Calculus of Set Valued Functions and Control

HENRY HERMES†

Communicated by JURGIN MOSER

Introduction. A beginning of what might be called a calculus of set valued functions can be found in [1]. Many references pertaining to this field are motivated by the mathematical theory of economics with problems of a similar nature appearing in the study of generalized differential equations having the form $\dot{x}(t) \in R(t, x(t))$, ($\dot{x}(t) = d/dt x(t)$) where R is a set valued function. Such equations, commonly called contingent equations, arise naturally from control systems.

Following the notation of [1], if P is a function defined on a real interval $[0, T]$ with values $P(t)$ nonempty subsets of Euclidean n space, E^n , we define

$$\int_0^T P(\tau) d\tau = \left\{ \int_0^T p(\tau) d\tau : p \text{ is measurable, } p(\tau) \in P(\tau) \right\}.$$

We shall use measurable to mean Lebesgue measurable and prefix the word when it is used in a different sense.

It is always true that $\int_0^T P(\tau) d\tau$ is convex. Furthermore, as shown in [1], if the graph of P is a Borel subset when considered as a subset of $[0, T] \times E^n$, then $\int_0^T P(\tau) d\tau$ is nonempty. In this case P is termed Borel measurable. If P is Borel measurable with values subsets of a fixed compact subset of E^n , $\int_0^T P(\tau) d\tau = \int_0^T \text{co } P(\tau) d\tau$ where $\text{co } P$ is the set valued function with values $\text{co } P(t)$ the convex hull of $P(t)$. If, in addition, the values of P are closed subsets, $\int_0^T P(\tau) d\tau$ is compact. These results will find constant use in this paper; they generalize previous work in [4] and [10].

If P is a Borel measurable set valued function with values nonempty, compact subsets of a given compact set in E^n and $\mathfrak{A}(t) = \int_0^t P(\tau) d\tau$, we shall call $\mathfrak{A}$ the integral of P . In Section 1 we use a result of Plíš, [8], on a Lusin theorem for measurable set valued functions to show that $\lim_{h \rightarrow 0} (1/h) \int_{t-h}^{t+h} P(\tau) d\tau = \text{co } P(t)$ for almost all $t \in [0, T]$. From this and Theorem 3 of [1], it can be shown that a necessary and sufficient condition that $\int_0^t P^1(\tau) d\tau = \int_0^t P^2(\tau) d\tau$ for all $t \in [0, T]$ is that $\text{co } P^1(t) = \text{co } P^2(t)$ almost everywhere. Thus if $\mathfrak{A}$ is the integral

† This research was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under AFOSR grant Nr. 1243-67.

of P one could say P is the derivative of $\mathfrak{A}$ and the derivative is unique within the class of Borel measurable, convex set valued functions.

In Section 2, the question of when a set valued function $\mathfrak{A}$ is an integral is studied. The main result, Theorem 2.1, is that the uniform limit (in the Hausdorff topology) of integrals is an integral.

We define a *contingent* R to be a map from $[0, T] \times E^n$ into the space of nonempty, compact subsets of E^n . With R we associate the contingent equation

$$(0-1) \quad \dot{x}(t) \in R(t, x(t)), \quad x(0) = x^0.$$

A solution of (0-1) is any absolutely continuous function φ such that $\dot{\varphi}(t) \in R(t, \varphi(t))$ for almost all t in some neighborhood of zero and $\varphi(0) = x^0$. The attainable set for (0-1) at time $t \in [0, T]$ will be denoted by $\mathfrak{A}(t)$ and is defined as

$$\mathfrak{A}(t) = \{\varphi(t) : \varphi \text{ is a solution of (0-1)}\}.$$

A linear control system has the form

$$(0-2) \quad \dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(0) = x^0$$

where x is an n vector, A and B are $n \times n$ and $n \times r$ matrix valued functions, respectively, and the control u may take values in some given compact set $U \subset E^r$. With (0-2) we may associate the set valued function $Q(t, x) = A(t)x + P(t)$ where $P(t) = \{B(t)u : u \in U\}$ and reformulate (0-2) as the contingent equation $\dot{x}(t) \in Q(t, x(t))$, $x(0) = x^0$. This motivates the definition: a contingent Q is *linear* if it can be written in the form $Q(t, x) = A(t)x + P(t)$ for A with components in $\mathcal{L}_\infty[0, T]$ and P a Borel measurable, compact (possibly empty) set valued function. If in addition the components of A are essentially bounded by K and the values of P lie in a compact, origin centered ball of radius K in E^n , the linear contingent will be said to be *K bounded*.

Partially order contingents by $R^1 \leq R^2$ if $\mathfrak{A}^1(t) \subset \mathfrak{A}^2(t)$ for all $t \in [0, T]$ where $\mathfrak{A}^i$ is the attainable set associated with R^i . Let R be a given contingent, continuous in the Hausdorff topology, and such that there exists a $c > 0$ with $\max \{x \cdot r : r \in R(t, x)\} \leq c(1 + |x|^2)$ for all x and $t \in [0, T]$. This assures us that if a solution φ of the associated contingent equation exists, then $|\varphi(t)|^2 \leq (1 + |x^0|^2) \exp 2ct$. Let $\mathcal{O}_K$ denote the partially ordered set of K bounded linear contingents $Q \geq R$. Using the results of the previous sections, in Section 3 we obtain

Theorem 3.1. *For K sufficiently large, $\mathcal{O}_K$ is nonempty and has a minimal element.*

Notationally, $\mathfrak{N}(\epsilon, A)$ will denote an $\epsilon > 0$ neighborhood of a set $A \subset E^n$. For two sets A, B in E^n , the set sum $A + B = \{a + b : a \in A, b \in B\}$, while $A - B$ consists of those points of A which are not in B . If the set A is clear from the context we will use B' to denote $A - B$. If P is a subset of E^n and X an $n \times n$ matrix, $XP = \{Xp : p \in P\}$. The unit sphere of E^n , i.e. $\{y \in E^n : |y| = 1\}$, will be denoted by S^{n-1} and μ will denote Lebesgue measure on the line.

1. Calculus for set valued functions. The collection of all nonempty compact subsets of E^n can be made into a metric space by defining the distance $\rho(A, B)$ between two elements A, B to be the smallest real number d so that A is contained in a d neighborhood of B and B in a d neighborhood of A . The metric induced topology is called the *Hausdorff topology*. Many properties of this and other interesting topologies for spaces of subsets may be found in [6]. In general we will deal with functions which need not be continuous in the Hausdorff topology. Throughout, P will be used to denote a set valued function, defined on the real interval $[0, T]$, $T > 0$, with values nonempty closed subsets of a compact ball $B_K \subset E^n$, where B_K is origin centered and of radius K .

Definition ([1]). *P is Borel measurable if $\{(t, x) : t \in [0, T], x \in P(t)\}$ is a Borel subset of E^{n+1} . Intuitively this means that the graph of P considered as a subset of E^{n+1} is a Borel subset.*

Definition ([3], [8]). *P is measurable if for any closed set $D \subset E^n$, $\{t \in [0, T] : P(t) \cap D \neq \emptyset\}$ is Lebesgue measurable. ($\emptyset$ denotes the empty set.)*

Letting π_1 denote projection of E^{n+1} onto its first coordinate, the latter definition can be restated as : P is measurable if $\pi_1([0, T] \times D \cap \{(t, x) : t \in [0, T], x \in P(t)\})$ is Lebesgue measurable for every closed set $D \subset E^n$. Since π_1 is continuous and the continuous image of a Borel set (termed an analytic set) is Lebesgue measurable ([5] pg. 482) we have: *If P is Borel measurable it is measurable.*

If P is either Borel measurable, or measurable, it contains a measurable point valued function, i.e. there is a measurable function $p : [0, T] \rightarrow E^n$ with $p(t) \in P(t)$ for $t \in [0, T]$. (See, respectively, [1] Prop. 2.1, and [3] pg. 942.)

Pliś, [8], has given the following generalization of Lusin's theorem.

Theorem (PLIŚ). *Let P be a measurable set valued function, defined on $[0, T]$, with values nonempty closed subsets of a compact metric space X . Then given any $\epsilon > 0$ there exists a closed set $B \subset [0, T]$, with measure differing from T by less than ϵ , on which P is continuous (in the Hausdorff topology).*

If the set B in the statement of the above theorem were open, one would have that for every $t \in B$ and $\epsilon > 0$ there is an $h > 0$ such that $\rho(P(\tau), P(t)) < \epsilon$ if $|t - \tau| < h$. A condition of this type will be essential in computing $\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} P(\tau) d\tau$ but for B closed, this condition is not possible even for point valued functions. We proceed, as in the case of real valued functions, to replace continuity with the slightly weaker concept of approximate continuity.

A *point of density* of a measurable set $E \subset [0, T]$ is a point $t \in E$ such that $\lim_{h \rightarrow 0} (1/2h) \mu([t - h, t + h] \cap E) = 1$. It is known (Theorem 1, pg. 261, [7]) that almost all points of E are points of density.

Let P be a measurable, compact, set valued function defined on $[0, T]$ and let $t \in [0, T]$. If there exists a measurable subset $B \subset [0, T]$ having t as a point of density and such that the restriction of P to B is continuous (in the relative topology of B) at t , P is said to be *approximately continuous* at t .

Lemma 1.1. *If P is measurable on $[0, T]$ with values compact nonempty subsets of a given compact subset of E^n , then P is approximately continuous at almost all points of $[0, T]$.*

Proof. Let $\epsilon > 0$ be given. By the Plí's generalization of Lusin's theorem there exists a closed set $B \subset [0, T]$ with measure differing from T by less than ϵ and on which P is continuous. Let $D \subset B$ be the set of points of density of B . Then $\mu(D) = \mu(B) > T - \epsilon$. If $t \in D$ then P is approximately continuous at t hence the set $H \subset [0, T]$ of all points of approximate continuity of P has inner measure greater than $T - \epsilon$. Since $\epsilon > 0$ is arbitrary, the inner measure of H is greater than or equal to T . But $H \subset [0, T]$ hence its outer measure is less than or equal to T showing H is measurable with measure T .

Lemma 1.2. *Let P be a measurable set valued function defined on $[0, T]$ with values compact nonempty subsets of B_K . Then for almost all $t \in [0, T]$, $\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} P(\tau) \subset \text{co } P(t)$.*

Proof. Let t be a point of approximate continuity of P . Then there exists a measurable subset $B \subset [0, T]$ such that t is a point of density of B and the restriction of P to B is continuous. From continuity, for any $\delta > 0$ let $\alpha > 0$ be such that $\rho(P(\tau), P(t)) < \delta/2$ for $\tau \in [t - \alpha, t + \alpha] \cap B$. Since t is a point of density of B let $h_0 > 0$ (we shall assume $h_0 < \alpha$) be such that for $0 < |h| < h_0$, $(1/|h|)\mu([t, t+h] \cap B') < \delta/4K$.

Now let p be any measurable function with values $p(\tau) \in P(\tau)$, $\tau \in [0, T]$. Then

$$(1/h) \int_t^{t+h} p(\tau) d\tau = \int_0^1 p(h\sigma + t) d\sigma.$$

Clearly $\mu\{\sigma \in [0, 1] : \sigma h + t \in [t, t+h] \cap B\} = (1/|h|)\mu([t, t+h] \cap B)$. Therefore for $0 < |h| < h_0$, $p(h\sigma + t) \in \mathfrak{N}(\delta/2, P(t))$ for all $\sigma \in [0, 1]$ except a set of measure less than $\delta/4K$. Let q be measurable, agree with p for values σ such that $p(h\sigma + t) \in \mathfrak{N}(\delta/2, P(t))$, and $q(h\sigma + t) \in P(t)$ for all other σ . Then $q(h\sigma + t) \in \mathfrak{N}(\delta/2, P(t))$ for all $\sigma \in [0, 1]$ hence

$$\int_0^1 q(h\sigma + t) d\sigma \in \text{co } \mathfrak{N}(\delta/2, P(t)) = \mathfrak{N}(\delta/2, \text{co } P(t)).$$

Since $|p(h\sigma + t) - q(h\sigma + t)| \leq 2K$ and this can occur at most on a set of values σ having measure less than $\delta/4K$,

$$\left| \int_0^1 q(h\sigma + t) d\sigma - \int_0^1 p(h\sigma + t) d\sigma \right| < \delta/2.$$

Therefore

$$\int_0^1 p(h\sigma + t) d\sigma = (1/h) \int_t^{t+h} p(\tau) d\tau \in \mathfrak{N}(\delta, \text{co } P(t))$$

for any measurable function p with values $p(\tau) \in P(\tau)$ and $0 < |h| < h_0$. Since

co $P(t)$ is closed and $\delta > 0$ is arbitrary, $\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} P(\tau) d\tau \subset \text{co } P(t)$ for every point t of approximate continuity of P , therefore by Lemma 1.1, for almost all $t \in [0, T]$.

Lemma 1.2 is included since it has future application and a direct proof. By using results from [1], the following strengthened version can be obtained.

Lemma 1.3. *Let P be a Borel measurable set valued function defined on $[0, T]$ with values nonempty, compact, subsets of $B_K \subset E^n$. Then $\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} P(\tau) d\tau = \text{co } P(t)$ for almost all $t \in [0, T]$.*

Proof. Let t be a point of approximate continuity of P and B the measurable subset of $[0, T]$ such that t is a point of density of B and P restricted to B is continuous. We may, without loss of generality, take B to be closed. Then P has a continuous extension Q to all of $[0, T]$. Indeed, let (t_1, t_2) be a connected component of the complement of B . Join every point p^1 of $P(t_1)$ to every point p^2 of $P(t_2)$ by a straight line segment parametrized by $t \in [t_1, t_2]$. Then

$$Q(t) = \{1/(t_2 - t_1)[(t - t_1)p^2 + (t_2 - t)p^1] : p^2 \in P(t_2), p^1 \in P(t_1)\}$$

for $t \in (t_1, t_2)$ is as desired.

Now

$$\begin{aligned} (1/h) \int_t^{t+h} P(\tau) d\tau &= \int_0^1 P(h\sigma + t) d\sigma \\ &= \int_{E(h)} P(h\sigma + t) d\sigma + \int_{E'(h)} P(h\sigma + t) d\sigma, \end{aligned}$$

where $E(h) = \{\sigma \in [0, 1] : \sigma h + t \in [t, t+h] \cap B\}$ and $E'(h)$ is its complement in $[0, 1]$. Then $\mu(E(h)) = (1/h)\mu([t, t+h] \cap B)$; since t is a point of density of B , $\mu(E(h)) \rightarrow 1$ as $h \rightarrow 0$ and $\mu(E'(h)) \rightarrow 0$. Thus

$$\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} P(\tau) d\tau = \lim_{h \rightarrow 0} \int_{E(h)} P(h\sigma + t) d\sigma.$$

It is this latter limit which we shall compute.

Since Q is a continuous extension of the restriction of P to B ,

$$\int_0^1 Q(h\sigma + t) d\sigma = \int_{E(h)} P(h\sigma + t) d\sigma + \int_{E'(h)} Q(h\sigma + t) d\sigma.$$

But $\mu(E'(h)) \rightarrow 0$ as $h \rightarrow 0$ so it suffices to compute $\lim_{h \rightarrow 0} \int_0^1 Q(h\sigma + t) d\sigma$.

Let $\{h_k\}$ be any sequence of reals tending to zero as $k \rightarrow \infty$. Define $Q_k(\sigma) = Q(h_k\sigma + t)$. Then each set valued function Q_k is Borel measurable on $[0, 1]$ with values in an origin centered ball of radius K . Also, for any $\sigma \in [0, 1]$, $\lim_{k \rightarrow \infty} Q_k(\sigma) = Q(t) = P(t)$, the limit being taken in the Hausdorff topology. Therefore this limit exists in the sense used by Aumann. ([1], Theorem 5). Applying this theorem gives

$$\lim_{k \rightarrow \infty} \int_0^1 Q(h_k\sigma + t) d\sigma = \lim_{k \rightarrow \infty} \int_0^1 Q_k(\sigma) d\sigma = \int_0^1 P(t) d\sigma = \text{co } P(t);$$

the first equality being the definition of Q_k , the second using theorem 5 of [1] and the third using theorem 3 of [1]. This holds for every point t of approximate continuity of P , i.e. for almost all $t \in [0, T]$.

As a simple example, with P continuous, consider $P(t) = \{-1, 1\} \subset \mathbb{R}^1$. Then for every $t > 0$,

$$\int_0^t P(\tau) d\tau = [-t, t] \quad \text{while} \quad \int_t^{t+h} P(\tau) d\tau = [-h, h].$$

Hence

$$\lim_{h \rightarrow 0} (1/h) \int_t^{t+h} P(\tau) d\tau = [-1, 1] = \text{co } P(t).$$

It is clear from Theorem 3 of [1] that a sufficient condition for two Borel measurable set valued functions P^1 and P^2 to have the same integral is that they satisfy $\text{co } P^1(t) = \text{co } P^2(t)$ for almost all t . We may now show that this is also a necessary condition.

Theorem 1.1. *Let P^1, P^2 be Borel measurable set valued functions defined on $[0, T]$ with values nonempty compact subsets of $B_K \subset E^n$. Let $\mathfrak{A}^1(t) = \int_0^t P^1(\tau) d\tau$, $\mathfrak{A}^2(t) = \int_0^t P^2(\tau) d\tau$. Then a necessary and sufficient condition that $\mathfrak{A}^1(t) = \mathfrak{A}^2(t)$ for all $t \in [0, T]$ is that $\text{co } P^1(t) = \text{co } P^2(t)$ for almost all $t \in [0, T]$.*

Proof. Sufficiency is Theorem 3 of [1]. To show necessity assume that $\mathfrak{A}^1(t) = \mathfrak{A}^2(t)$ for all $t \in [0, T]$ but that $E = \{t \in [0, T] : \text{co } P^1(t) \neq \text{co } P^2(t)\}$ has positive measure. Let $E^1 = \{t \in [0, T] : \text{co } P^1(t) - \text{co } P^2(t) \neq \emptyset\}$ and $E^2 = \{t \in [0, T] : \text{co } P^2(t) - \text{co } P^1(t) \neq \emptyset\}$. By assumption at least one of the sets E^1, E^2 has positive measure, let it be E^1 . Then $\text{co } P^1 - \text{co } P^2$ is a nonempty, Borel measurable, set valued function on E^1 . Hence ([1], Prop. 2.1) it contains a measurable point valued function p^1 . Let $t' \in E^1$ be in the Lebesgue set of p^1 and a point of approximate continuity of P^2 . The existence of such a t' follows easily from Lemma 1.1 and the fact that almost every point is in the Lebesgue set of a measurable function.

Since $E^1 \subset E$, $p^1(t') \notin \text{co } P^2(t')$ and there exists an $\eta \in S^{n-1}$ such that

$$(1-1) \quad \eta \cdot p^1(t') > \sup \{\eta \cdot p : p \in P^2(t')\}.$$

For convenience, extend p^1 as a measurable function on $[0, T]$ with values $p^1(t) \in P^1(t)$; let p^1 also denote this extension.

Now let $a(t')$ be any point in $\mathfrak{A}^1(t') = \mathfrak{A}^2(t')$ having maximum projection on η . For any $\epsilon > 0$, define

$$a(t' + \epsilon) = \left(a(t') + \int_{t'}^{t'+\epsilon} p^1(\tau) d\tau \right) \in \mathfrak{A}^1(t' + \epsilon).$$

Suppose $a(t' + \epsilon) \in \mathfrak{A}^2(t' + \epsilon)$. Then there exists a measurable function p^2 with values $p^2(\tau) \in P^2(\tau)$ such that

$$\begin{aligned}
 (1-2) \quad a(t' + \epsilon) &= \int_0^{t'+\epsilon} p^2(\tau) d\tau = \int_0^{t'} p^2(\tau) d\tau + \int_{t'}^{t'+\epsilon} p^2(\tau) d\tau \\
 &= a(t') + \int_{t'}^{t'+\epsilon} p^1(\tau) d\tau.
 \end{aligned}$$

From our choice of $a(t')$ it follows that

$$\eta \cdot \int_0^{t'} p^2(\tau) d\tau \leq \eta \cdot a(t')$$

thus for equality to hold in (1-2) we must have

$$\eta \cdot \int_{t'}^{t'+\epsilon} p^2(\tau) d\tau \geq \eta \cdot \int_{t'}^{t'+\epsilon} p^1(\tau) d\tau \quad \text{for all } \epsilon > 0$$

or

$$(1-3) \quad \eta \cdot \left(\frac{1}{\epsilon} \int_{t'}^{t'+\epsilon} p^2(\tau) d\tau \right) \geq \eta \cdot \left(\frac{1}{\epsilon} \int_{t'}^{t'+\epsilon} p^1(\tau) d\tau \right).$$

But t' is in the Lebesgue set of p^1 and a point of approximate continuity of P^2 which, by lemma 1.2, implies $\lim_{\epsilon \rightarrow 0} (1/\epsilon \int_{t'}^{t'+\epsilon} P^2(\tau) d\tau) \subset \text{co } P^2(t')$. Taking a limit as $\epsilon \rightarrow 0$ in (1-3) yields $\eta \cdot p^2 \geq \eta \cdot p^1(t')$ for some $p^2 \in \text{co } P^2(t')$ contradicting (1-1) as desired.

Let P be a measurable, nonempty set valued function defined on $[0, T]$. For any measurable point valued function p with $p(t) \in P(t)$ we shall call the function a defined by $a(t) = \int_0^t p(\tau) d\tau$ a *trajectory* of P . As noted, we may have two different measurable set valued functions P^1 and P^2 such that $\mathfrak{A}^1(t) = \int_0^t P^1(\tau) d\tau = \int_0^t P^2(\tau) d\tau = \mathfrak{A}^2(t)$ for all $t \in [0, T]$. However, it is not true that the trajectories of P^1 and P^2 will be the same. This has the following implication to control theory. If the functional to be minimized in a linear optimal control problem depends only on the state of the system at a single time (for example a time optimal control problem) then knowledge of the attainable set $\mathfrak{A}(t)$ is all that is required for solutions. The fact that a measurable set valued function P and the set valued function $\text{co } P$ have the same integral is the “bang-bang principle” in the special case where $P(t)$ is the set of vertices of a unit cube in E^n for each $t \in [0, T]$.

2. When is a set valued function an integral?

Theorem 2.1. *Let $\{P^k\}$ be a sequence of Borel measurable, nonempty, compact set valued functions, defined on $[0, T]$ and with values in $B_K \subset E^n$. Let $\mathfrak{A}^k(t) = \int_0^t P^k(\tau) d\tau$, $t \in [0, T]$, and assume that $\{\mathfrak{A}^k(\cdot)\}$ converges uniformly on $[0, T]$ to a compact set valued function $\mathfrak{A}(\cdot)$. Then there exists a nonempty, compact set valued, Borel measurable function P with values in B_K , such that $\mathfrak{A}(t) = \int_0^t P(\tau) d\tau$, $t \in [0, T]$. (Briefly, the uniform limit of integrals is an integral).*

Proof. (a) We first will show that for any k , there exists a function a_k ,

defined on $[0, T] \times S^{n-1}$, which is Lipschitz continuous (Lip. constant K) in t for fixed $\eta \in S^{n-1}$, Borel measurable in η , has values $a_k(t, \eta) \in \mathfrak{A}^k(t)$ and for any $\eta \in S^{n-1}$, $t \in [0, T]$, satisfies $a_k(t, \eta) \cdot \eta \geq a \cdot \eta$ for all $a \in \mathfrak{A}^k(t)$.

$\mathfrak{A}^k(T)$ is a compact convex set. For $\eta \in S^{n-1}$ let $a^*(\eta) \in \mathfrak{A}^k(T)$ be such that $a^*(\eta) \cdot \eta = \max \{a \cdot \eta : a \in \mathfrak{A}^k(T)\}$. Then there exists a measurable function $p_k(\cdot, \eta)$, with values $p_k(\tau, \eta) \in P^k(\tau)$ such that $a^*(\eta) = \int_0^T p_k(\tau, \eta) d\tau$. Define $a_k(t, \eta) = \int_0^t p_k(\tau, \eta) d\tau$. Then $a_k(\cdot, \eta)$ is Lipschitz continuous, with Lipschitz constant K , $a_k(t, \eta) \in \mathfrak{A}^k(t)$ and $a_k(T, \eta) = a^*(\eta)$.

Suppose $a_k(t, \eta) \cdot \eta < a \cdot \eta$ for some $t \in [0, T]$ and $a \in \mathfrak{A}^k(t)$. Then $a = \int_0^t q_k(\tau) d\tau$ for some measurable function q_k with values $q_k(\tau) \in P^k(\tau)$. Define

$$p(\tau) = \begin{cases} q_k(\tau) & \text{for } 0 \leq \tau \leq t \\ p_k(\tau, \eta) & \text{for } t < \tau \leq T \end{cases}$$

and $a' = \int_0^T p(\tau) d\tau \in \mathfrak{A}^k(T)$. Clearly $\eta \cdot a' > \eta \cdot a^*$, a contradiction to the choice of a^* . Therefore $a_k(t, \eta) \cdot \eta \geq a \cdot \eta$ for all $a \in \mathfrak{A}^k(t)$, $t \in [0, T]$. From this inequality and [5, pg. 513 Th. 3] it follows that $a_k(t, \cdot)$ is measurable as a function of η . With every measurable function we may associate a Borel measurable function, which differs from it on at most a set of measure zero. Thus we may consider $a_k(t, \cdot)$ Borel measurable.

An argument similar to that given above shows that for any measurable subset $E \subset [0, T]$ and measurable function q_k with values $q_k(\tau) \in P^k(\tau)$,

$$\eta \cdot \int_E p_k(\tau, \eta) d\tau \geq \eta \cdot \int_E q_k(\tau) d\tau$$

and in particular

$$(2-1) \quad \eta \cdot \int_E p_k(\tau, \eta) d\tau \geq \eta \cdot \int_E p_k(\tau, \xi) d\tau, \quad \xi \in S^{n-1}.$$

(b) For each $\eta \in S^{n-1}$, $\{a_k(\cdot, \eta)\}$ is an equicontinuous family of functions defined on $[0, T]$, hence has a uniformly convergent subsequence (we assume it to be the original sequence for notational convenience) converging to a function $a(\cdot, \eta)$. Then $a(\cdot, \eta)$ is Lipschitz continuous, with Lipschitz constant K ; for any $t \in [0, T]$, $a(t, \eta) \in \mathfrak{A}(t)$ and $a(t, \eta) \cdot \eta \geq a \cdot \eta$ for all $a \in \mathfrak{A}(t)$. From this, we may conclude that a is Borel measurable on $[0, T] \times S^{n-1}$.

For any $\eta \in S^{n-1}$ define $p(t, \eta) = \dot{a}(t, \eta)$. Then p is measurable and has values almost everywhere in B_K . Also $\{p_k(\cdot, \eta)\}$, considered as a sequence in $L_\infty[0, T]$ with its weak* topology, converges to $p(\cdot, \eta)$. (See [2], pg. 342, Ex. 27). This implies, using (2-1), that

$$(2-2) \quad \eta \cdot \int_E p(\tau, \eta) d\tau \geq \eta \cdot \int_E p(\tau, \xi) d\tau$$

for any measurable subset $E \subset [0, T]$ and $\xi \in S^{n-1}$.

Define

$$\begin{aligned} F(\tau) &= \{p(\tau, \eta) : \eta \in S^{n-1}\} \\ P(\tau) &= \text{closure } F(\tau). \end{aligned}$$

Then P is a nonempty, compact, set valued function. To show it is Borel measurable, we shall show F is Borel measurable.

Let $\mathfrak{F}(\tau, \eta, h) = (1/h)(a(\tau + h, \eta) - a(\tau, \eta))$, $h > 0$. Then for almost all τ ,

$$\limsup_{h \rightarrow 0} \mathfrak{F}(\tau, \eta, h) = \liminf_{h \rightarrow 0} \mathfrak{F}(\tau, \eta, h) = p(\tau, \eta).$$

From [5], pg. 513, p is both an A function and a CA function, hence Borel measurable, [5], pg. 512). From [5, Th. 1, pg. 384], considering p as a map from $[0, T] \times S^{n-1} \rightarrow E^n$, its graph, which is the "graph" of F , is Borel measurable. This shows F is a Borel measurable set valued function. The same is then true for P .

From (2-2) one may conclude that if g is a measurable function with values $g(\tau) \in F(\tau)$,

$$(2-3) \quad \eta \cdot \int_0^t g(\tau) d\tau \leq \eta \cdot \int_0^t p(\tau, \eta) d\tau.$$

(c) Verification that $\mathfrak{U}(t) = \int_0^t P(\tau) d\tau$. Since $\mathfrak{U}^k(t) \rightarrow \mathfrak{U}(t)$ in the Hausdorff topology and each $\mathfrak{U}^k(t)$ is convex, it follows that $\mathfrak{U}(t)$ is convex; it is closed by hypothesis.

For any $\eta \in S^{n-1}$,

$$a(t, \eta) = \int_0^t d(\tau, \eta) d\tau = \int_0^t p(\tau, \eta) d\tau \in \int_0^t P(\tau) d\tau.$$

Therefore $\{a(t, \eta) : \eta \in S^{n-1}\} \subset \int_0^t P(\tau) d\tau$; since the right side is closed and convex, the closed convex hull of the left side, which is $\mathfrak{U}(t)$, is contained in $\int_0^t P(\tau) d\tau$.

To obtain the reverse inclusion, we first will show that for a Borel measurable, nonempty set valued function F with values in B_K ,

$$(2-4) \quad \int_0^t \text{cl } F(\tau) d\tau = \text{cl } \int_0^t F(\tau) d\tau.$$

Indeed, Aumann, [1, pg. 8], obtains $\text{cl } \int_0^t F(\tau) \subset \int_0^t \text{cl } F(\tau) d\tau$. To get the inclusion in the reverse order, let $\mathfrak{N}(\epsilon, F)$ denote a $1 > \epsilon > 0$ neighborhood of the set $F \subset E^n$. Then if $F(\cdot)$ is Borel measurable with values in a compact set, the same is true for $\mathfrak{N}(\epsilon, F(\cdot))$. Also

$$\int_0^t \mathfrak{N}(\epsilon, F(\tau)) d\tau \subset \mathfrak{N}\left(t\epsilon, \int_0^t F(\tau) d\tau\right)$$

therefore

$$\int_0^t \text{cl } F(\tau) d\tau \subset \int_0^t \mathfrak{N}(\epsilon, F(\tau)) d\tau \subset \mathfrak{N}\left(t\epsilon, \int_0^t F(\tau) d\tau\right).$$

From this, and the fact that the left side is closed, it follows that $\int_0^t \text{cl } F(\tau) d\tau \subset \text{cl } \int_0^t F(\tau) d\tau$. This establishes (2-4).

Now let F be as defined in part b). In view of (2-4) it suffices to show that

$\int_0^t F(\tau) d\tau \subset \mathfrak{A}(t)$. Indeed, since $\mathfrak{A}(t)$ is closed, this will yield

$$\text{cl} \int_0^t F(\tau) d\tau = \int_0^t \text{cl} F(\tau) d\tau = \int_0^t P(\tau) d\tau \subset \mathfrak{A}(t),$$

as desired.

But, from (2-3), for any point $q = \int_0^t g(\tau) d\tau \in \int_0^t F(\tau) d\tau$ and $\eta \in S^{n-1}$ we have $\eta \cdot q \leq \max \{ \eta \cdot a : a \in \mathfrak{A}(t) \}$, implying that q belongs to the convex closure of $\mathfrak{A}(t)$ which is $\mathfrak{A}(t)$. Thus $\int_0^t P(\tau) d\tau \subset \mathfrak{A}(t)$, completing the proof.

We have defined the set valued function $\mathfrak{A}$ to be an integral if there exists a Borel measurable, nonempty, compact set valued function P , with values subsets of $B_K \subset E^n$, such that $\mathfrak{A}(t) = \int_0^t P(\tau) d\tau$ for $t \in [0, T]$. From this it is clear that if an arbitrary set valued function $\mathfrak{A}$ is to be an integral, it is necessary that

- a) $\mathfrak{A}(t)$ be compact and convex for each $t \in [0, T]$.
- b) If ρ is the Hausdorff metric, $\rho(\mathfrak{A}(t), \mathfrak{A}(t')) \leq K |t - t'|$.
- c) If $\mathfrak{A}(t_1)$ has nonempty r dimensional interior, then the same is true for $t \geq t_1$.

Another necessary condition can easily be seen from part a) of the proof of Theorem 2.1. Suppose $\mathfrak{A}$ is an integral, $a^* \in \partial \mathfrak{A}(\tau)$, and $\eta \in S^{n-1}$ is such that $\eta \cdot a^* \geq \eta \cdot a$ for all $a \in \mathfrak{A}(\tau)$. Then there is an absolutely continuous function (see part a) of the proof of Theorem 2.1) $a(\cdot, \eta)$ with values $a(t, \eta) \in \partial \mathfrak{A}(t)$ such that $a(T, \eta) = a^*$ and $\eta \cdot a(t, \eta) \geq \eta \cdot a$ for all $a \in \mathfrak{A}(t)$, $t \in [0, T]$. Think, for the moment, of η as a unit normal to a support plane to $\mathfrak{A}(T)$ at a^* . If a^* is a "vertex" or "corner" of $\mathfrak{A}(T)$ there may be a cone of unit normals to support planes at a^* , call this cone $N(T, a^*)$. Then if $\eta, \eta^1 \in N(T, a^*)$, we have $\eta \cdot a^* \geq \eta \cdot a$ and $\eta^1 \cdot a^* \geq \eta^1 \cdot a$ for all $a \in \mathfrak{A}(T)$. Let the function $a(\cdot, \eta)$ be as above. Then it easily follows that $\eta^1 \cdot a(t, \eta) \geq \eta^1 \cdot a$ for all $a \in \mathfrak{A}(t)$, $t \in [0, T]$. Thus if a^* is a vertex of $\mathfrak{A}(T)$ and $\eta \in N(T, a^*)$, for $t \leq T$, $\mathfrak{A}(t)$ has a vertex at $a(t, \eta)$ and $\eta \in N(t, a(t, \eta))$. *This shows that if $\mathfrak{A}$ is an integral, any "corners" on $\partial \mathfrak{A}$ may "smooth out" with increasing time, but if $\mathfrak{A}(t_1)$ has a corner for some $t_1 > 0$, the same must be true of $\mathfrak{A}(t)$ for $0 < t \leq t_1$.* (Compare [9], p. 316.)

In order to obtain necessary and sufficient conditions that $\mathfrak{A}$ be an integral, we abstract the preceding property as follows.

d) (Normal cone property). There exists a function a defined on $[0, T] \times S^{n-1}$ with values $a(t, \eta) \in \mathfrak{A}(t)$ which is Lipschitz continuous in t and satisfies $a(t, \eta) \cdot \eta \geq a \cdot \eta$ for all $a \in \mathfrak{A}(t)$, $t \in [0, T]$. (As in the proof of Theorem 2.1, part a), we may conclude that $a(t, \cdot)$ is measurable in η).

Note: This implies that if $a^* \in \partial \mathfrak{A}(t_1)$ and $\eta \in N(t_1, a^*)$, then $\eta \in N(t, a(t, \eta))$ for all $0 < t \leq t_1$.

We next show that the normal cone property, property d) implies property c).

Lemma 2.1. *Let $\mathfrak{A}$ be a compact, convex, nonempty set valued function which satisfies property d). Then if $\mathfrak{A}(t_1)$ has nonempty r dimensional interior, the same is true for $\mathfrak{A}(t)$, $t \geq t_1$.*

Proof. Suppose $\mathfrak{A}(t_1)$ has nonempty r dimensional interior but $\mathfrak{A}(t_2)$ has empty r dimensional interior for some $t_2 > t_1$. We will show a contradiction.

Consider everything in E^r ; boundary, interior, etc. will be relative to this. Let $a(t_1)$ be any point on $\partial\mathfrak{A}(t_1)$. Then for any $\eta \in N(t_1, a(t_1))$, since $\mathfrak{A}(t_1)$ has nonempty r dimensional interior, there exists an $a \in \mathfrak{A}(t_1)$ such that $\eta \cdot (a(t_1) - a) > 0$. This shows $-\eta \notin N(t_1, a(t_1))$.

If $\mathfrak{A}(t_2)$ has empty interior, it lies in a hyperplane in E^r , let $\eta' \in S^{n-1}$ be a normal to this hyperplane. If $a(t_2) \in \mathfrak{A}(t_2)$, both η and $-\eta$ belong to $N(t_2, a(t_2))$. By the normal cone property there would be a point $a(t_1) \in \partial\mathfrak{A}(t_1)$ such that $-\eta$ and η belong to $N(t_1, a(t_1))$; a contradiction.

We now easily obtain: *A necessary and sufficient condition that $\mathfrak{A}$ be an integral is that it satisfy properties a) and d).*

Necessity has been shown above. For sufficiency we use the function a of property d) in the same role as its counterpart in part b) of the proof of Theorem 2.1 and proceed to construct P as in the proof of that theorem.

3. Some problems concerning contingent equations. If R is a continuous contingent, it seems natural to try to show the existence (locally) of a solution to the associated contingent equation by first showing that there is a continuous point valued function r , with values $r(t, x) \in R(t, x)$, and then using the existence theorem for ordinary differential equations. It is *not* true, however, that every continuous set valued function contains a continuous point valued function. A standard example of this is constructed as follows. For $0 < t \leq 1$ let $S(t) = \{(x, y) \in E^2 : x = \cos \theta, y = \sin \theta, t \leq \theta \leq 2\pi\}$ and

$$A(t) = \begin{bmatrix} \cos 1/t & \sin 1/t \\ -\sin 1/t & \cos 1/t \end{bmatrix}.$$

Define $R(t) = A(t)S(t)$ for $t > 0$ and $R(0) = S^1$, the unit circle. To see that R is continuous on $[0, 1]$ one need only check continuity at 0. But $\rho(R(0), R(t)) \leq \pi t$ showing this. Any point valued function r in R has its graph on the cylinder $S^1 \times [0, 1]$. If r were continuous its graph would be a connected set. But the "gap" $S^1 - S(t)$, rotated by $A(t)$, disconnects the graph of any continuous function on the cylinder for $t > 0$ and sufficiently small.

To the author's knowledge, the existence of a single solution to a continuous contingent equation is an open question. However, for a linear contingent equation $\dot{x}(t) \in A(t)x(t) + P(t)$, $x(0) = x^0$, one may easily check that if $X(t)$ is a fundamental solution of the ordinary differential equation $\dot{x}(t) = A(t)x(t)$ which is the identity at zero, then solutions do exist and $\mathfrak{A}(t) = X(t)x^0 + X(t) \int_0^t X^{-1}(\tau)P(\tau) d\tau$.

Another question, of practical importance, is the following. Given a continuous contingent equation $\dot{x}(t) \in R(t, x(t))$, $x(0) = x^0$, with nonempty attainable set function $\mathfrak{A}$ defined on $[0, T]$, when is the closed convex hull of $\mathfrak{A}$ the attainable set of a linear contingent? Theorem 3.1, as stated in the introduction, may be considered a first step toward answering this question.

The bound $\max \{x \cdot r : r \in R(t, x)\} \leq c(1 + |x|^2)$ imposed on R in the introduction implies that the attainable set $\mathfrak{A}$ associated with R has values $\mathfrak{A}(t)$ contained in an origin centered ball of radius $(1 + |x^0|^2)^{1/2} \exp(cT)$ for $0 \leq t \leq T$. Thus in the proof of Theorem 3.1 we may consider R defined on the compact set $D = \{t, x : 0 \leq t \leq T, |x|^2 \leq (1 + |x^0|^2) \exp 2cT\}$.

Proof of Theorem 3.1. To avoid trivialities, we assume $\mathfrak{A}(t)$ is nonempty for $t \in [0, T]$. Since each $\mathfrak{A}(t) \subset D$ and R is continuous, $M(t) = \text{closure}(\bigcup_{x \in \mathfrak{A}(t)} R(t, x))$ is compact and contained in some origin centered ball B_K for all $t \in [0, T]$. Then $Q(t, x) = M(t)$ is a K bounded linear contingent which clearly satisfies $Q \geq R$.

Now let $\{Q^\beta\}_{\beta \in B}$ be a totally ordered subset of $\mathcal{O}_K$. If we can show that $\{Q^\beta\}$ has a lower bound, call it Q^* , in $\mathcal{O}_K$, the existence of a minimal element will follow from Zorn's lemma.

Let $\mathfrak{A}^\beta$ denote the attainable set function associated with the linear contingent Q^β , where $Q^\beta(t, x) = A^\beta(t)x + P^\beta(t)$. Let $X_\beta, X_\beta(0) = I$, denote the fundamental solution matrix of the linear differential equation $\dot{x}(t) = A^\beta(t)x$. Then

$$(3-1) \quad \mathfrak{A}^\beta(t) = X_\beta(t)x_0 + X_\beta(t) \int_0^t X_\beta^{-1}(\tau)P^\beta(\tau) d\tau.$$

Thus for every $t \in [0, T]$, $\mathfrak{A}^\beta(t)$ is a nonempty (it contains $\mathfrak{A}(t)$), compact, convex subset of E^n .

Define

$$\mathfrak{A}^*(t) = \bigcap_{\beta \in B} \mathfrak{A}^\beta(t), \quad t \in [0, T].$$

Then $\mathfrak{A}^*(t)$ is nonempty, compact and convex. The problem becomes to show that $\mathfrak{A}^*$ is the attainable set function of some K bounded linear contingent Q^* ; i.e. we must produce A^* and P^* so that $Q^*(t, x) = A^*(t)x + P^*(t)$ is a K bounded linear contingent with attainable set function $\mathfrak{A}^*$.

We next reduce the problem to one of a sequential nature. Since $\{Q^\beta\}_{\beta \in B}$ is totally ordered, we can use it to totally order the index set B , i.e. if $Q^{\beta_1} \geq Q^{\beta_2}$ let $\beta_1 \geq \beta_2$. Assume this has been done. Again, using the total order of $\{Q^\beta\}$, for each positive integer k , we may choose β_k such that the attainable set $\mathfrak{A}^{\beta_k}(t)$ corresponding to Q^{β_k} is contained in a $1/k$ neighborhood of $\mathfrak{A}^*(t)$ for all $t \in [0, T]$. For notational convenience let the attainable set function and contingent associated with the index β_k now be $\mathfrak{A}^k$ and Q^k respectively. Clearly $\{\mathfrak{A}^k\}$ converges, uniformly on $[0, T]$, to $\mathfrak{A}^*$ in the Hausdorff topology.

If $Q^k(t, x) = A^k(t)x + P^k(t)$, our assumption that the elements of A^k are essentially bounded by K implies that each $X_k(\cdot)$ is Lipschitz continuous, with a Lipschitz constant independent of k . Therefore $\{X_k(\cdot)\}$ is an equicontinuous family; it has a uniformly convergent subsequence (which we will not distinguish from the original sequence) converging to a Lipschitz continuous, matrix function X . Since $\det X_k(t) = \exp(\int_0^t \text{trace } A^k(\tau) d\tau)$, it is clear that $\det X(t)$ is nonzero for $t \in [0, T]$. Also, $X(0) = I$, therefore X is a fundamental solution

matrix of the differential equation $\dot{x}(t) = A^*(t)x(t)$ where $A^*(t) = \dot{X}(t)X^{-1}(t)$. A simple argument shows that $\{A^k(\cdot)\}$, considered in the weak* topology of $L_\infty[0, T]$, converges to A^* and hence the elements of A^* are essentially bounded by K .

Now consider, from (3-1), $X_k^{-1}(t)[\mathfrak{U}^k(t) - x_0] = \int_0^t X_k^{-1}(\tau)P^k(\tau) d\tau$, where in this case $(\mathfrak{U}^k(t) - x_0) = \{a - x_0 : a \in \mathfrak{U}^k(t)\}$. Clearly $\{X_k^{-1}(\cdot)[\mathfrak{U}^k(\cdot) - x_0]\}$ converges uniformly on $[0, T]$ to the compact set valued function $X^{-1}(\cdot)[\mathfrak{U}^*(\cdot) - x_0]$. From Theorem 2.1, there exists a nonempty compact set valued, Borel measurable function P such that $X^{-1}(t)[\mathfrak{U}^*(t) - x_0] = \int_0^t P(\tau) d\tau$ for $t \in [0, T]$. Also, from the proof of Theorem 2.1, it is clear that P has values $P(t)$ in the ball $X^{-1}(t)B_K$. Now define $P^*(t) = X(t)P(t)$. Then P^* has values in B_K and $\mathfrak{U}^*$ is the attainable set function associated with the K bounded linear contingent $Q^*(t, x) = A^*(t)x + P^*(t)$.

REFERENCES

- [1] R. J. AUMANN, Integrals of set-valued functions, *J. Math. Anal. and Appl.*, 12 (1965) 1-12.
- [2] N. DUNFORD & J. T. SCHWARTZ, *Linear Operators I*, Interscience, N. Y., 1958.
- [3] A. F. FILIPPOV, Differential equations with many-valued discontinuous right-hand sides, *Soviet Math.* (English trans.), 4 (1963) 941-945.
- [4] H. HERMES, A note on the range of a vector measure; application to the theory of optimal control, *J. Math. Anal. and Appl.*, 8 (1964) 78-83.
- [5] K. KURATOWSKI, *Topology* (English trans.), Academic Press, N. Y., 1966.
- [6] E. MICHAEL, Topologies on spaces of subsets, *Trans. Amer. Math. Soc.*, 71 (1951) 152-182.
- [7] I. P. NATANSON, *Theory of Functions of a Real Variable, Vol. I* (English Translation), P. Ungar Publ Co., New York, N. Y., 1955.
- [8] A. PLIS, Remark on measurable set-valued functions, *Bull. Acad. Polon. Sci. Math. Astr. & Phys.*, IX (1961) 857-859.
- [9] E. ROXIN, A geometric interpretation of pontryagins maximum principle, *Proc. Symp. Nonlinear Diff. Eqs. & Nonlinear Mech.*, Academic Press., N. Y., (1963), 303-324.
- [10] L. M. SONNEBORN & F. S. VAN VLECK, The bang bang principle for linear control systems, *J. SIAM Control*, 2 (1965) 151-159.

University of Colorado

Communication date: DECEMBER 10, 1967