JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **179**, 483–499 (1993)

Parametrizations of Integrals of Set-Valued Mappings and Applications

B. D. DOITCHINOV

Institute of Mathematics, Bulgarian Academy of Sciences, ul. Acad. G. Bontchev 8, 1113 Sofia, Bulgaria

AND

V. VELIOV

International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

Submitted by J.-P. Aubin

Received September 14, 1990

1. INTRODUCTION

Consider the integral (in the sense of Aumann [1]) of a set-valued mapping in the form $t \to F(t)\,U$, where $F(t)$ is $n \times m$ matrix and $U \subset \Re^m$. That is,

$$I = \int_{t_0}^{T} F(t)\,U dt = \left\{ \int_{t_0}^{T} F(t)\,u(t)\,dt;\, u(t) \in U,\, u(\cdot) \in \mathscr{L}_1(t_0, T) \right\}. \tag{1}$$

By the very definition I is in parametric form depending linearly on the "parameter"

$$u(\cdot) \in \mathscr{U} = \{u(\cdot) \in \mathscr{L}_1(t_0, T);\, u(t) \in U\}.$$

The numerical treatment of such integrals (which implicitly or explicitly arise in control theory) requires approximation of I by means of sets being parametrized by a finite dimensional parameter, i.e.,

$$I_\varepsilon = \{J_\varepsilon(p);\, p \in P_\varepsilon \subset R^{N(\varepsilon)}\}. \tag{2}$$

There is a number of ways to do this and many of them are exploited in the numerical methods for optimal control problems, guaranteed estimation of uncertain systems, etc. One simple way is to consider only piece-wise constant selections $u(\cdot)$ of U in (1). If $t_1 < \cdots < t_{N-1}$ are the jump points

0022-247X/93 $5.00

and $u_i \in U$ are the values of $u(\cdot)$ in $[t_i, t_{i+1}]$, $i = 0, ..., N-1$, both $t_1, ... t_{N-1}$ and $u_0, ..., u_{N-1}$ considered as parameters, then

$$I_N = \left\{ \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} F(s)\, ds\, u_i;\ t_0 \leqslant t_1 \leqslant \cdots \leqslant t_N = T, u_0, ..., u_{N-1} \in U \right\} \quad (3)$$

is in the form of (2) and under quite general assumptions.

$$\lim_{N \to +\infty} H(I_N, I) = 0,$$

where H is the Hausdorff distance between compact sets. It happens even that $I_N = I$ for sufficiently large N (for instance if U is a polyhedron and the entries of $F(\cdot)$ are analytic functions). The parameter p in (2) is $p = (t_1, ..., t_{N-1}, u_0, ..., u_{N-1})$ and the dependence of $J_\varepsilon(p)$ on the second group of parameters is linear. (In our consideration we of course presume computability of all the integrals

$$\Phi(s, t) = \int_s^t F(\tau)\, d\tau,$$

in order to concentrate on the specificity of the set-valuedness.) An essential shortcoming of the approximation (3) is the possibly complicated dependence of $J_\varepsilon(\cdot)$ on $t_1, ..., t_{N-1}$. Moreover, the constraints on $t_1, ..., t_{N-1}$ in (3) result in interconnection of the summands in (3) which causes computational inconvenience. For this reason, it may be preferable to fix the jump points t_i rather than to consider them as parameters.

If we fix the grid $t_0, ..., t_N = T$ setting for simplicity $t_k = t_0 + kh$, $h = (T - t_0)/N$, then (3) should be replaced with

$$I_N = \left\{ \sum_{k=0}^{N-1} \Phi(t_k, t_{k+1})\, u_k;\ u_k \in \operatorname{co} U \right\}. \quad (4)$$

If $F(\cdot)$ is Lipschitz continuous then it is easy to verify that

$$\Delta_N^k = H\left(\operatorname{co} \Phi(t_k, t_{k+1})\, U, \int_{t_k}^{t_{k+1}} F(t)\, U dt \right) \leqslant ch^2$$

for some constant c which is independent of N and k. On the other hand, simple examples show (and this is typical) that even for analytic $F(\cdot)$

$$\Delta_N^k \geqslant c' h^2,$$

where c' is also independent of N and k (take for instance $n = 2$, $m = 1$,

$F(t) = (t, 1)^*$, $U = [-1, 1]$, $c' = 0.25$). From these inequalities one can conclude that

$$H(I_N, I) \leqslant c(T - t_0)^2/N$$

and that better order than $1/N$ is not probable in this estimation.

The good news is that in fact the estimate

$$H(I_N, I) \leqslant C/N^2$$

does hold if $\dot{F}(\cdot)$ is of bounded variation. In other words, despite that errors Δ_N^k of all the summands in (4) can be of order $1/N^2$, the error of the sum remains of order $1/N^2$, which means that the errors *do not accumulate*. This follows from the result proven in Section 3.

In Section 2 we investigate a more general problem, namely the order of approximation of I by using selections of U in (1) which are piece-wise constant and have a limited number of jumps (say p) in each of the intervals $[t_i, t_{i+1}]$. We establish a similar result to the one above (the latter corresponding to $p = 0$), but depending also on some geometric properties of U.

More precisely, let $\mathcal{U}_p(s, t)$ be the set of all piece-wise constant selections of U on $[s, t]$, which have at most p jumps. Then consider

$$I_{N,p} = co\left\{\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} F(s)\, u_i(s)\, ds;\ u_i(\cdot) \in \mathcal{U}_p(t_i, t_{i+1}),\ i = 0, ..., N-1\right\}. \qquad (5)$$

We suppose that U is a compact convex polyhedron and define $q = q(U)$ to be the number of the nonparallel edges of U; that is, the maximal number q for which there are q edges of U, every two of them nonparallel. Under the appropriate smoothness condition on $F(\cdot)$ we prove the estimate

$$H(I_{N,p}, I) \leqslant c/N^{2+[p/q]}, \qquad (6)$$

where $[p/q]$ is the integer part of p/q.

From (6) one can guess that for a nonpolyhedral set U (say an ellipsoid) no better estimate than $1/N^2$ is true, no matter how large a value of p, is used. This is actually the case as shown by an example.

In Section 4 we indicate some applications of the above results and some open problems.

2. Higher than Second Order Approximations

In this section we consider the approximation (5) of the integral (1), where $p \geqslant 0$. The set U is supposed to be a convex compact polyhedron and the number $q = q(U)$ is defined as in Section 1.

We note that the variation $\bigvee_{t_0}^{T} f$ of a function f defined almost everywhere in $[t_0, T]$ is the supremum of the sums

$$\sum_{i=1}^{k} |f(t_i) - f(t_{i+1})|$$

over all the finite collections of points $t_0 \leqslant t_1 \cdots < t_{k+1} \leqslant T$ for which $f(t_i)$, $i = 1, ..., t_{k+1}$, is defined.

Denote $r = [p/q]$ as the integer part of p/q.

Theorem 1. *Let $F(\cdot)$ be r-times differentiable, let the rth derivative be Lipschitz continuous, and the $(r+1)$st derivative be of bounded variation. Then there exists a constant C such that* (6) *holds for every integer N.*

In the proof we use the following two lemmas.

Lemma 1. *Let φ: $[t_0, T] \to \mathfrak{R}$ be with the same differentiability properties as F in Theorem* 1. *Then for every $\alpha_0, \alpha_1, ..., \alpha_r \in [t_0, T]$ it holds*

$$\varphi(t) = \varphi(\alpha_0) + \varphi'(\alpha_1) \int_{\alpha_0}^{t} ds_1 + \varphi''(\alpha_2) \int_{\alpha_0}^{t} \int_{\alpha_1}^{s_1} ds_2\, ds_1 + \cdots$$

$$+ \varphi^{(r)}(\alpha_r) \int_{\alpha_0}^{t} \cdots \int_{\alpha_{r-1}}^{s_{r-1}} ds_r \cdots ds_1 + \int_{\alpha_0}^{t} \cdots \int_{\alpha_r}^{s_r} \varphi^{(r+1)}(s_{r+1})\, ds_{r+1} \cdots ds_1.$$

Lemma 2. *Let $\psi(\cdot)$: $[t_0, T] \to R^n$ be Lipschitz continuous and let $\dot{\psi}(\cdot)$ be of bounded variations. Denote by $K_1(\psi)$ the set of those k for which the function $\psi(\cdot)$ vanishes at least once in (t_k, t_{k+1}).*
Then

$$\sum_{k \in K_1(\psi)} \int_{t_k}^{t_{k+1}} |\dot{\psi}(t)|\, dt \leqslant h\left(2 \bigvee_{t_0}^{T} \dot{\psi} + \|\dot{\psi}\|_{L\infty}\right).$$

Proof. This inequality is obvious if $K_1(\psi)$ is empty or if it contains only one number, thus let (t_{i_j}, t_{i_j+1}), $j = 1, ..., s > 1$, be the intervals in which $\psi(\cdot)$ vanishes. To simplify the notations denote $\alpha_j = t_{i_j}$, $\beta_j = t_{i_j+1}$. According to Theorem 2.3.7 in Clarke [2] there are $s_j \in (\alpha_j, \beta_{j+1})$ such that $0 \in \partial\psi(s_j)$, where $\partial\psi$ is the Clarke subdifferential of ψ.

There are $\tau_j \in (\alpha_j, \beta_j)$ such that

$$\int_{\alpha_j}^{\beta_j} |\dot{\psi}(t)|\, dt \leqslant h|\dot{\psi}(\tau_j)|. \tag{7}$$

Moreover, τ_j can be chosen to be different from s_i, $i = 1, ..., s-1$. Theorem 2.5.1 of [2] implies that for a given $\varepsilon > 0$ there are $s_j' \leqslant s_j \leqslant s_j''$ such that the derivatives $\dot{\psi}(s_j')$ and $\dot{\psi}(s_j'')$ exist,

$$|\delta_j \dot{\psi}(s_j') + (1-\delta_j)\dot{\psi}(s_j'')| \leqslant \varepsilon$$

for some $\delta_j \in [0, 1]$ and $\tau_j, \tau_{j+1} \notin [s_j', s_j'']$.
Consider the following two cases

(1) $s_j \in (\tau_j, \tau_{j+1})$. Then

$$\begin{aligned}|\dot{\psi}(\tau_j)| + |\dot{\psi}(\tau_{j+1})| &\leqslant |\dot{\psi}(\tau_j) - \dot{\psi}(s_j') + (1-\delta_j)(\dot{\psi}(s_j') - \dot{\psi}(s_j''))| \\ &\quad + |\dot{\psi}(\tau_{j+1}) - \dot{\psi}(s_j'') + \delta_j(\dot{\psi}(s_j'') - \dot{\psi}(s_j'))| + 2\varepsilon \\ &\leqslant \bigvee_{\alpha_j}^{\beta_{j+1}} \dot{\psi} + 2\varepsilon.\end{aligned}$$

(2) $s_j \leqslant \tau_j$ ($s_j > \tau_{j+1}$ is analogous). Then similarly

$$|\dot{\psi}(\tau_j)| + |\dot{\psi}(\tau_{j+1})| \leqslant 2 \bigvee_{\alpha_j}^{\beta_{j+1}} \dot{\psi} + 2\varepsilon. \tag{8}$$

Thus (8) holds in any case and implies

$$\sum_{j=1}^{s} |\dot{\psi}(\tau_j)| \leqslant 2 \bigvee_{t_0}^{T} \dot{\psi} + \|\dot{\psi}\|_{L\infty} + 2\varepsilon s.$$

Taking into account (7) and the free choice of ε we complete the proof.

Q.E.D.

LEMMA 3. *Let $\varphi(\cdot)\colon [t_0, T] \to \mathfrak{R}$ be as in Lemma 1. Denote by $K_{r+1}(\varphi)$ the set of those K for which $\varphi(\cdot)$ vanishes at least at $(r+1)$ points in $[t_k, t_{k+1}]$. Then*

$$\sum_{k \in K_{r+1}(\varphi)} \int_{t_k}^{t_{k+1}} |\varphi(t)|\, dt \leqslant h^{r+2}\left(\|\varphi^{(r+1)}\|_{L\infty} + 2 \bigvee_{t_0}^{T} \varphi^{(r+1)}\right). \tag{9}$$

Proof. For each $k \in K_{r+1}(\varphi)$ there are $\alpha_0, ..., \alpha_r \in [t_k, t_{k+1}]$ such that $\varphi^{(i)}(\alpha_i) = 0$, $i = 0, ..., r$. Then (9) follows from Lemma 1 and Lemma 2 applied to $\psi(\cdot) = \varphi^{(r)}(\cdot)$.

Q.E.D.

Proof of Theorem 1. Since

$$H(I_{N,p}, I) = \max_{|l|=1} (\rho(l|I) - \rho(l|I_{N,p})),$$

where $\rho(l|X) = \max_{x \in X} \langle l, x \rangle$ is the support function of the compact set X, we need to estimate

$$\int_{t_0}^{T} \max_{u \in U} \langle \varphi(t), u \rangle \, dt - \sum_{k=0}^{N-1} \max_{u_k(\cdot) \in \mathscr{U}_p(t_k, t_{k+1})} \int_{t_k}^{t_{k+1}} \langle \varphi(t), u_k(t) \rangle \, dt,$$

where $\varphi(t) = F(t)^* l$ and * means the transposition.

Let $g_1, ..., g_q$ be unit vectors such that every edge e of U is parallel to some of these vectors, indicated by $g_{i(e)}$.

Denote

$$\varphi_i(t) = \langle \varphi(t), g_i \rangle$$

and

$$J_k = \{i; \varphi_i(t) = 0 \text{ at more than } r \text{ points from } (t_k, t_{k+1})\}.$$

For each $k = 0, ..., N-1$ we define a relation of equivalence $\sim_k$ in the set $\mathscr{V}$ of all vertices of U in the following way. For two vertices v' and v'' we say that $v' \sim_k v''$ (v' equivalent to v'') if there is a sequence $e_1, ..., e_s$ of neighbouring edges of U connecting v' and v'' and such that $i(e_j) \in J_k$ for every $j = 1, ..., s$.

Let V_k be the factor set of $\mathscr{V}$ with respect to $\sim_k$ and let $\mathscr{P}_k : \mathscr{V} \to V_k$ be the natural mapping of $\mathscr{V}$ on V_k. For every $\omega \in V_k$ we fix arbitrarily an element $S_k(\omega) \in \mathscr{V}$ such that $\mathscr{P}_k S_k(\omega) = \omega$.

Denote

$$U(t) = \{v \in \mathscr{V}; \langle \varphi(t), v \rangle = \rho(\varphi(t) | U)\}$$

and

$$V_k(t) = \mathscr{P}_k U(t), \qquad t \in [t_k, t_{k+1}].$$

We prove that $V_k(t)$ is single valued excepting at most p points from (t_k, t_{k+1}) and that it is constant in every interval of single-valuedness.

The second claim is obvious since $U(\cdot)$ is upper semicontinuous and the set $\mathscr{V}$ is finite.

If $t \in (t_k, t_{k+1})$ is a point at which $V_k(t)$ is not single valued then there is a sequence of neighbouring edges of U connecting two vertices v' and v'' of U, v', $v'' \in U(t)$, $\mathscr{P}_k v' \neq \mathscr{P}_k v''$ such that $\langle \varphi(t), g_{i(e)} \rangle = 0$ for everyone edge e of them. At least one of these edges is such that $i(e) \notin J_k$ since otherwise v' and v'' would be equivalent. Thus for at least one $i \in \{1, ..., q\} \backslash J_k$ we have $\varphi_i(t) = 0$. This implies that the number of the points $t \in (t_k, t_{k+1})$ at which $V_k(t)$ is not single valued cannot be greater than $(q - |J_k|) r \leqslant qr \leqslant p$, where $|J_k|$ is the number of the elements of J_k.

Now we can define correctly

$$\hat{u}(t) = S_k V_k(t)$$

as a single valued function excepting at most p points in (t_k, t_{k+1}), $\hat{u}(\cdot)$ being a piece-wise constant with not more than p jumps. Thus $\hat{u}(\cdot) \in \mathscr{U}_p(t_k, t_{k+1})$ for every $k = 0, ..., N-1$. We prove that

$$\int_{t_0}^{T} \langle \varphi(t), \hat{u}(t) \rangle \, dt \geqslant \int_{t_0}^{T} \rho(\varphi(t) \mid U) \, dt - Ch^{r+2},$$

where C does not depend on N and l and this will complete the proof of the theorem.

Take $t \in (t_k, t_{k+1})$ for which $V_k(t)$ is single-valued and take an arbitrary $v_0 \in U(t)$. Then estimate

$$\Delta(t) = \rho(\varphi(t) \mid U) - \langle \varphi(t), \hat{u}(t) \rangle = \langle \varphi(t), v_0 - \hat{u}(t) \rangle.$$

Since $v_0 \sim_k \hat{u}(t)$ there is sequence of neighbouring (and different) edges $[v_0, v_1], ..., [v_{s-1}, v_s]$, $v_s = \hat{u}(t)$, every one of which is parallel to some $g_{j_i} \in J_k$, $i = 1, ..., s$. Thus

$$\Delta(t) \leqslant \sum_{i=1}^{s} \langle \varphi(t), v_i - v_{i-1} \rangle \leqslant c \sum_{i \in J_k} |\langle \varphi(t), g_i \rangle|,$$

where c is the length of the longest edge of U.

Let $K(i)$ be the set of those k for which $g_i \in J_k$. Then

$$\int_{t_0}^{T} \Delta(t) \, dt \leqslant c \sum_{k=0}^{N-1} \sum_{i \in J_k} \int_{t_k}^{t_{k+1}} |\varphi_i(t) \, dt| = c \sum_{i=1}^{q} \sum_{k \in K(i)} \int_{t_k}^{t_{k+1}} |\varphi_i(t)| \, dt$$
$$\leqslant ch^{r+2} \sum_{i=1}^{q} \left(\|\varphi_i^{(r+1)}\|_{L_\infty} + 2 \bigvee_{t_0}^{T} \varphi_i^{(r+1)} \right) \leqslant Ch^{r+2}.$$

Q.E.D.

Remark. If U is a coordinate polyhedron in $\mathfrak{R}^m$, i.e.,

$$U = \{u = (u_1, ..., u_m); u_i \in [a_i, b_i], i = 1, ..., m\},$$

then ovbiously $q=m$ and $p=m$ switching points in every interval $[t_k, t_{k+1}]$ are enough to ensure approximation of I of order $1/N^3$. The proof of Theorem 1 implies in fact a slightly stronger result, namely that one jump of every component of $u(\cdot)$ in every interval $[t_k, t_{k+1}]$ is enough for third order accuracy. In Section 4 we indicate some applications of the last fact.

Now we show by examples that: 1) the estimation (6) is exact at least in the sense that $p<q$ in general is not enough for third order accuracy in (6); and 2) if U is a ball then (6) holds only with c/N^2 in the right-hand side, independently of how large p is.

EXAMPLE 1. Take $n=r=2$, $[t_0, T]=[0, 1]$.

$$F(t)=\begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad U=\operatorname{co}\left\{\begin{pmatrix} \cos\dfrac{k\Pi}{4q} \\ \sin\dfrac{k\Pi}{4q} \end{pmatrix}, k=0, ..., q\right\}.$$

It is easy to calculate that

$$H\left(\int_0^h F(t)\, U\, dt, \operatorname{co}\left\{\int_0^h F(t)\, u(t)\, dt; (\cdot)\in \mathscr{U}_{q-1}(0, h)\right\}\right)\geqslant ch^2$$

for an appropriate $c>0$.

EXAMPLE 2. Take $n=r=2$, $[t_0, T]=[0, 1]$.

$$F(t)=\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}, \qquad U=\{u\in \Re^2; |u|\leqslant 1\}.$$

It turns out that

$$H\left(\int_0^1 F(t)\, U\, dt, \left\{\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} F(t)\, u_k(t)\, dt; u_k(\cdot)\in \mathscr{U}_p(t_k, t_{k+1})\right\}\right)$$
$$\geqslant \frac{h^2}{48}\left(\frac{1}{(p+1)^3}-\frac{h^2}{80}\right)$$

whatever the number p of possible jumps in $[t_k t_{k+1}]$ is. In the next section we prove that the left-hand side of the above inequality is actually of order 2 with respect to h, even for $p=0$.

3. Second Order Approximation

In this section we investigate the approximation of the integral (1) by finite sums of the type of (4). The more general case

$$I=\int_{t_0}^{T} F(t, V)\, dt \tag{10}$$

will be considered, where $F(\cdot)$: $[t_0, T] \to \mathfrak{R}^n$ is a maping and V is an arbitrary (abstract) set.

By definition $F(t, V)=\{F(t, v); v \in V\}$. The approximation (4) is now in the form

$$I_N=\operatorname{co}\left\{\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} F(t, v_i)\, dt; v_i \in V, i=0, ..., N-1\right\}. \tag{11}$$

We assume the following:

(A1) $F(t, V)$ is compact for every $t \in [t_0, T]$;

(A2) $F(\cdot, v)$ is Lipschitz continuous in $[t_0, T]$, uniformly in $v \in V$.

Under the above conditions I is a convex and compact set and

$$H(I_N, I)=\max_{|l|=1} (\rho(l \mid I)-\rho(l \mid I_N)).$$

In order to estimate the right-hand side we introduce the following notion of joint variation of a family of functions. Let $\{f(\cdot, v)\}_{v \in V}$ be a family of functions parametrized by v (where as above V is an arbitrary set) every one of them defined for a.e. $t \in [t_0, T]$. We say that the family is of bounded joint variation if there is a number w such that

$$\sum_{i=1}^{s} |f(t_i, v_i)-f(t_{i+1}, v_i)| \leqslant w \tag{12}$$

for all finite collections $t_0 \leqslant t_1 < \cdots < t_{s+1} \leqslant T$, $v_1, ... v_s \in V$ for which $f(t_i, v_i)$ and $f(t_{i+1}, v_i)$ are defined $i=1, ..., s$. If $\{f(\cdot, v)\}_{v \in V}$ is of bounded joint variation then the intimum of those numbers w which satisfy (12) will be called the joint variation of the family and will be denoted by $J\bigvee_{t_0}^{T} f(\cdot, V)$.

We mention that boudedness of the joint variation of a family is a stronger property than uniform boundedness of the variations of the functions from this family. The following proposition is a motivation for the above definition.

PROPOSITION 1. *Let conditions* (A1) *and* (A2) *be satisfied and let the variation of* $(\partial/\partial t)\,\rho(l\,|\,F(\cdot, V))$ *be bounded uniformly in* l, $|l| = 1$, *and the joint variation of the family* $\{(\partial/\partial t)\,F(\cdot, v)\}_{v \in V}$ *be bounded. Then*

$$H(I_N, I) \leqslant \frac{(T-t_0)^2}{8N^2}\left(\sup_{|l|=1} \bigvee_{t_0}^{T} \frac{\partial}{\partial t}\rho(l\,|\,F(\cdot, V)) + J\bigvee_{t_0}^{T} \frac{\partial}{\partial t} F(\cdot, V)\right) \tag{13}$$

for every integer N.

The claim is closely related to the more general result in Ivanov [3, Theorem 8.1] and we only sketch the proof, which is similar to that in Sendov and Popov [6].

Proof. From (11) we have

$$H(I_N, I) \leqslant \max_{|l|=1} \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (\rho(l\,|\,F(t, V)) - \langle l, F(t, v_i)\rangle)\, dt,$$

where $v_i \in V$ is such that

$$\langle l\,|\,F(t_i + 0.5h, v_i)\rangle = \rho(l\,|\,F(t_i + 0.5h, V)).$$

Hence, denoting $\xi_i = t_i + 0.5h$, $\varphi(t) = \rho(l\,|\,F(t, V))$, and $\psi(t) = \langle l, F(t, v_i)\rangle$ for $t \in [t_i, t_{i+1})$, and using that $\varphi(\xi_i) = \psi(\xi_i)$ we have

$$\begin{aligned}
H(I_N, I) &\leqslant \max_{|l|=1} \sum_{i=0}^{N-1} \left[\int_0^{h/2} (\varphi(\xi_i - t) + \varphi(\xi_i + t) - \psi(\xi_i - t) - \psi(\xi_i + t))\right] dt \\
&\leqslant \max_{|l|=1} \sum_{i=0}^{N-1} \left[\int_0^{h/2} (\varphi(\xi_i - t) + \varphi(\xi_i + t) - 2\varphi(\xi_i))\, dt \right.\\
&\qquad \left. - \int_0^{h/2} (\psi(\xi_i - t) + \psi(\xi_i + t) - 2\psi(\xi_i))\, dt\right] \\
&\leqslant \frac{h^2}{8} \sup_{|l|=1} \sum_{i=0}^{N-1} \left(\bigvee_{t_i}^{t_{i+1}} \dot{\varphi} + \bigvee_{t_i}^{t_{i+1}} \left\langle l, \frac{\partial}{\partial t} F(\cdot, v_i)\right\rangle\right)
\end{aligned}$$

which implies (13). Q.E.D.

The last term in (13) is often easy to be estimated in a more constructive way. For instance, if $F(\cdot, v)$ is differentiable and $(\partial/\partial t)\,F(\cdot, v)$ is Lipschitz continuous with a constant L' for every $v \in V$, then

$$J\bigvee_{t_0}^{T} \frac{\partial}{\partial t} F(\cdot, V) \leqslant L'(T - t_0).$$

The next step will be to estimate the variation of $(\partial/\partial t)\,\rho(l\,|\,F(\cdot, V))$, which will be done also in the terms of the joint variation of the family $\{(\partial/\partial t)\,F(\cdot, v)\}_{v\in V}$. We use the following more general result.

PROPOSITION 2. *Let $g(\cdot, v)$: $[t_0, T] \to \mathfrak{R}$, $v \in V$, be a family of functions (V is an arbitrary set) such that*

(i) $|g(t_1, v) - g(t_2, v)| \leqslant L|t_1 - t_2|$ *for every* $t_1, t_2, \in [t_0, T]$ *and* $v \in V$;

(ii) *the family* $\{(\partial g/\partial t)(\cdot, v)\}_{v\in V}$ *is of bounded joint variation.*

Then the function $g(t) = \sup\{g(t, v); v \in V\}$ is Lipschitz continuous and

$$\bigvee_{t_0}^{T} \dot{g}(\cdot) \leqslant 2L + 2J \bigvee_{t_0}^{T} \frac{\partial}{\partial t} g(\cdot, V).$$

Proof. From Clarke [2, Theorem 2.8.6] it follows that $g(\cdot)$ is Lipschitz continuous and the Clarke's generalized derivative $\partial g(t)$ satisfies

$$\partial g(t) \subset \operatorname{cl}\operatorname{co}\{\lim_{i\to+\infty} g'_t(t_i, u_i); t_i \to t, g(t, u_i) \to g(t)\} = [a(t), b(t)], \tag{14}$$

where in the right-hand side it stands the closed convex hull of the condensation points of all sequences $g'_t(t_i, u_i)$, for which this derivative exists and $t_i \to t$, $g(t, u_i) \to g(t)$. Since the mapping $t \to \partial g(t)$ is u.s.c. [2, Proposition 2.1.5] we can extend for convenience $\dot{g}(\cdot)$ to the whole interval $[t_0, t]$ in such a way that $\bigvee_{t_0}^{T} \dot{g}(\cdot)$ does not change and $\dot{g}(t) \in [a(t), b(t)]$ for every $t \in [t_0, T]$. Actually, at a point t where $\dot{g}(t)$ does not exist one can define it as an arbitrary condensation point of a sequence $g(t_i)$ such that $t_i \to t$ and $\dot{g}(t_i)$ exists.

Take arbitrarily $t_0 \leqslant t_1 < \cdots < t_{r+1} \leqslant T$ and consider

$$w = \sum_{i=1}^{r} |\dot{g}(t_i) - \dot{g}(t_{i+1})|.$$

Without any restriction we can assume that $\dot{g}(t_i) - \dot{g}(t_{i+1})$ changes its sign alternatively with i, since otherwise one can remove some of the points $\{t_i\}$ ensuring this property without changing w. Thus

$$w = \left| \sum_{i=1}^{r} (-1)^i (\dot{g}(t_i) - \dot{g}(t_{i+1})) \right|. \tag{15}$$

Since $\dot{g}(t) \in [a(t), b(t)]$ for every $t \in [t_0, T]$ we have

$$a(t_i) - b(t_{i+1}) \leqslant \dot{g}(t_i) - \dot{g}(t_{i+1}) \leqslant b(t_i) - a(t_{i+1}).$$

Hence using (15) we estimate

$$w \leqslant \sum_{i=1}^{r} |c_i - c_{i+1}|,$$

where c_i is either $a(t_i)$ or $b(t_i)$, $i = 1, ..., r+1$.

Denote

$$\delta_1 = 0.5(t_2 - t_1), \qquad \delta_{r+1} = 0.5(t_{r+1} - t_r),$$
$$\delta_i = 0.5 \min \{t_i - t_{i-1}, t_{i+1} - t_i\}, i = 2, ... r.$$

Since the c_i are extreme points of the right-hand side of (14), for every $\varepsilon > 0$ and each $i = 1, ..., r+1$ there exist $\theta_i \in [t_0, T]$ and $v_i \in V$ such that $g'_t(\theta_i, v_i)$ exists and

$$|\theta_i - t_i| \leqslant \varepsilon \delta_i^2 \tag{16}$$

$$g(t_i, v_i) \geqslant g(t_i) - \varepsilon \delta_i^2 \tag{17}$$

$$|g'_t(\theta_i, v_i) - c_i| \leqslant \varepsilon \delta_i. \tag{18}$$

Using (18) we obtain

$$w \leqslant \sum_{i=1}^{r} |g'_t(\theta_i, v_i) - g'_t(\theta_{i+1}, v_{i+1})| + 2\varepsilon \sum_{i=1}^{r+1} \delta_i \leqslant w_1 + 2\varepsilon(T - t_0). \tag{19}$$

Consider the function

$$\psi(s) = -g(s, v_i) + g(s, v_{i+1}) \qquad \text{for} \quad t \in [t_i, t_{i+1}].$$

From (17) we have

$$\psi(t_i) \leqslant -g(t_i) + \varepsilon \delta_i^2 + g(t_i) = \varepsilon \delta_i^2$$
$$\psi(t_{i+1}) \geqslant -g(t_{i+1}) + g(t_{i+1}) - \varepsilon \delta_{i+1}^2 = -\varepsilon \delta_{i+1}^2.$$

Hence

$$-\varepsilon(\delta_i^2 + \delta_{i+1}^2) \leqslant \int_{t_i}^{t_{i+1}} \dot{\psi}(s)\, ds \leqslant 2\varepsilon(\delta_i^2 + \delta_{i+1}^2) L + \int_{t_i + \varepsilon\delta_i^2}^{t_{i+1}}$$
$$\leqslant 2\varepsilon(\delta_i^2 + \delta_{i+1}^2) L + (t_{i+1} - t_i - \varepsilon(\delta_i^2 + \delta_{i+1}^2)) \operatorname{essup}$$
$$\{\dot{\psi}(s); s \in [t_i + \varepsilon\delta_i^2, t_{i+1} - \varepsilon\delta_{i+1}^2]\}$$

and there is $\tau_i \in (t_i + \varepsilon\delta_i^2, t_{i+1} - \varepsilon\delta_{i+1}^2)$ at which $\dot{\psi}(\tau_i)$ exists and

$$\dot{\psi}(\tau_i) \geqslant -\frac{2\varepsilon(1 + 2L)(\delta_i^2 + \delta_{i+1}^2)}{t_{i+1} - t_i} \geqslant -\varepsilon(1 + 2L)(\delta_i + \delta_{i+1}).$$

Thus we can estimate

$$w_1 \leqslant \sum_{i=1}^{r} |\dot{\psi}(\tau_i)| + \sum_{i=1}^{r} |g'_t(\theta_i, v_i) - g'_t(\tau_i, v_i)|$$
$$+ \sum_{i=1}^{r} |g'_t(\theta_{i+1}, v_{i+1}) - g'_t(\tau_i, v_{i+1})|$$

and since $\theta_i \leqslant \tau_i \leqslant \theta_{i+1}$ we have

$$w_1 \leqslant \sum_{i=1}^{r} \dot{\psi}(\tau_i) + \varepsilon(T - t_0)(2 + 4L) + J \bigvee_{t_0}^{T} g'_t(\cdot, V) \tag{20}$$

and we can express

$$\sum_{i=1}^{r} \dot{\psi}(\tau_i) = \sum_{i=1}^{r-1} (g'_t(\tau_i, v_{i+1}) - g'_t(\tau_{i+1}, v_{i+1})) + g'_t(\tau_r, v_{r+1}) - g'_t(\tau_1, v_1)$$
$$\leqslant J \bigvee_{t_0}^{T} g'_t(\cdot, \mathrm{V}) + 2L.$$

The last inequality combined with (19) and (20) implies that claim of the proposition, since $\varepsilon > 0$ is arbitrary. Q.E.D.

Propositions 1 and 2 result in the following theorem.

THEOREM 2. *Let conditions* (A1) *and* (A2) *be fulfilled and let the family* $\{(\partial/\partial t) F(\cdot, V)\}_{v \in V}$ *be of bounded joint variation. Then* I *and* I_N *defined by* (10) *and* (11), *correspondingly, satisfy*

$$H(I_N, I) \leqslant \frac{(T - t_0)^2}{8N^2} \left(2L + 3J \bigvee_{t_0}^{T} \frac{\partial}{\partial t} F(\cdot, V) \right), \tag{21}$$

where L *is the Lipschitz constant of* $F(\cdot, v)$ *and* N *is an arbitrary integer.*

Proof. We apply Proposition 2 for $g(t, v) = \langle l, F(t, v) \rangle$ to obtain an estimation of the variation of $(\partial/\partial t)\, \rho(l \mid F(\cdot, V))$ and then (21) follows from Proposition 1. Q.E.D.

4. SOME APPLICATIONS AND OPEN PROBLEMS

We start with an application of Proposition 2 which concerns the numerical integration of functions of max-type or differential equations with such functions in the right-hand side. Namely, consider

$$\int_{t_0}^{T} f(t)\, dt \qquad \text{or} \qquad \dot{x} = g(x, t). \tag{22}$$

where

$$f(t) = \sup\{ f_\alpha(x); \alpha \in \mathscr{A} \}, \qquad g(x, t) = \sup\{ g_\alpha(x, t); \alpha \in \mathscr{A} \},$$

where $\mathscr{A}$ is an arbitrary set. Proposition 2 combined with Theorem 8.1 of [3] implies that if the $f_\alpha(\cdot)$ satisfy similar conditions as in Proposition 2 then any linear composite quadrature formula, which is exact for polynomials of first degree, provides second order approximation to the integral in (22).

A similar fact holds for the differential equation in (22). Let g_α be differentiable and let $\partial g_\alpha/\partial x$ and $\partial g_\alpha/\partial t$ be Lipschitz continuous uniformly in $\alpha \in \mathscr{A}$. Applying Proposition 2 one can find that the solution of an initial value problem for (22) has a Lipschitz continuous first derivative and second derivative with bounded variation. This means that the solution admits a second order global approximation by means of Runge–Kutta or Adams type methods (see Sendov and Popov [6]).

Another field of applications is in control theory, where Theorem 2 provides a basis for second order approximations to either optimal control problems or the guaranteed systems estimation problem (see, e.g., Kurzhanski [4]) for linear systems

$$\dot{x} = A(t)x + B(t)u. \tag{23}$$

Here $x \in \mathfrak{R}^n$ is the state variable and $u \in \mathfrak{R}^m$ is either a control parameter or a disturbance, in both cases known to be bounded: $u(t) \in U \subset \mathfrak{R}^r$. Details in this direction are contained in Veliov [8]. Here we indicate an application of Theorem 1 for approximation of the reachable set of (23) on a given interval $[t_0, T]$, starting from x_0:

$$R(t_0, T; x_0) = \{x(T); x(t_0) = x_0, x(\cdot) \text{ solves (23) for some } \mathscr{L}_1 \text{ selection } u(\cdot) \text{ of } U\}.$$

The reachable set can be presented by an integral in the form of (1) with $F(t) = \Phi(T, t)\, B(t)$, where $\Phi(t, s)$ is the fundamental matrix solution of (23) normalized at $t = s$.

We suppose that U is a coordinate polyhedron in $\mathfrak{R}^m$ and since the integral (1) decomposes to a sum of m independent integrals we may assume without any restriction that $m = 1$ and $U = [-1, 1]$. Thus we are interested in

$$\int_{t_0}^{T} f(t)[-1, 1]\, dt,$$

where $f(t)=\Phi(T, t)\,B(t)$. Supposing that $f(\cdot)$ is twice continuously differentiable and applying Theorem 1 we obtain that

$$H\left(\int_{t_0}^{T} f(t)[-1,1]\,dt, \left\{\operatorname{co}\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}} \left(f_k^0 u_k(t)+(t-t_k)f_k^1 u_k(t)\right.\right.\right.$$
$$\left.\left.\left.+(t-t_k)^2 f_k^2 u_k(t)\right)dt;\, u_k(\cdot)\in\mathscr{U}_1(t_k, t_{k+1})\right\}\right)\leqslant ch^3, \tag{24}$$

where $f_k^0=f(t_k)$, $f_k^1=f'(t_k)$, $f_k^2=0.5f''(t_k)$, and as above $t_k=t_0+kh$, $h=(T-t_0)/N$. Presenting the single jump point of $u_k(\cdot)$ on $[t_k, t_{k+1}]$ in the form $t_k+h\alpha_k$, $\alpha_k\in[0,1]$ one can express the sum in (24) as

$$h\operatorname{co}\sum_{k=0}^{N-1}\{\pm(f_k^0(1-2\alpha_k)+hf_k^1(\tfrac{1}{2}-\alpha_k^2)+h^2 f_k^2(\tfrac{1}{3}-\tfrac{2}{3}\alpha_k^3));\,\alpha_k\in[0,1]\}. \tag{25}$$

Thus the reachable set of (23) is approximated by means of (the convex hull of) $2N$ cubic curves in $\mathfrak{R}^n$ with accuracy $0(1/N^3)$. We stress the fact that each summand in (25) approximates the corresponding integral $\int_{t_k}^{t_{k+1}} f(t)[-1,1]\,dt$ with accuracy also $O(1/N^3)$. Thus we observe again the effect of nonaccumulation of the errors of the summands. We mention that in order to ensure accuracy $O(1/N^4)$ of every summand (which also gives error $O(1/N^3)$ of the sum) one needs to use selections of U having two jumps in every interval $[t_k, t_{k+1}]$. This straightforward third order approximation of the reachable set would involve pieces of 2-dimensional manifolds parametrized by means of cubic functions (instead of the cubic curves used in (25)), which is a principle complication.

Let us return again to the control system (23) considering it as a differential inclusion

$$\dot{x}\in A(t)\,x+B(t)\,U. \tag{26}$$

We note that every absolutely continuous function $x(\cdot)$ on $[t_0, T]$ which satisfies (26) for a.e. t is a solution to (26). Moreover, the set of solutions of (26) coincides with the set of solutions of (23) when u in (23) runs over all $\mathscr{L}_1$-selections of U.

Supposing that $A(\cdot)$ and $B(\cdot)$ have Lipschitz continuous derivatives and that U is compact and convex, one can interpret the result of Theorem 2 as follows.

Property A. Using constant selections of U on each interval $[t_k, t_{k+1}]$, $k=0,\ldots, N-1$, one can approximate the reachable set of (26) with accuracy $O(1/N^2)$.

We mention that the "straightforward" estimation of this accuracy is $O(1/N)$, which holds also in the case of a nonlinear control system (c.f.

Nikol'skii [5]). On the other hand Property A was proven in Veliov [7] also for the differential inclusion

$$\dot{x} \in f(x, t) + B(x, t)\, U, \qquad x(t_0) \in X_0 \tag{27}$$

under natural smoothness conditions and the additional (restrictive) hypothesis that $B(x, t)\, U$ is strongly convex in $\mathfrak{R}^n$ for every x and t. Thus the following problem arises (and it is still open): Which is the class of differential inclusion in the form of (27) for which Property A holds?

As mentioned above the linear differential inclusion (26) with arbitrary convex and compact set U, and the differential inclusion (27) with strongly convex right-hand side belongs to this class.

As a final remark we mention that if (27) possesses the above property then it can be effectively approximated by means of set-valued analogs to second (and higher) order finite difference schemes. For simplicity we show this in the case of the simplest Runge–Kutta formula with third order local accuracy. Namely, suppose the following:

(i) f and B are differentiable and the first derivatives with respect to x and t are Lipschitz continuous;

(ii) $U \subset \mathfrak{R}^m$ is convex and compact, $X_0 \subset \mathfrak{R}^n$ is compact;

(iii) there is a compact set $Z \subset \mathfrak{R}^n$ such that every trajectory $x(\cdot)$ of (27) which exists on some interval $[t_0, T]$ satisfies $x(t) \in \operatorname{int} Z$, $t \in [t_0, T]$.

Consider the discrete inclusion

$$x_{k+1} \in x_k + 0.5h\{ g(x_k, t_k, u) + g(x_k + hg(x_k, t_k, u), t_{k+1}, u) : u \in U\},$$
$$x_0 \in X_0, \tag{28}$$

where $g(x, t, u) = f(x, t) + B(x, t)\, u$ and as above $t_k = kh$, $h = (T - t_0)/N$. Let R be the reachable set of (27) on $[t_0, T]$ and R_N be the reachable set of (28):

$$R_N = \{x_N;\ \text{there are } x_0, \ldots, x_N \text{ satisfying (28)}\}.$$

THEOREM 3. *Let* (27) *possess Property* A. *Then there is a constant c such that*

$$H(R_N, R) \leqslant c/N^2 \tag{29}$$

for every integer N.

Proof. Denote by

$$\bar{R}_N = \{x(T);\ \dot{x}(t) = f(x(t), t) + B(x(t), t)\, u(t),\ x(t_0) \in X_0,\ u(\cdot) \in \mathscr{U}_0(t_k, t_{k+1}),\ k = 0, \ldots, N-1\}$$

the reachable set of (27) in the class of those selctions of U which are constant on every interval $[t_k, t_{k+1}]$. By Property A

$$H(R, \bar{R}_N) \leqslant c_1/N^2$$

for an appropriate constant c_1. The estimation

$$H(\bar{R}_N, R_N) \leqslant c_2/N^2$$

follows in a standard way from the classical estimation of the local accuracy of the refined Euler method for differential equations and the assumptions (i) and (iii). Q.E.D.

REFERENCES

1. R. AUMANN, Integrals of set valued functions, *J. Math. Anal. Appl.* **12** (1965), 1–12.
2. F.CLARKE, "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
3. K. IVANOV, New estimates of error of quadrature formulae, formulae of numerical differentiation and integration, *Anal. Math.* **6** (1980), 281–303.
4. A. KURZHANSKI, "Control and Observation under Uncertainty," Nauka, Moscow, 1978. [In Russian]
5. M. NIKOL'SKII, On the approximation of the reachable set of a control system, *Mat. Zametki* **41** (1987), 71–76. [In Russian]
6. B. SENDOV AND V. POPOV, "Averaged Moduli of Smoothness," Wiley, New York, 1988.
7. V. VELIOV, Second order discrete approximations to strongly convex differential inclusion, *Systems Control Lett.* **13** (1389), 263–269.
8. V. VELIOV, Second order discrete approximations to linear differential inclusions, *SIAM J. on Numerical Analysis* **29** (2) (1992), 439–451.