and define

$$\Phi^*(t) = \begin{cases} \Phi(t) & t \neq 0 \\ \Phi_o & t = 0 \end{cases}$$

Since Φ^* is also Borel measurable, compact, and satisfies diam Φ $(t) \leq M$, $t \in [-\tau,0]$ we could replace Φ by Φ^* and Theorem 3.2 could be applied.

§4. Necessary Conditions for an Optimal Control. The properties of the attainable sets deduced in section 3 suggest that the main geometric ideas involved in proving the maximal principle for ordinary linear control problems (see [L-3]) are going to retain their validity for certain of the optimization problems formulated in Remark 3.3. We shall only consider the time optimal control problem mentioned in Remark 3.3. It will be clear from the discussion that the results can be used to prove a maximal principle for the other problem discussed in the aforementioned remark if we add additional assumptions which assure that on compact convex subsets $\mathscr{A}_t$ of R^n the mapping P assumes its minimum on $\partial\mathscr{A}_t$, the boundary of $\mathscr{A}_t$, e.g., when P is linear (cf. [H-2]).

The following lemma is true and the proof is in effect given in [L-3].

Lemma 4.1. Let $\mathfrak{F}, \mathfrak{G}: [a,b] \to R^n$ be compact multifunctions which are continuous with respect to the Hausdorff metric. Let $\mathfrak{F}(t)$ be convex for $a \leqq t \leqq b$. Let $t^* \in (a,b]$ be such that $\mathfrak{F}(t^*) \cap \mathfrak{G}(t^*) \neq \emptyset$ and $\mathfrak{F}(t) \cap \mathfrak{G}(t) = \emptyset$ if $a \leqq t < t^*$. Then $x^* \in \mathfrak{F}(t^*) \cap \mathfrak{G}(t^*)$ implies $x^* \in \partial\mathfrak{F}(t^*)$.

We shall use $\langle x,y\rangle$ to denote the scalar product, $x,y \in R^n$, and A' to denote the transpose of a matrix A.

Theorem 4.1. Let the hypotheses of Theorem 3.2 be satisfied. If $\{\varphi^*,u^*,t^*\}$ is an optimal solution to the time optimal control problem in Remark 3.3, then there is a function $\psi\colon [0,t^*] \to R^n$ which is of bounded variation and satisfies the adjoint equation

$$\psi(s) + \int_s^{t^*} \psi(\alpha)F'(\alpha,s-\alpha)d\alpha = e, \qquad 0 \le s \le t^*,$$

where e is an outward normal to a support hyperplane to the set $\mathscr{A}_{t^*}(\Phi,U)$ through the point $x(t^*,\varphi^*,u^*)$ on the boundary of $\mathscr{A}_{t^*}(\Phi,U)$, such that

$$1^\circ)\quad \langle\varphi^*(0),\psi(0)\rangle \ge \langle\varphi_0,\psi(0)\rangle, \qquad \varphi_0 \in \Phi(0);$$

$$2^\circ)\quad \int_0^\tau \Big\langle\int_{-\tau}^0 \varphi^*(s)d_sF(\alpha,s-\alpha),\psi(\alpha)\Big\rangle d\alpha \ge$$

$$\int_0^\tau \Big\langle\int_{-\tau}^0 \varphi(s)d_sF(\alpha,s-\alpha),\psi(\alpha)\Big\rangle d\alpha$$

for every admissible φ;

$$3^\circ)\quad \int_0^{t^*} \Big\langle\int_{-\tau}^{t^*} h(u^*(s),s)d_sG(\alpha,s),\psi(\alpha)\Big\rangle d\alpha \ge$$

$$\int_0^{t^*} \Big\langle\int_{-\tau}^{t^*} h(u(s),s)d_sG(\alpha,s),\psi(\alpha)\Big\rangle d\alpha.$$

for every admissible u. Moreover if $\mathscr{T}(t)$ is equal to a fixed compact convex set $\mathfrak{T} \subset R^n$ for $t \ge 0$, then e can be picked to satisfy the transversality condition: e is normal to a common support hyperplane separating $\mathscr{A}_{t^*}(\Phi,U)$ and $\mathfrak{T}$.

<u>Proof.</u> Let $x^* = x(t^*, \varphi^*, u^*)$. By Theorem 3.2 and Lemma 4.1 we infer that x^* belongs to the boundary of $\mathscr{A}_{t^*}$. There is a vector $e \in R^n$ with $|e| = 1$ such that

$$\max\{\langle e,x\rangle \mid x \in \mathscr{A}_{t^*}\} = \langle e,x^*\rangle.$$

Using the fact that the value of $\varphi(0)$ does not affect the value of the second term on the right hand side of (2.7) (cf. proof of Lemma 3.4), and some elementary reasoning involving formula (2.7) it can be shown that

$$(4.1)\quad (a)\quad \langle \varphi^*(0),\ eY'(0,t^*)\rangle \geq \langle \varphi_0, eY'(0,t^*)\rangle,\quad \varphi_0 \in \Phi(0);$$

$$(b)\quad \langle \int_{-\tau}^{0} \varphi^*(s)d_s\left[\int_0^\tau F(\alpha,s-\alpha)Y(\alpha,t^*)d\alpha\right],\ e\rangle \geq$$

$$\langle \int_{-\tau}^{0} \varphi(s)d_s\left[\int_0^\tau F(\alpha,s-\alpha)Y(\alpha,t^*)d\alpha\right],\ e\rangle$$

for every admissible φ;

$$(c)\quad \langle \int_0^{t^*}\left[\int_{-\tau}^{t^*} h(u^*(s),s)d_sG(\alpha,s)\right]Y(\alpha,t^*)d\alpha,\ e\rangle \geq$$

$$\langle \int_0^{t^*}\left[\int_{-\tau}^{t^*} h(u(s),s)d_sG(\alpha,s)\right]Y(\alpha,t^*)d\alpha,\ e\rangle$$

for every admissible u. Define $\psi(\alpha) = eY'(\alpha,t^*)$, $0 \leq \alpha \leq t^*$.

Then by appropriately using the unsymmetric Fubini theorem [C-1] and some standard manipulation with the scalar product in (4.1)b,c, relations 2°) and 3°) are proved. The fact that $\alpha \to \psi(\alpha)$, $0 \leq \alpha \leq t^*$ is of bounded variation and satisfies the adjoint equation is an immediate consequence of (2.8).

The transversality condition is just a geometric property. In proving this condition we use the norm in R^n defined by $|x|^2 = \langle x,x\rangle$. We have $\mathscr{A}_t \cap \mathfrak{T} = \emptyset$ for $0 \leq t < t^*$. Let $t_n \in [0,t^*)$ be such that $t_n \to t^*$ as $n \to \infty$. Let $a_n \in \mathscr{A}_{t_n}$, $b_n \in \mathfrak{T}$ be such that $|a_n-b_n|$ is the minimum value that the function $(x,y) \to |x-y|$, $(x,y) \in \mathscr{A}_{t_n} \times \mathfrak{T}$ assumes. Then $a_n - b_n \neq 0$ and $e_n = (b_n-a_n)/|a_n-b_n|$ is a unit outer normal to $\mathscr{A}_{t_n}$ at a_n and a unit inner normal to $\mathfrak{T}$ at b_n. Hence

$$(4.2) \qquad \{x | \langle e_n, x-a_n\rangle \leq 0\} \supset \mathscr{A}_{t_n}$$

$$\{x | \langle e_n, x-b_n\rangle \geq 0\} \supset \mathfrak{T}, \quad n = 1,2,3,\ldots .$$

We might as well assume $e_n \to e$ and $b_n \to b$ as $n \to \infty$. Then a_n also converges to b. Using (4.2) and the fact that $\mathscr{A}_{t_n} \to \mathscr{A}_{t^*}$ as $n \to \infty$ (the limit is taken with respect to the Hausdorff metric) we find that

$$\{x | \langle e, x-b\rangle \leq 0\} \supset \mathscr{A}_{t^*}$$

$$\{x | \langle e, x-b\rangle \geq 0\} \supset \mathfrak{T}$$

so that $\pi = \{x | \langle e, x-b\rangle = 0\}$ is a hyperplane satisfying the transversality condition.

Remark 4.1. We can put conditions 2^o) and 3^o) of Theorem 4.1 in a form which will in many cases be more manageable if we assume that $F(t,\cdot)$ and $G(t,\cdot)$ have no singular part and if the functions θ_i, ζ_i, $i = 1,2,3,\ldots$ introduced in properties (Δ_1) and (Δ_2) are of class C^1. Let us indicate the form which 2^o) and 3^o) take in this case. We use the decompositions $F = A_F + N_F$ and $G = A_G + N_G$ which were introduced in the proof of Theorem 3.2. According to our assumptions $N_F(t,\cdot)$ and $N_G(t,\cdot)$ are absolutely continuous. By some rather involved analysis, which includes several applications of the unsymmetric Fubini theorem [C-1], it can be shown that condition 2^o) of Theorem 4.1 implies

$$2^o_1)\quad \langle\varphi^*(s), P(s)\rangle \geq \langle\varphi(s), P(s)\rangle \qquad \text{a.e. on } [-\tau, 0]$$

for every admissible φ where P is defined by

$$P(s) \equiv -\psi(s+\tau)B'_{-1}(s+\tau) + \sum_{i=1}^{\infty} \psi(\rho_i^{-1}(s))B'_i(\rho_i^{-1}(s))K_i(s)\nu_i(s)$$

$$+ \int_0^{\tau} \psi(\alpha)\frac{\partial N'_F}{\partial s}(\alpha, s-\alpha)d\alpha, \qquad s \in [-\tau, 0],$$

and $\rho_i(\alpha) \equiv \alpha - \theta_i(\alpha)$ for $\alpha \in [0,\tau]$, K_i denotes the characteristic function of $[\rho_i(0), \rho_i(\tau)] \cap [-\tau, 0]$, and $\nu_i(s) \equiv 1/\dot{\rho}_i(\rho_i^{-1}(s))$ for $s \in [-\tau, 0]$, $i = 1,2,3,\ldots$. By a similar type of analysis which is again omitted 3^o) can be shown to imply

$3^{\circ}_{1})$ $\langle h(u^{*}(-\tau),-\tau),\ \psi(\xi)C'_{-1}(\xi)\rangle \geqq \langle h(u(-\tau),-\tau),\ \psi(\xi)C'_{-1}(\xi)\rangle,$ a.e. on $[0,t^{*}]$,

and

$$\langle h(u^{*}(\xi),\xi),\ Q(\xi)\rangle \geqq \langle h(u(\xi),\xi),\ Q(\xi)\rangle, \quad \text{a.e. on} \quad [-\tau,t^{*}]$$

for every admissible u where Q is defined by

$$Q(\xi) \equiv \psi(\xi)(C'_{0}(\xi) - C'(\xi))\mathscr{K}_{0}(\xi) + \sum_{i=1}^{\infty}\psi(\zeta_{i}^{-1}(\xi))C'_{i}(\zeta_{i}^{-1}(\xi))\mathscr{K}_{i}(\xi)\beta_{i}(\xi) + \int_{0}^{t^{*}}\psi(\alpha)\frac{\partial N'_{G}}{\partial\xi}(\alpha,\xi)d\alpha$$

and $\mathscr{K}_{0}$ is the characteristic function of $[0,t^{*}]$, $\mathscr{K}_{i}$ is the characteristic function of $[\zeta_{i}(0),\ \zeta_{i}(t^{*})]$, and $\beta_{i}(\xi) \equiv 1/\dot{\zeta}_{i}(\zeta_{i}^{-1}(\xi))$, $-\tau \leqq \xi \leqq t^{*}$.

§5. Analyticity Results for Solutions of FDE's. As the reader is by now well aware, the representation (2.7) of solutions to (2.3) in terms of "fundamental" or "adjoint" matrix solutions [B-1] is of immense importance in the study of control of such systems. In this section we investigate analyticity properties of these fundamental matrix solutions for certain types of linear (in the state variable) systems which are special cases of (2.3); namely

$$(5.1) \qquad \dot{x}(t) = \sum_{i=0}^{K} x(t-\theta_i)A_i(t) + \int_{-\tau}^{0} x(t+s)A(t,s)ds + h(u(t),t)$$

with $0 = \theta_0 < \theta_1 < \dots < \theta_k \leqq \tau$, which correspond to an $F(t,\cdot)$ consisting of an absolutely continuous function plus a saltus function with a finite number of constant (in t) jump points. The associated fundamental matrices $X(t,\sigma)$ satisfy (as a function of t)

$$(5.2) \qquad \dot{X}(t) = \sum_{i=0}^{K} X(t-\theta_i)A_i(t) + \int_{-\tau}^{0} X(t+s)A(t,s)ds, \quad t > \sigma$$

$$X(\sigma) = E, \quad X(t) = 0 \quad \text{for} \quad t < \sigma.$$

Since the corresponding adjoint matrices $Y(\sigma,t)$ satisfy (in σ) systems [H-1] which can be put in a form similar to that of (5.2) and since we have $X(t,\sigma) = Y(\sigma,t)$, to investigate analyticity properties of X and Y in σ or t it suffices to examine the analyticity in t of solutions to (5.2). Considering the following two examples one sees that systems of the type (5.2) with analytic coefficients and analytic initial function need not possess an analytic solution.

Example 5.1. The scalar system

$$\dot{x}(t) = x(t-1), \quad t > 0$$
$$x(t) = 1, \quad t \in [-1,0]$$

has a unique solution on $[-1,2]$ given by

$$x(t) = \begin{cases} 1, & t \in [-1,0] \\ 1 + t, & t \in [0,1] \\ 3/2 + t^2/2, & t \in [1,2] \end{cases}$$

which is not analytic at $t = 1$.

Example 5.2. The scalar system

$$\dot{x}(t) = \int_{-1}^{0} x(t+s)ds, \quad t > 0$$
$$x(t) = 1, \quad t \in [-1,0]$$

has a unique solution on $[-1,2]$ given by

$$x(t) = \begin{cases} 1, & t \in [-1,0] \\ 1 + \sinh t, & t \in [0,1] \\ 1 + \sinh t - \sum_{n=1}^{\infty} (n-1)\frac{(t-1)^{2n-1}}{(2n-1)!}, & t \in [1,2], \end{cases}$$

which is not analytic at $t = 1$.

Remark 5.1. Example 5.1 can be used to contradict a theorem of Pinney [P-1, p. 237] while Example 5.2 contradicts a result due to Oguztöreli [O-2, p. 52]. (It is not difficult to show that the right side of the system in Example 5.2 is analytic in x in the sense of Volterra [V-1, O-2] as required in Oguztöreli's theorem.)

In light of the previous examples and remarks one might expect under reasonable assumptions on the coefficients to obtain not analyticity but some type of piecewise analyticity for solutions to (5.2). We are thus motivated to introduce the following concepts (see also Halkin [H-4] and Levinson [L-4]). A function $f: R \to R$ is analytic on $[a,b]$ if there exist $\epsilon > 0$ and a function g analytic on $(a-\epsilon, b+\epsilon)$ such that $f = g$ on $[a,b]$. We say that f is piecewise analytic (pwa) on $[a,b]$ if there exists a partition $a = s_0 < s_1 < \ldots < s_\nu = b$ such that f is analytic on $[s_{i-1}, s_i]$, $i = 1,2,\ldots,\nu$. Finally, f is said to be quasi piecewise analytic (qpwa) on $[a,b]$ if there exists a partition $a = s_0 < s_1 < \ldots < s_\nu = b$ such that f is analytic on (s_{i-1}, s_i), $i = 1,2,\ldots,\nu$.

Combining a modification of the step-method [E-1] with known results for ordinary linear differential equations we can prove

Theorem 5.1. Let $A(t,s) \equiv 0$ in (5.2) and $t \to A_i(t)$, $i = 0,1,\ldots,K$ be (real) analytic on $[\sigma,\infty)$ into $\mathcal{L}_{nn}$. If the lags θ_i, $i = 1,2,\ldots,K$, are commensurate, then the solution to (5.2) is pwa on $[\sigma,\sigma+T]$ for any $T > 0$.

Proof. We shall give the proof for $K = 1$, $\theta_1 = 1$ since it will then be clear how one extends the results to cover the case of a finite number of commensurate lags. Thus we consider the system

$$\text{(5.3)} \qquad \begin{aligned} \dot{X}(t) &= X(t)A_o(t) + X(t-1)A_1(t), \quad t > \sigma \\ X(\sigma) &= E \\ X(t) &= 0, \quad t < \sigma \end{aligned}$$

and denote by $\mathfrak{W}$ the solution of

$$\begin{aligned} \dot{\mathfrak{W}}(t) &= \mathfrak{W}(t)A_o(t) \\ \mathfrak{W}(\sigma) &= E. \end{aligned}$$

From the theory of ordinary differential equations it is known that $\mathfrak{W}$ and $\mathfrak{W}^{-1}$ exist and are analytic on $(\sigma-\epsilon,\ \sigma+T+\epsilon)$ for some $\epsilon > 0$. Since the solution X of (5.3) agrees with $\mathfrak{W}$ on $[\sigma,\sigma+1]$, we have that X is analytic on $[\sigma,\sigma+1]$. Furthermore, we see that

$$(5.4) \quad X(t) = \{X(\sigma+n)\mathfrak{W}^{-1}(\sigma+n) + \int_{\sigma+n}^{t} X(s-1)A_1(s)\mathfrak{W}^{-1}(s)ds\}\mathfrak{W}(t)$$

for $t \in [\sigma+n,\sigma+n+1]$, $n \geq 1$. Hence the analyticity of $\mathfrak{W}$, $\mathfrak{W}^{-1}$, A_1 on $[\sigma+1,\sigma+2]$, and that of X on $[\sigma,\sigma+1]$ imply that X is analytic on $[\sigma+1,\sigma+2]$, which by the same reasoning leads to the analyticity of X on $[\sigma+2,\sigma+3]$. A finite number of repetitions of this reasoning using (5.4) completes the proof.

Remark 5.2. From the definition of the determinant it follows immediately that if $[\alpha,\beta]$ is any interval of analyticity of X (the solution to (5.2) with $A(t,s) \equiv 0$), then either X is singular on $[\alpha,\beta]$ or else there are at most a finite number of points in $[\alpha,\beta]$ where $X^{-1}(t)$ fails to exist.

Just as the step-method fails in existence proofs for (5.2) whenever $A(t,s) \not\equiv 0$, this form of the step-method will not be of use in proving analyticity results for solutions to the general system (5.2). We can, however, obtain the following result by utilization of successive approximations with step-like procedures.

Theorem 5.2. Suppose that $(t,s) \to A(t,s)$ and $t \to A_i(t)$, $i = 0,1,\ldots,K$, are (real) analytic on $[\sigma,\infty) \times [-\tau,0]$ and $[\sigma,\infty)$ respectively into $\mathcal{L}_{nn}$. If the lags $\theta_1,\theta_2,\ldots,\theta_K,\tau$ are commensurate, then the solution to (5.2) is qpwa on $[\sigma,\sigma+T]$ for any $T > 0$.

Again, we shall here give a proof of this theorem only for the special case

$$(5.5) \quad X(t) = X(t)A_0(t) + X(t-1)A_1(t) + \int_{-1}^{0} X(t+s)A(t,s)ds, \quad t \in [0,T],$$

$$X(0) = E, \quad X(t) = 0 \text{ for } t < 0,$$

as it will then be easily seen how one modifies the ideas to obtai the result for commensurate lags on $[\sigma, \sigma+T]$. Since the uniform limit of a sequence of real analytic functions need not be analyti if we wish to use successive approximation techniques to obtain analyticity results, then we must work with complex systems. That is, we must somehow replace (5.5) by a system defined on a domain the complex plane $\mathbb{C}$ which contains $[-1,T]$ so that the system is equivalent to (5.5) on $[-1,T]$. Before beginning the proof we give some preliminary results which will be needed.

Lemma 5.1. If f is analytic in a region $\mathcal{S}(a,b) = \{z = x+iy | a < x < b, -d < y < d\}$ and continuous at $z = a$ from within $\mathcal{S}(a,b)$, then F defined by $F(z) \equiv \int_a^z f(\zeta)d\zeta$ is an analytic function on $\mathcal{S}(a,b)$.

Proof. From the extended form of Cauchy's theorem [W-1] it follows that F is independent of path in $\mathcal{S}(a,b)$ and is thus well defined. For $z \in \mathcal{S}(a,b)$ in a neighborhood of $z_0 \in \mathcal{S}(a,b)$ we have $F(z) = F(z_0) + \int_{z_0}^{z} f(\zeta)d\zeta$ which is analytic at z_0 by well-known results.

Lemma 5.2. Suppose $(t,s) \to \alpha(t,s)$ is real analytic on $(-\epsilon, T+\epsilon) \times (-1-\epsilon, \epsilon) \subset R^2$. Then there are sets $\mathscr{T}_\delta$ and $\mathscr{S}_\delta$ in $\mathbb{C}$ of the form

$$\mathscr{T}_\delta = \{z = x+iy \mid x\in(-\delta, T+\delta),\ y\in(-\delta,\delta)\}$$

$$\mathscr{S}_\delta = \{z = x+iy \mid x\in(-1-\delta, \delta),\ y\in(-\delta,\delta)\}$$

and a function $(z,w) \to \alpha^*(z,w)$ analytic on $\mathscr{T}_\delta \times \mathscr{S}_\delta \subset \mathbb{C}^2$ such that $\alpha^*|K = \alpha$ where $K \equiv \{(x,0) \mid x\in[0,T]\} \times \{(x,0) \mid x\in[-1,0]\}$ (or $K = [0,T] \times [-1,0]$ as a subset of R^2).

Proof. Define $\mathscr{D} = (-\epsilon, T+\epsilon) \times (1-\epsilon, \epsilon)$, which is an open region in R^2 on which α is analytic. It then follows [N-1, p. 5] and [H-8, p. 41-42] that there is an open set $\mathscr{D}^*$ in $\mathbb{C}^2$ with $\mathscr{D}^* \cap R^2 = \mathscr{D}$ and an analytic function α^* on $\mathscr{D}^*$ such that $\alpha^*|\mathscr{D} = \alpha$. The set K defined in the lemma is compact in $\mathbb{C}^2$ and $K \cap (\mathbb{C}^2 - \mathscr{D}^*)$ is empty since $K \cap R^2 \subset \mathscr{D}$. It is then not difficult to show that there is a $\delta > 0$ such that $\mathscr{T}_\delta$ and $\mathscr{S}_\delta$ as defined in the lemma satisfy $K \subset \mathscr{T}_\delta \times \mathscr{S}_\delta \subset \mathscr{D}^*$.

Proof of Theorem 5.2. Our first task is to somehow extend system (5.5) (or, as in the usual case of successive approximations, its equivalent in integral form) to a system on a complex domain where of course we want all coefficients involved to be analytic. From Lemma 5.2 and standard arguments it follows that there exist domains

$\mathcal{S}_\delta$, $\mathcal{T}_\delta$ (see Lemma 5.2) and analytic continuations (which we again denote by A and A_i) of the mappings $(t,s) \to A(t,s)$ and $t \to A_i(t)$ to $\mathcal{T}_\delta \times \mathcal{S}_\delta$ and $\mathcal{T}_\delta$ respectively. Let $\mathcal{D}_\mathcal{S}$ and $\mathcal{D}_\mathcal{T}$ be $\delta/2$ neighborhoods in $\mathbb{C}$ (using the usual norm in $\mathbb{C}$) of the sets $[-1,0]$ and $[0,T]$ respectively. These are the regions on which we shall work throughout the remainder of the proof.

For k any integer, we define $S_k = \{z \in \mathbb{C} \mid k < \mathrm{Re}\ (z) < k+1\} \cap (\mathcal{D}_\mathcal{S} \cup \mathcal{D}_\mathcal{T})$. We shall consider the system defined for $z \in S_k \cup \{k+1\}$, $k \geq 0$ by

$$(5.6) \qquad X(z) = E + \int_{[0,z]} \{X(\zeta)A_0(\zeta) + X(\zeta-1)A_1(\zeta) + \int_{[-1,0]} X(\zeta+w)A(\zeta,w)dw\}d\zeta$$

where we must indicate the paths of integration to be used. The path $[0,z]$ for $z \in S_k \cup \{k+1\}$ consists of straight line segments joining z and k, k and $z-1$, $z-1$ and $k-1, \ldots, z-(k-1)$ and 1, 1 and $z-k$, $z-k$ and 0. Note that the path will always lie in $\mathcal{D}_\mathcal{T}$. The integral $\int_{[-1,0]} X(\zeta+w)A(\zeta,w)dw$, for ζ on the polygonal path joining 0 and z described above and $\zeta \in S_m \cup \{m+1\}$, is to be integrated along the w-path of straight line segments joining 0 and $-\zeta + m$, $-\zeta + m$ and -1. Hence for any $\zeta \in S_m \cup \{m+1\}$ this latter integral depends on the values of X along the segments joing ζ and m, m and $\zeta - 1$. Note that for z real the system (5.6) with the proper initial conditions reduces to the integrated form (in the usual sense) of (5.5).

We next obtain a quasi-slabwise analytic (i.e., analytic on S_k, $k = 0,1,\ldots$) solution to (5.6). To do this we define successive approximations X_n, show that each is analytic on S_k, $k = 0,1,\ldots$, and that $\{X_n\}$ converges uniformly on each S_k. The limit function will be the desired solution. Define for $n = 0,1,2,\ldots, X_n(z) = 0$ for $\operatorname{Re}(z) < 0$ and $X_n(0) = E$. For $z \in S_k \cup \{k+1\}$, $k = 0,1,2,\ldots$ define $X_o(z) = E$ and

$$(5.7) \quad X_n(z) = E + \int_{[0,z]} \{X_{n-1}(\zeta)A_o(\zeta) + X_{n-1}(\zeta-1)A_1(\zeta)$$

$$+ \int_{[-1,0]} X_{n-1}(\zeta+w)A(\zeta,w)dw\}d\zeta$$

for $n = 1,2,\ldots,$ where the paths of integration are the polygonal paths described above. Note that each X_n, $n \geq 1$ is defined on $\mathscr{D}_{\mathscr{S}} \cup \mathscr{D}_{\mathscr{T}}$ less the rays $\bigcup_{k\geq 0} \{z = k+iy \mid y \neq 0\}$.

We shall say that a function g is left continuous at $z = k$, $k \geq 0$, if $g(\zeta) \to g(k)$ as $\zeta \to k$, $\zeta \in S_{k-1}$. A similar meaning is attached to "right continuous at $z = k$". Finally, we shall say that g is continuous at $z = k$ if it is both left and right continuous at $z = k$ in the above sense. We now state and prove an induction lemma which will yield analyticity of X_n on the S_k.

<u>Induction Lemma</u>. Let $n \geq 1$. Let $k \geq 0$. Then X_{n-1} analytic on

$S_{-1}, S_o, S_1, \ldots, S_k$ and continuous at $z = 0,1,2,\ldots,k$ imply X_n analytic on S_k and continuous at $z = k$.

Note. Since clearly none of the approximations are left continuous at $z = 0$, we understand "continuous at $z = 0$" to mean "right continuous at $z = 0$" in the above lemma.

Proof. Suppose the assumptions of the induction lemma are true. We can then establish

Lemma 5.3. Let m be a fixed integer, $0 \leqq m \leqq k$. For $\zeta \in S_m \cup \{m\} \cup \{m+1\}$ define

$$F(\zeta) = \int_{[-1,0]} X_{n-1}(\zeta+w)A(\zeta,w)dw.$$

Then F is analytic on S_m, right continuous at $z = m$, and, if $m < k$, left continuous at $z = m + 1$.

Use of the hypotheses of the induction lemma and Lemma 5.3 yield that the integrand

$$\mathscr{F}(\zeta) = X_{n-1}(\zeta)A_o(\zeta) + X_{n-1}(\zeta-1)A_1(\zeta) + \int_{[-1,0]} X_{n-1}(\zeta+w)A(\zeta,w)dw$$

in (5.7) is analytic on $S_o, S_1, \ldots, S_{k-1}$ and continuous at $z = 0,1,\ldots,k-1$, left continuous at $z = k$. Hence by the extension of Cauchy's theorem (see the proof of Lemma 5.1) the part of the integral in (5.7) from 0 to k along the polygonal paths is actually independent of path (as long as the paths cross the lines $\mathrm{Re}(z) = m$ through the point $z = m$). Thus, (5.7) may be written

$$(5.8) \qquad X_n(z) = E + \int_0^k \mathscr{I}(\zeta)d\zeta + \int_k^z \mathscr{I}(\zeta)d\zeta$$

where as usual $\int_{z_1}^{z_2}$ denotes integration along the straight line segment joining z_1 and z_2. The first integral in (5.8) is now independent of $z \in S_k$. Thus we need only show that the second integral is analytic for $z \in S_k$. But this follows immediately from the hypotheses of the induction lemma, Lemma 5.3, and Lemma 5.1. We therefore have X_n analytic on S_k.

We shall next argue that X_n is right continuous at $z = k$; the arguments for left continuity are not dissimilar and will be omitted. From (5.8) we have $X_n(z) - X_n(k) = \int_k^z \mathscr{I}(\zeta)d\zeta$ for $z \in S_k$, the integrand $\mathscr{I}$ being analytic on S_k and right continuous at $z = k$. Thus $\mathscr{I}$ is bounded in some "right neighborhood" of $z = k$, from which the desired result follows immediately. To complete the proof of the induction lemma it remains only to establish the validity of Lemma 5.3.

Proof of Lemma 5.3. Making the assumptions given in the statement of the induction lemma, we let m be a fixed integer, $0 \leq m \leq k$. Then $F(\zeta)$, $\zeta \in S_m$, can be written

$$F(\zeta) = \int_{-1}^{-\zeta+m} X_{n-1}(\zeta+w)A(\zeta,w)dw + \int_{-\zeta+m}^{0} X_{n-1}(\zeta+w)A(\zeta,w)dw$$

$$= \int_{\zeta-1}^{m} X_{n-1}(w)A(\zeta,w-\zeta)dw + \int_{m}^{\zeta} X_{n-1}(w)A(\zeta,w-\zeta)dw.$$

The right continuity of F at $z = m$ follows from the continuity of A, the boundedness of X_{n-1} in right and left neighborhoods of $z = m$ and a right neighborhood of $z = m - 1$, and the theorem on dominated convergence. For $m < k$ the proof that F is left continuous at $z = m + 1$ is similar. (If $m = k$ these arguments are no longer valid in obtaining left continuity of F at $m + 1$ since at this stage in the induction we do not have that X_{n-1} is left continuous at $k + 1$, which is needed for the boundedness conclusions about X_{n-1}

We turn next to the analyticity arguments for F on S_m. We shall argue that the function f defined by $f(\zeta) = \int_{m}^{\zeta} X_{n-1}(w)A(\zeta,w-\zeta)dw$ is analytic on S_m, similar arguments being valid for the term $\int_{\zeta-1}^{m} X_{n-1}(w)A(\zeta,w-\zeta)dw$ in F above.

Fix $\zeta_0 \in S_m$. For ζ in a sufficiently small neighborhood $\mathcal{N}_0$ of ζ_0 we can write

$$\int_{m}^{\zeta} X_{n-1}(w)A(\zeta,w-\zeta)dw = \int_{m}^{\zeta_0} X_{n-1}(w)A(\zeta,w-\zeta)dw$$
$$+ \int_{\zeta_0}^{\zeta} X_{n-1}(w)A(\zeta,w-\zeta)dw \equiv h_1(\zeta) + h_2(\zeta)$$

where the integrands are analytic in ζ on $\mathcal{N}_0$ for each fixed

$w \in \mathscr{N}_o$ and analytic in w on $\mathscr{N}_o$ for each fixed $\zeta \in \mathscr{N}_o$. A straightforward application of Morera's theorem establishes the analyticity of h_1 on $\mathscr{N}_o$. Use of a theorem of Hartogs-Osgood [H-9, p. 28] yields that h_2 is of the form $\int_{\zeta_o}^{\zeta} g(w,\zeta)dw$ where g is analytic in $\mathscr{N}_o \times \mathscr{N}_o$, from which the analyticity of h_2 follows easily.

Having confirmed the validity of the induction lemma, we point out that it follows directly from the analyticity properties of X_o (recall $X_o(z) = E$ for $z \in S_k \cup \{k+1\}$, $k = 0,1,\ldots$ and $X_o(z) = 0$ for $\operatorname{Re}(z) < 0$) and the induction lemma that each X_n is analytic on each S_k.

We next prove that the sequence $\{X_n\}$ converges uniformly on the region of interest. Let ℓ be the positive integer $(\ell > T)$ such that $S_k = \emptyset$ for $k \geq \ell$ and $S_{\ell-1} \neq \emptyset$. (Recall the definition of S_k, $\mathscr{D}_{\mathscr{S}}$, and $\mathscr{D}_{\mathscr{T}}$.) We shall show that the sequence $\{X_n\}$ converges uniformly on $\mathfrak{R} \equiv \bigcup_{k=o}^{\ell-1} S_k \cup \{k\}$. We note that we trivially have uniform convergence of $\{X_n\}$ on $S_{-1} \cup \{0\}$ to the function X defined by $X(z) = 0$ for $z \in S_{-1}$, $X(0) = E$. Recall now the definition of X_n given in (5.7) and the integration paths employed. For any $z \in \mathfrak{R}$ let $s(z)$ denote the <u>arclength</u> of the polygonal path described above (see (5.6)) which joins 0 to z. Let M be a bound for $|A_o(\zeta)|$, $|A_1(\zeta)|$, $\zeta \in \mathscr{D}_{\mathscr{T}}$ and $|A(\zeta,w)|$, $(\zeta,w) \in \mathscr{D}_{\mathscr{T}} \times \mathscr{D}_{\mathscr{S}}$. Then for $z \in \mathfrak{R}$ we have

$$|X_1(z) - X_o(z)| = \Big|\int_{[0,z]} \{X_o(\zeta)A_o(\zeta) + X_o(\zeta-1)A_1(\zeta)$$

$$+ \int_{[-1,0]} X_o(\zeta+w)A(\zeta,w)dw\}d\zeta\Big|$$

$$\leq \int_{[0,z]} \{|A_o(\zeta)| + |A_1(\zeta)| + \int_{[-1,0]} |A(\zeta,w)|\,|dw|\}\,|d\zeta|$$

$$\leq \int_{[0,z]} \{M + M + M(1+2(\tfrac{\delta}{2}))\}\,|d\zeta|$$

$$\leq 3M(1+\delta)s(z) \equiv \rho s(z).$$

Furthermore,

$$|X_2(z) - X_1(z)| \leq \int_{[0,z]} \{M|X_1(\zeta) - X_o(\zeta)| + M|X_1(\zeta-1) - X_o(\zeta-1)|$$

$$+ \int_{[-1,0]} M|X_1(\zeta+w) - X_o(\zeta+w)|\;|dw|\}\,|d\zeta|.$$

For $\zeta \in S_o$ we have

$$\int_{[-1,0]} M|X_1(\zeta+w) - X_o(\zeta+w)|\;|dw| = \int_{-\zeta}^{o} M|X_1(\zeta+w) - X_o(\zeta+w)|\;|dw|$$

$$\leq \int_{-\zeta}^{o} M\rho s(\zeta+w)|dw| \leq \int_{-\zeta}^{o} M\rho s(\zeta)|dw|$$

$$\leq M\rho s(\zeta)(1+\delta).$$

For $\zeta \in S_k$, $k \geq 1$ we find that

$$(5.9) \qquad s(\zeta+w) \leq s(\zeta)$$

for any w lying on the path consisting of straight line segments joing -1 to $-\zeta + k$ and $-\zeta + k$ to 0. Hence the above estimate is also valid for these values of ζ. It follows that

$$\begin{aligned} |X_2(z) - X_1(z)| &\leq \int_{[0,z]} \{M\rho s(\zeta) + M\rho s(\zeta) + M\rho s(\zeta)(1+\delta)\} |d\zeta| \\ &\leq 3M(1+\delta)\rho \int_{[0,z]} s(\zeta)|d\zeta| \\ &= \rho^2 \frac{[s(z)]^2}{2} . \end{aligned}$$

Using this estimate and the above ideas it is easily shown that

$$|X_3(z) - X_2(z)| \leq \rho^3 \frac{[s(z)]^3}{3!}$$

and in general

$$|X_n(z) - X_{n-1}(z)| \leq \rho^n \frac{[s(z)]^n}{n!}$$

for $z \in \Re$. Hence for $n > m$ we have

$$(5.10) \quad |X_n(z) - X_m(z)| \leq \sum_{j=m+1}^{n} |X_j(z) - X_{j-1}(z)| \leq \sum_{j=m+1}^{n} \rho^j \frac{[s(z)]^j}{j!}$$

But for $z \in \Re$ we have that $s(z) \leq \ell(1+2(\delta/2)) = \ell(1+\delta)$. Using this with (5.10) yields the uniform convergence of $\{X_n\}$ on $\Re$. Let us denote by X this limit function on $\Re \cup S_{-1}$. Since each X_n is analytic on S_k and continuous at $z = k$, we have that X also possesses these properties. Furthermore, for each n, $X_n(z)$ is real valued whenever z is real, from which it follows that X is real analytic on $(0,1)$, $(1,2)$, etc. Finally, since $X_n \rightrightarrows X$ on $[-1,T]$ it is not difficult to argue that X is the unique solution to (5.5), which completes the proof of Theorem 5.2.

One might reasonably expect a stronger type of analyticity (say pwa) than that obtained in Theorem 5.2 to be true for systems of the type (5.2) even with $A \neq 0$. The authors have tried unsuccessfully so far to obtain these stronger results. Several ideas using different integration paths in defining the successive approximations (see the proof of Theorem 5.2) and stronger assumptions on the coefficients have been tried. These lead to either a lack of analyticity of the estimates in the desired regions, or else an inability to obtain uniform convergence of the estimates. The authors were able to prove that the solution to (5.5) is analytic on $[0,1]$, but could not adapt these methods to prove analyticity on $[k,k+1]$ for $k > 0$. The fact that one is using a zero initial matrix on $[-1,0)$ appears to be essential in obtaining analyticity on $[0,1]$. (Note that in this case the system loses some of its lag behavior on $[0,1]$ and is much more like an integral equation.)

The analyticity results obtained in this section can be used to study the zeros of the multipliers in the maximum principle for control problems involving functional differential equations (see Remark 5.2 and [B-2, B-3, B-4, H-2]. The information thus obtained can be especially useful when the maximum principle is also a sufficient condition for optimality (see [H-2]). Another application of these analyticity results is discussed in the next section.

§6. Application of the Analyticity Results. We shall use pwc as an abbreviation for piecewise continuous, and when we say a function $f: [a,b] \to R^p$ is piecewise continuous we are taking the standard definition. We shall say that f is almost piecewise continuous (apwc) if there is a finite number of points $s_i \in [a,b]$, $i = 0,1,\ldots,N$ with the property that $f|[\alpha,\beta]$ is pwc for every $[\alpha,\beta] \subset [a,b]$ for which $s_i \notin [\alpha,\beta]$, $i = 0,1,\ldots,N$.

In this section we shall demonstrate how some of the work with subintegrals of multifunctions by Halkin and Hendricks [H-6] and the related existence theory for piecewise continuous optimal controls [H-5] can be applied in special cases of (2.3) to give analogs of Theorem 3.2 when the admissible triples $\{\varphi,u,t\}$ are required to be pwc (or apwc) (i.e., φ and u are pwc (or apwc)).

Lebesgue measure will be understood in all of the integrals appearing in this section. Suppose a multifunction $H: [a,b] \to R^p$ is given. Then we have defined $\int_a^b H(t)dt$ and with $[a,b]$ understood we denote this by $\int H$. We define

$$\int^* H \doteq \{\int_a^b g(t)dt \,|\, g: [a,b] \to R^p \text{ is apwc and } g(t) \in H(t),\ t \in [a,b]\}.$$

Lemma 6.1 (Halkin-Hendricks). $\int^* H$ is convex.

We omit the proof. Let it suffice to say that the proof of Theorem 1 [H-6, pg. 365] may in effect be repeated. One need only take f_1 and f_2 to be apwc in that proof and observe that $f_1 \cdot \chi_{[a,b]\setminus E} + f_2 \cdot \chi_E$ is apwc if $E \subset [a,b]$ is the union of a finite number of intervals.

A set $E \subset R^q$ is said to be semianalytic (see Lojasiewicz

[L-7] or Halkin and Hendricks [H-6]) if for every point in R^q there exists a neighborhood V of that point such that

$$E \cap V = \bigcup_{i=1}^{k} \{x \in R^q \mid f_i(x) = 0 \text{ and } g_{ij}(x) > 0 \text{ for } j = 1,2,\ldots,\ell\}$$

where g_{ij} and f_i are real valued functions which are analytic on V.

Lemma 6.2. (Halkin-Hendricks). Let $H\colon [a,b] \to R^p$ be a compact multifunction and suppose the graph of H is bounded. Let there exist a finite set of points s_i, $i = 0,1,\ldots,N$ such that $a \leq s_0 < s_1 < \ldots < s_N \leq b$ and such that for each compact interval $[\alpha,\beta] \subset [a,b]$ which contains none of the points s_i the graph of H restricted to $[\alpha,\beta]$ is semianalytic. Then $\int H = \int^* H$.

Again this is only a slight extension of the main result (Theorem 2) in [H-6]. Indeed, the proof is clear upon examining the proof of that theorem. In effect one observes that $\int H \supset \int^* H \supset \mathscr{E}(\int H)$ where $\mathscr{E}(\int H)$ denotes the set of extreme points of the convex set $\int H$, and that $\int^* H$ is convex, and then the proof is immediate. To show that $\int^* H \supset \mathscr{E}(\int H)$ one need only show that the function $g\colon [a,b] \to R^p$ is apwc, where g is defined by the condition that $g(t)$ is the lexicographic maximum (with respect to an arbitrary orthonormal basis for R^p as in Olech [O-1]) of $H(t)$, $t \in [a,b]$. If the s_i, $i = 0,1,\ldots,N$ and an interval $[\alpha,\beta]$ are chosen as in the hypotheses of Lemma 6.2, then Halkin and Hendricks [H-6] have shown $g|[\alpha,\beta]$ is pwc. Hence $g\colon [a,b] \to R^p$ is apwc.

Lemma 6.3 (Halkin). Let $B\colon [a,b] \times R^q \to R^p$ be a continuous function with the property that there is a finite set of points s_i, $i = 0,1,\ldots,N$, such that $a \leqq s_0 < s_1 < \ldots < s_N \leqq b$ and such that $B|(a,s_0) \times R^q$, $B|(s_N,b) \times R^q$, and $B|(s_{i-1},s_i) \times R^q$, $i = 1,2,\ldots,N$ are analytic. Define $\mathscr{B}\colon [a,b] \times R^q \to R^{p+q}$ by the relation $\mathscr{B}(t,u) = (u,B(t,u))$. Let $\Omega\colon [a,b] \to R^q$ be a compact multifunction satisfying the generic hypotheses of Lemma 6.2. Then the compact multifunction $\mathscr{W}\colon [a,b] \to R^{p+q}$ defined by $\mathscr{W}(t) = \mathscr{B}(t,\Omega(t))$, $t \in [a,b]$ also satisfies the generic hypotheses of Lemma 6.2.

This result is a modification of a statement of Halkin's [H-5]. Since Halkin omitted a proof and since the above lemma differs somewhat from his result, we shall suggest a proof which is straightforward. There will be no loss in generality if we assume that the same points satisfy the hypothesis of Lemma 6.2 with $H\colon [a,b] \to R^p$ replaced by $\Omega\colon [a,b] \to R^q$. It will suffice for us to show that if $[\alpha,\beta] \subset [a,b]$ and $s_i \notin [\alpha,\beta]$, $i = 0,1,\ldots,N$, then $\mathscr{W}|[\alpha,\beta]$ has a semianalytic graph. Let $\mathscr{G}$ denote the graph of $\mathscr{W}|[\alpha,\beta]$ and let $P_0 = (t_0,u_0,x_0)$ be an arbitrary point in $R \times R^q \times R^p = R^{p+q+1}$. Then there is a neighborhood V_0 of (t_0,u_0) in $R \times R^q = R^{q+1}$ and analytic functions f_i, g_{ij}, $i = 1,\ldots,k$, $j = 1,\ldots,\ell$ on V_0 such that

$$E \cap V_0 = \bigcup_{i=1}^{k} \{(t,u) \in R^{q+1} | f_i(t,u) = 0 \text{ and } g_{ij}(t,u) > 0, \text{ for } j = 1,\ldots,\ell\}$$

where E is the graph of $\Omega|[\alpha,\beta]$. Let $\mathfrak{B}$ denote the set $\{(t,u,x) \in R^{p+q+1} | (t,u) \in V_o\}$, then $\mathfrak{B}$ is a neighborhood of P_o in R^{p+q+1}. A generic point (t,u,x) in R^{p+q+1} is also denoted by $(t,u^1,\ldots,u^q,x^1,\ldots,x^p)$ and we write $(B^1,\ldots,B^p)$ for the function B. Let $\pi: R^{p+q+1} \to R^{q+1}$ be defined by $\pi(t,u,x) = (t,u)$. Define functions $\bar{f}_i$ and $\bar{g}_{ij}$ on $\mathfrak{B}$ by the equations

$$\bar{f}_i(t,u,x) = [f_i(\pi(t,u,x))]^2 + \sum_{n=1}^{p} [B^n(t,u) - x^n]^2$$

$$\bar{g}_{ij}(t,u,x) = g_{ij}(\pi(t,u,x))$$

for $i = 1,\ldots,k$, $j = 1,\ldots,\ell$. Then we have that there are real numbers $\bar{\alpha},\bar{\beta}$ such that $(\bar{\alpha},\bar{\beta}) \supset [\alpha,\beta]$ and such that B is analytic on $(\bar{\alpha},\bar{\beta}) \times R^q$. One can now verify that

$$\mathscr{G} \cap \mathfrak{B} \cap ((\bar{\alpha},\bar{\beta}) \times R^{p+q}) = \bigcup_{i=1} \{(t,u,x) \in R^{p+q+1} | \bar{f}_i(t,u,x) = 0 \text{ and } \bar{g}_{ij}(t,u,x) > 0 \text{ for } j = 1,2,\ldots,\ell\}$$

where the functions $\bar{f}_i$ and $\bar{g}_{ij}$ are analytic on $\mathfrak{B} \cap ((\bar{\alpha},\bar{\beta}) \times R^{p+q})$. We can assume $P_o \in (\bar{\alpha},\bar{\beta}) \times R^{p+q}$ since the contrary case can be dealt with trivialty. Thus $\mathscr{G}$ is semianalytic and this proves the lemma.

We now turn our attention to the control system (5.1). Let $\mathscr{A}_t^*(\Phi,U)$ denote the collection of all points in $\mathscr{A}_t(\Phi,U)$ which are attainable from admissible triples $\{\varphi,u,t\}$ where

$\varphi\colon [-\tau,0] \to R^n$ and $u\colon [0,t] \to R^m$ are apwc. Let $\underline{\mathscr{A}}_t(\Phi,U)$ denote the collection of all points in $\mathscr{A}_t(\Phi,U)$ which are attainable from admissible triples $\{\varphi,u,t\}$ where both φ and u are pwc. The variation of parameters formula* (2.7) when applied to the FDE (5.1) gives

$$(6.1) \qquad x(t,\varphi,u) = \varphi(0)Y(0,t) + \int_0^t h(u(\alpha),\alpha)Y(\alpha,t)d\alpha + \int_{-\tau}^0 \varphi(\alpha)\,\mathfrak{K}(\alpha,t)d\alpha$$

where $\mathfrak{K}(\alpha,t)$ is defined by the equation

$$\mathfrak{K}(\alpha,t) \equiv \sum_{j=1}^{k} A_i(\alpha+\theta_i)Y(\alpha+\theta_i,t)\chi_{[-\theta_i,0]}(\alpha)$$

$$+ \int_0^{\alpha+\tau} A(s,\alpha-s)Y(s,t)ds, \quad -\tau \leqq \alpha \leqq 0.$$

Let a function $\mathfrak{M}\colon [-\tau,0] \times R^n \to R^{2n}$ be defined by

$$\mathfrak{M}(\alpha,\varphi) \equiv (\varphi,\varphi\,\mathfrak{K}(\alpha,t)), \qquad -\tau \leqq \alpha \leqq 0, \quad \varphi \in R^n,$$

and let $\mathfrak{H}\colon [0,t] \times R^m \to R^{m+n}$ be the function defined by

$$\mathfrak{H}(u,\alpha) \equiv (u,h(u,\alpha)Y(\alpha,t)), \quad 0 \leqq \alpha \leqq t, \quad u \in R^m.$$

---

*We remark that the representation theorems can easily be shown to be valid under the analyticity hypotheses placed on (5.1) in section 5 (Theorems 5.1, 5.2).

Define projections $\pi_1\colon R^{m+n} \to R^n$ and $\pi_2\colon R^{2n} \to R^n$ by the equations

$$\pi_1(u,x) = x, \qquad (u,x) \in R^m \times R^n = R^{m+n}$$

$$\pi_2(\varphi,x) = x, \qquad (\varphi,x) \in R^n \times R^n = R^{2n}$$

In each of the following three formulas the first integral on the right hand side of the equation is over the interval $[0,t]$ and the second integral is over the interval $[-\tau,0]$. Using (6.1) it can be shown that

$$(6.2) \qquad \mathscr{A}_t(\Phi,U) = \Phi(0)Y(0,t) + \pi_1[\int \mathfrak{H}(U(\alpha),\alpha)d\alpha] + \pi_2[\int \mathfrak{M}(\alpha,\Phi(\alpha))d\alpha],$$

$$\mathscr{A}_t^*(\Phi,U) = \Phi(0)Y(0,t) + \pi_1[\int^* \mathfrak{H}(U(\alpha),\alpha)d\alpha] + \pi_2[\int^* \mathfrak{M}(\alpha,\Phi(\alpha))d\alpha],$$

and*

$$\underline{\mathscr{A}}_t(\Phi,U) = \Phi(0)Y(0,t) + \pi_1[\underline{\int} \mathfrak{H}(U(\alpha),\alpha)d\alpha] + \pi_2[\underline{\int} \mathfrak{M}(\alpha,\Phi(\alpha))d\alpha],$$

whenever the left hand sides are nonempty.

<u>Theorem 6.1.</u> Let the homogeneous part of (5.1) satisfy the hypotheses of Theorem 5.2, and let the functions $h|R^m \times [0,t]$; $t \geq 0$ satisfy the same conditions as the function B in Lemma 6.3. Let $\Phi\colon [-\tau,0] \to R^n$ and $U\colon [0,\infty) \to R^m$ be compact multifunctions satisfying the

---

*Here $\underline{\int}$ denotes the subintegral in Halkin and Hendricks [H-6], i.e., if $H\colon [a,b] \to R^p$ is some multifunction, then $\underline{\int} H \equiv \{\int_a^b g(t)dt \mid g\colon [a,b] \to R^p$ is pwc and $g(t) \in H(t)$, for $t \in [a,b]\}$.

generic hypotheses of Lemma 6.2. Then

$$\mathscr{A}_t(\Phi,U) = \mathscr{A}_t^*(\Phi,U), \qquad t \geq 0.$$

Proof. With the aid of Theorem 5.2 and a few rudimentary deductions. one can show that the function $\mathfrak{H}$: $[0,t] \times R^m \to R^{m+n}$ and the multifunction U: $[0,t] \to R^m$ satisfy the generic hypotheses of Lemma 6.3, and similarly for the function $\mathfrak{M}$: $[-\tau,0] \times R^n \to R^{2n}$ and the multifunction Φ: $[-\tau,0] \to R^n$. Thus Lemmas 6.3 and 6.2 and the second formula in relation (6.2) apply to give the derived conclusion.

Theorem 6.2. Let the homogeneous part of (5.1) satisfy the hypotheses of Theorem 5.1 and let the function h be analytic on $R^m \times [0,\infty)$. Let Φ: $[-\tau,0] \to R^n$ and U: $[0,\infty) \to R^m$ be compact multifunctions such that the graph of Φ and the graph of $U|[0,t]$ for $t \geq 0$ are bounded and semianalytic. Then

$$\mathscr{A}_t(\Phi,U) = \underline{\mathscr{A}}_t(\Phi,U), \qquad t \geq 0.$$

With the aid of Theorem 5.1 and the above remarks the proof of this theorem will be so similar to Halkin's proof [H-5] of the corresponding result for nondelay systems that it can safely be omitted.

Recalling Remark 3.3 one sees that Theorems 6.1 and 6.2 give new existence theorems for certain optimal control problems in the class of apwc admissible triples $\{\varphi,u,t\}$ and the class of pwc admissible triples $\{\varphi,u,t\}$ respectively.

Bibliography

[A-1] P. Alexandroff and H. Hopf, "Topologie", erster Band. Springer-Verlag, Berlin, 1935.

[A-2] R. J. Aumann, Integrals of set-valued functions, J. Math. Anal. Appl. 12(1965), 1-12.

[B-1] H. T. Banks, Representations for solutions of linear functional differential equations, J. Differential Equations 5(1969), 399-409.

[B-2] __________, Variational problems involving functional differential equations, SIAM J. Control 7(1969), 1-17.

[B-3] __________, A maximum principle for optimal control problems with functional differential systems, Bull. Amer. Math. Soc. 75(1969), 158-161.

[B-4] __________, Necessary conditions for control problems with variable time lags, SIAM J. Control 6(1968), 9-47.

[B-5] R. Bellman and K. L. Cooke, "Differential Difference Equations", Academic Press, New York, 1963.

[B-6] D. Blackwell, The range of certain vector integrals, Proc. Amer. Math. Soc. 2(1951), 390-395.

[C-1] R. H. Cameron and W. T. Martin, An unsymmetric Fubini theorem, Bull. Amer. Math. Soc. 47(1941), 121-125.

[C-2] C. Castaing, "Sur les multi-applications measurables", Doctoral Dissertation. Universite de Caen. France, 1967.

[C-3] __________, Sur les multi applications mesurables, Rev. Francaise Inf. Rech. Oper. 1(1967), 91-126.

[C-4] L. Cesari, Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints, I, Trans. Amer. Math. Soc. 124(1966), 369-412.

[C-5] D. H. Chyung and E. B. Lee, Optimal systems with time delays, Proc. 3rd IFAC Conf. London, 1966.

[D-1] K. S. Day and T. C. Hsai, Optimal control of linear time-lag systems, Proceedings JACC, Ann Arbor, Michigan, 1968, 1046-1055.

[D-2] G. Debreu, Integration of Correspondences, Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1966, 351-372.

[D-3] N. Dinculeanu, "Vector Measures", Pergamon, New York, 1967.

[D-4] N. Dunford and J. Schwartz, "Linear Operators", Part I, Interscience, New York, 1958.

[E-1] L. E. El'sgol'ts, "Introduction to the Theory of Differential Equations with Deviating Arguments", Holden-Day, San Francisco, 1966.

[G-1] W. C. Grimmell, The existence of piecewise-continuous fuel optimal controls, SIAM J. Control 5(1967), 515-519.

[H-1] A. Halanay, "Differential Equations", Academic Press, New York, 1966.

[H-2] __________, Le "principe du maximum" pour les systemes optimaux lineaires a rétardement, C. R. Acad. Sci. Paris 254(1962), 2277-2279.

[H-3] __________, Optimal controls for systems with time lag, SIAM J. Control 6(1968), 215-234.

[H-4] H. Halkin, A generalization of LaSalle's "Bang-Bang" principle, SIAM J. Control 2(1964), 199-202.

[H-5] __________, A new existence theorem in the class of piecewise continuous control functions, To appear in "Calculus of Variations and Control Theory", edited by A. V. Balakrishnan, Academic Press, 1969.

[H-6] H. Halkin and E. C. Hendricks, Subintegrals of set-valued functions with semianalytic graphs, Proc. Nat. Acad. Sci. U.S.A. 59(1968), 365-367.

[H-7] H. Hermes, Calculus of set-valued functions and control, J. Math. Mech. 18(1968), 47-60.

[H-8] M. Hervé, "Several Complex Variable-Local Theory", Oxford University Press, 1963.

[H-9] L. Hörmander, "An Introduction to Complex Analysis in Several Variables", Van Nostrand, New York. 1966.

[H-10] C. J. Himmelberg, M. Q. Jacobs, and F. S. VanVleck, Measurable multifunctions, selectors, and Filippov's implicit functions lemma, J. Math. Anal. Appl. 25(1969), 276-284.

[H-11] D. K. Hughes, Variational and optimal control problems with delayed argument, J. Optimization Theory and Applications 2(1968), 1-14.

[J-1] M. Q. Jacobs, Measurable multivalued mappings and Lusin's theorem, Trans. Amer. Math. Soc. 134(1968), 471-481.

[J-2] __________, Remarks on some recent extensions of Filippov's implicit functions lemma, SIAM J. Control 5 (1967), 622-627.

[J-3] __________, On the approximation of integrals of multivalued functions, SIAM J. Control 7(1969), 158-177.

[K-1] G. L. Kharatishvili, "Extremal Problems in Linear Topological Spaces", Doctoral Dissertation, Tbilisi State University, USSR, 1968.

[K-2] __________, A maximal principle in extremal problems with delays, Mathematical Theory of Control, edited by A. V. Balakrishnan and L. W. Neustadt. Academic Press, New York, 1967.

[K-3] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, Bull. Acad. Polon. Sci. Math. Astronom. Phys. 13(1965), 397-403.

[K-4] K. Kuratowski, Les fonctions semi-continues dans l'espace des ensemble fermés, Fund. Math. 18(1932), 148-160.

[L-1] E. B. Lee, Variational problems for systems having delay in the control action, IEEE Trans. AC-13(1968), 697-699.

[L-2] __________, Geometric theory of linear controlled systems, to appear.

[L-3] E. B. Lee and L. Markus, "Foundations of Optimal Control Theory", John Wiley and Sons, Inc., New York, 1967.

[L-4] N. Levinson, Minimax, Liapunov and "Bang-Bang", J. Diff. Equations 2(1966), 218-241.

[L-5] W. Li, "Mathematical Models in the Biological Sciences", Masters Thesis, Brown University, Providence, R.I., 1970.

[L-6] A. A. Liapunov, Sur les fonctions-vecteurs complètement additives, Izv. Akad. Nauk SSSR, Ser. Mat. 8(1940), 465-478.

[L-7] S. Lojasiewicz, "Ensembles Semi-Analytiques", Lecture Notes from the Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette, 1965.

[N-1] R. Narasimhan, "Introduction to the Theory of Analytic Spaces", Springer-Verlag Lecture Notes in Mathematics. Berlin-Heidelberg, 1966.

[N-2] L. W. Neustadt, The existence of optimal controls in the absence of convexity conditions, J. Math. Anal. Appl. 7 (1963), 110-117.

[O-1] C. Olech, Extremal solutions to a control system, J. Diff. Eqs. 2(1966), 74-101.

[O-2] M. N. Oguztöreli, "Time-Lag Control Systems", Academic Press, New York, 1966.

[P-1] E. Pinney, "Ordinary Difference-Differential Equations", University of California Press, 1958.

[R-1] H. L. Royden, "Real Analysis", 2nd Edition, MacMillan Co., New York, 1968.

[V-1] V. Volterra, "Theory of Functionals", Blackie, London, 1930.

[W-1] J. L. Walsh, The Cauchy-Goursat theorem for rectifiable Jordan curves, Proc. Nat. Acad. Sci. USA, 19(1933), 540-541.

[Z-1] A. M. Zverkin, G. A. Kemenskii, S. B. Norkin, and L. E. El'sgol'ts, Differential equations with a perturbed argument, I and II. Russian Math. Surveys 17(1962), 61-146 and Trudy Sem. Teor. Differencial. Uravnenii s Otklon. Argumentom 2(1963), 3-49.