JOURNAL OF APPROXIMATION THEORY 26, 301–316 (1979)

Approximation of Convex Set-Valued Functions*

RICHARD A. VITALE

Claremont Graduate School, Claremont, California 91711

Communicated by Oved Shisha

Received January 13, 1978

Approximation of set-valued functions is introduced and discussed under a convexity assumption. In particular, a theorem on positive linear operators is given.

1. INTRODUCTION

Let $\mathbb{K}$ denote the collection of nonempty, compact subsets of $\mathbb{R}^d$. With the introduction of the Hausdorff metric, given by

$$h(K_1, K_2) = \inf\{\epsilon > 0 \mid K_1 \subseteq K_2 + \epsilon B, K_2 \subseteq K_1 + \epsilon B\}, \tag{1.1}$$

$\mathbb{K}$ can be regarded as a complete, separable, and locally compact metric space. Here B is the closed unit ball in $\mathbb{R}^d$, scalar multiplication of sets is defined in the usual way, and "+" denotes (Minkowski) addition of sets.

A *set-valued function F* is a map from [0, 1] into $\mathbb{K}$. Such maps (and more general versions) arise in a variety of contexts, including optimal control theory, mathematical economics, and probability theory. Analytical investigations have followed several lines, including the construction of a differential calculus (see, for instance, Artstein [2], Aumann [4], and Matheron [8]) and the investigation of *selections*, namely vector-valued functions f: $[0, 1] \to \mathbb{R}^d$ such $f(t) \in F(t)$ for each t (Wagner [18] provides an extensive survey of this area).

Our purpose here is to present some initial investigations into the possibilities of an *approximation theory* for set-valued functions. We take our lead from traditional notions and begin by posing the question, is it possible to approximate a given F by a "simpler" one? More concretely, we may look for *linear* approximants of the form

$$\sum_{j=0}^{n} \varphi_j K_j = \varphi_0 K_0 + \cdots + \varphi_n K_n , \tag{1.2}$$

* This work was done at the Mathematics Research Center (Madison) and supported by the United States Army under Contract DAAG29-75-C-0024.

0021-9045/79/080301-16$02.00/0

where the K_j are fixed elements of $\mathbb{K}$ and the φ_j are scalar valued maps defined on [0, 1]. A new ingredient in this traditional formulation is that (1.2) must be treated with some care in combining terms. Note that, although $\{0\}$ is the identity for addition of sets, i.e.,

$$K + \{0\} = K,$$

generally no additive inverse exists (one can easily verify that $K + K^I = \{0\}$ cannot be solved for K^I unless K reduces to a point). Moreover, the distributive law

$$\alpha K + \beta K = (\alpha + \beta)K$$

generally fails to hold (consider, for instance, the case when $K = \{0, 1\} \subseteq \mathbb{R}^1$). It is true that a restricted version holds for *convex* K, namely

$$\alpha K + \beta K = (\alpha + \beta)K \qquad \text{for} \quad \alpha, \beta \geqslant 0. \tag{1.3}$$

This suggests that the class of convex-valued F may be an appropriate place in which to begin considering approximation, and we will devote our discussion to this case.

An outline of the development is as follows. In Section 2, we present notation and generally well-known preliminaries. We take up Bernstein approximation in Section 3 to show the possibility of uniform approximation by linear approximants of polynomial type. We then make a brief excursion into the nonconvex case. Section 4 presents our main result concerning convergence of positive, linear operators. In Section 5 we return to Bernstein approximation to examine some of its other features.

2. $\mathbb{K}_c$

We denote by $\mathbb{K}_c$ the collection of elements of $\mathbb{K}$ which are also convex. We summarize some properties of $\mathbb{K}_c$ which can be found in standard references (see, for instance, Eggleston [7], Rockafellar [10], and Valentine [14]).

$\mathbb{K}_c$ is closed under addition and scalar multiplication of sets and enjoys the distributive property (1.3). $\mathbb{K}_c$ inherits its metric from $\mathbb{K}$ as a closed, separable and locally compact subspace. Given an element K, we may form its convex hull con K which is in $\mathbb{K}_c$. The map $K \to \operatorname{con} K$ is continuous and satisfies additionally

$$\operatorname{con}(\alpha K_1 + \beta K_2) = \alpha \operatorname{con} K_1 + \beta \operatorname{con} K_2$$

for $\alpha, \beta \geqslant 0$.

To each $K \in \mathbb{K}_c$ is associated its *support function*, given by

$$s(p, K) = \max\{p \cdot k \mid k \in K\}, \qquad p \in \mathbb{R}^d, \qquad \| p \| = 1. \tag{2.1}$$

One may consider the support function to give a convenient parameterization of the family of supporting hyperplanes to a set. A set $K \in \mathbb{K}_c$ and a point not in K can always be separated by some hyperplane, and this leads to the useful equivalence

$$K_1 \subseteq K_2 \Leftrightarrow s(p, K_1) \leqslant s(p, K_2) \qquad \forall p \tag{2.2}$$

and consequent uniqueness of support functions

$$K_1 = K_2 \Leftrightarrow s(p, K_1) = s(p, K_2) \qquad \forall p.$$

As a function of p, $s(p, K)$ is continuous; indeed the Schwarz inequality, together with (2.1), yields the uniform bound $| s(p_2, K) - s(p_1, K)| \leqslant \| p_2 - p_1 \| \| K \|$. Here we have used the symbol $\| K \|$ to denote the *norm* of K which is equal to $\max\{\| k \| \mid k \in K\}$ and, equivalently, $d(\{0\}, K)$.

Evidently we may use the map $K \mapsto s(\cdot, K)$ to embed $\mathbb{K}_c$ in the Banach space B_d of continuous functions defined on the surface of the unit $\mathbb{R}^d$ ball. Important structure is preserved under this mapping:

$$s(\cdot, \alpha K) = \alpha s(\cdot, K), \qquad \alpha \geqslant 0, \tag{2.3}$$

$$s(\cdot, K_1 + K_2) = s(\cdot, K_1) + s(\cdot, K_2), \tag{2.4}$$

$$h(K_1, K_2) = \| s_1 - s_2 \| \qquad \text{(uniform norm)}, \tag{2.5}$$

$(\| K \| = \| s(\cdot, K)\|)$.

Let us indicate briefly how (2.5) comes about: The support function of B is identically 1 so that (2.3) and (2.4) imply $s(p, K_2 + \epsilon B) = s(p, K_2) + \epsilon$. Together with (2.2) this yields $K_1 \subseteq K_2 + \epsilon B$ iff

$$s(p, K_1) \leqslant s(p, K_2) + \epsilon \qquad \text{for all } p.$$

The analogous expression holds iff $K_2 \subseteq K_1 + \epsilon B$. For both inclusions to hold, we must have

$$| s(p, K_1) - s(p, K_2)| \leqslant \epsilon \qquad \text{for all } p. \tag{2.6}$$

The infimum of all $\epsilon > 0$ satisfying (2.6) is at once $h(K_1, K_2)$ and $\| s_1 - s_2 \|$ (see (1.1)). Taking in particular $K_2 = \{0\}$ yields $\| K \| = \| s(\cdot, K)\|$.

$\mathbb{C}[\mathbb{K}_c]$ and $\mathbb{C}[\mathbb{K}_c]$ will denote the spaces of continuous maps from [0, 1] into $\mathbb{K}$ and $\mathbb{K}_c$, respectively. Given a map $F \in \mathbb{C}[\mathbb{K}]$ we denote the *norm* of F by

$$H(F) = \sup\{\|F(t)\| \mid t \in [0, 1]\}$$

and define the related metric by

$$H(F, G) = \sup\{h(F(t), G(t)) \mid t \in [0, 1]\}.$$

3. Bernstein Approximation

Given a set-valued function F, we define its nth *Bernstein approximant* to be

$$B_n(F; t) = \sum_{j=0}^{n} b_{jn}(t)\, F(j/n), \qquad b_{jn}(t) = \binom{n}{j} t^j(1-t)^{n-j}.$$

It is straightforward to show that this map necessarily lies in $\mathbb{C}[\mathbb{K}]$ and, indeed, in $\mathbb{C}[\mathbb{K}_c]$ if $F \in \mathbb{C}[\mathbb{K}_c]$.

THEOREM 1. *Let $F \in \mathbb{C}[\mathbb{K}_c]$. Then, as $n \to \infty$, $B_n(F;\ )$ converges uniformly to F (i.e., $H(F, B_n(F;\ )) \to 0$).*

Proof. We use the Banach space embedding. $F \in \mathbb{C}[\mathbb{K}_c]$ is equivalent to the continuity of the map from $[0, 1]$ into $\mathbb{B}_d$ given by $t \mapsto s(\cdot, F(t))$. Likewise, a Bernstein approximant of F corresponds to the map $t \mapsto \sum_{j=0}^{n} b_{jn}(t) s(\cdot, F(j/n))$. Hence, it is enough to show the uniform convergence (in $\mathbb{B}_d$) of the latter maps to $t \mapsto s(\cdot, F(t))$. This follows directly from classical arguments (see, for example, Davis [5] transposed to a Banach space setting). ∎

In Section 4, we shall view this result from a more general perspective. For now, let us turn to the case when $F \in \mathbb{C}[\mathbb{K}]$ does not necessarily have convex values. Of course, this does not preclude forming $B_n(F;\ )$ and, indeed, as we shall see, Bernstein approximation asymptotically "convexifies" F.

Let us digress for a moment to consider a simple example. If $K = \{0, 1\} \subseteq \mathbb{R}^1$, then

$$\frac{1}{n} \sum_{j=0}^{n} K = \frac{1}{n} [\underbrace{K + K + \cdots + K}_{n}] = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}$$

and hence $h(1/n \sum_{j=0}^{n} K, \operatorname{con} K = [0, 1]) \to 0$ as $n \to \infty$. This "filling in" of values is typical of what happens when nonconvex sets are summed. The following result quantifies this behavior.

PROPOSITION (Shapley–Folkman; see Arrow and Hahn [1, p. 396]). *Let* $K_j \in \mathbb{K}$, $j = 0, 1,..., n$, *be such that* $\| K_j \| \leqslant M$. *Then*

$$h\left(\sum_{j=0}^{n} K_j, \sum_{j=0}^{n} \operatorname{con} K_j\right) \leqslant M d^{1/2}. \tag{3.1}$$

We use this result to investigate the nonconvex case.

THEOREM 1. *Let* $F \in \mathbb{C}[\mathbb{K}]$. *Then in any subinterval* $[\epsilon, 1 - \epsilon]$, $0 < \epsilon < \frac{1}{2}$, $B_n(F; \ )$ *converges uniformly to* $\operatorname{con} F$ (*here* $(\operatorname{con} F)(t) = \operatorname{con} F(t)$, $t \in [0, 1]$).

Proof. With

$$B_n(F; t) = \sum_{j=0}^{n} b_{jn}(t) F(j/n),$$

we identify $K_j = b_{jn}(t) F(j/n)$ in (3.1). Now

$$\begin{aligned} \| K_j \| &\leqslant \| F(j/n) \| \, | b_{jn}(t) | \\ &\leqslant H(F) \sup\{b_{jn}(t) \mid \epsilon \leqslant t \leqslant 1 - \epsilon, j = 0, 1,..., n\}. \end{aligned}$$

The indicated supremum can be shown to be $O(n^{-1/2})$, so that by the proposition

$$h(B_n(F; t), B_n(\operatorname{con} F; t)) \leqslant H(F)\, O(n^{-1/2})\, d^{1/2}.$$

Theorem 1 applied to $B_n(\operatorname{con} F; \ )$ and the triangle inequality yield the assertion. ∎

We remark that the result cannot be extended to the full interval since at each endpoint, $t = 0, 1$, $B_n(F; t) = F(t)$ independent of n. Moreover, the $O(n^{-1/2})$ bound breaks down at the endpoints.

The convexification of F by Bernstein approximation is undoubtedly related to theories of integration of set-valued functions, which invariably yield integrals with convex values. It would be of interest to make this statement more precise via a general investigation of the behavior of linear operators on set-valued functions. We shall not consider this problem here but instead present another example which shows the difficulty of formulating approximation methods in the nonconvex case.

Let $F(t) \equiv \{0, 1\}$ (a constant set-valued function) be approximated by the piecewise linear scheme

$$\begin{aligned} F_n(t) &= ([\![nt]\!] + 1 - nt)\, F\left(\frac{[\![nt]\!]}{n}\right) + (nt - [\![nt]\!])\, F\left(\frac{[\![nt]\!] + 1}{n}\right), && 0 \leqslant t < 1, \\ &= F(1), && t = 1. \end{aligned} \tag{3.2}$$

Here $F_n(0) = F_n(1) \equiv \{0, 1\}$, whereas the sequence

$$F_n(t) = \{0, nt - [\![nt]\!], [\![nt]\!] + 1 - nt, 1\}$$

even fails to *converge* for any $t \in (0, 1)$.

4. A Convergence Result

Using (2.2) and the positivity of Bernstein approximation of real-valued functions, we see that for $F, G \in \mathbb{C}[\mathbb{K}_c]$

$$F \subseteq G \text{ (that is, } F(t) \subseteq G(t), \forall t) \Rightarrow B_n(F;\ ) \subseteq B_n(G;\ ) \qquad \forall n.$$

As in the real case, this suggests that a wider class of approximation methods may possess similar convergence properties. Let us agree to call a map $T: \mathbb{C}[\mathbb{K}_c] \to \mathbb{C}[\mathbb{K}_c]$ $\mathbb{K}_c$*-linear* if $T(\alpha F + \beta G) = \alpha TF + \beta TG$, $\forall \alpha, \beta \geqslant 0$ and $\forall F, G \in \mathbb{C}[\mathbb{K}_c]$, and $\mathbb{K}_c$*-positive* if $F \subseteq G \Rightarrow TF \subseteq TG$, $\forall F, G \in \mathbb{C}[\mathbb{K}_c]$ (note the restriction $\alpha, \beta \geqslant 0$, which, as noted above, is appropriate for convex sets). We then have the following result for such maps.

THEOREM 2. *Let $\{T_n\}$ be a sequence of $\mathbb{K}_c$-linear, $\mathbb{K}_c$-positive maps. In order that $T_nF \to F$ for each $F \in \mathbb{C}[\mathbb{K}_c]$, it is necessary and sufficient that*

(i) $T_nF^{(i)} \to F^{(i)}$, $i = 0, 1, 2$ *where* $F^{(i)}(t) = t^iB$ *and*

(ii) $\sup\{H(T_nF, F) \mid F(t) \equiv K, \|K\| = 1\} \to 0$.

Let us remark briefly on the hypotheses of the theorem before proceeding. Condition (i) is reminiscent of the vector-valued formulation and is perhaps even more striking here in that only a fixed *shape* (i.e., B, the closed unit ball) is involved. Condition (ii) asserts that the T_n behave uniformly well when applied to "constants" (including the case $F = F^{(0)}$ from (i)).

We first show necessity of the conditions: (i) is obvious. As for (ii), suppose the contrary. Then there is an $\epsilon > 0$ and a sequence of K_{n_j} such that $H(T_{n_j}K_{n_j}, K_{n_j}) \geqslant \epsilon$ (here we have abused notation slightly to let K_{n_j} stand for F_{n_j} where $F_{n_j}(t) \equiv K_{n_j}$). Local compactness of $\mathbb{K}_c$ and the uniform normalization $\|K_{n_j}\| = 1$ assure the existence of a convergent subsequence of the K_{n_j}. Without loss of generality, suppose that $K_n \to K_\infty$. Then by the triangle inequality,

$$H(T_nK_n, K_n) \leqslant H(T_nK_n, T_nK_\infty) + H(T_nK_\infty, K_\infty) + H(K_\infty, K_n).$$

Now $\epsilon_n = H(K_\infty, K_n) \to 0$. Moreover, the twin inclusions

$$K_n \subseteq K_\infty + \epsilon_n B,$$
$$K_\infty \subseteq K_n + \epsilon_n B$$

together with the properties of T_n imply

$$T_n K_n \subseteq T_n K_\infty + \epsilon_n T_n B$$

and

$$T_n K_\infty \subseteq T_n K_n + \epsilon_n T_n B$$

so that $H(T_n K_n, T_n K_\infty) \leqslant \epsilon_n H(T_n B) \to 0$. Hence $\lim H(T_n K_\infty, K_\infty) \geqslant \epsilon$, but this violates our assumption.

The proof of sufficiency is more involved and will require some preparation. We begin by formulating a quantitative result for families of real-valued functions. This will then be adapted to our needs by invoking the Banach space embedding.

Let P be an indexing set and let Σ denote the collection of all $\sigma = \langle \sigma_p \rangle$, $p \in P$. Here we have denoted by $\langle \sigma_p \rangle$ a bounded equicontinuous family of real-valued functions defined on $[0, 1]$. That is, given $\sigma = \langle \sigma_p \rangle$, each $\sigma_p \in C[0, 1]$ and

(i) $\exists M_\sigma$ such that $\sup_p \| \sigma_p \| \leqslant M_\sigma < \infty$, and

(ii) the modulus of continuity $\omega_\sigma(\delta) = \sup_{|t-x| \leqslant \delta} \sup_p | \sigma_p(t) - \sigma_p(x)|$ satisfies $\omega_\sigma(0+) = 0$.

Σ is a normed linear space under the definitions

$$\alpha\sigma^{(1)} + \beta\sigma^{(2)} = \langle \alpha\sigma_p^{(1)} + \beta\sigma_p^{(2)} \rangle$$

and

$$\| \sigma \| = \sup_t \sup_p | \sigma_p(t)|.$$

Moreover, we can define a partial ordering by

$$\sigma^{(1)} \prec \sigma^{(2)} \Leftrightarrow \sigma_p^{(1)}(t) \leqslant \sigma_p^{(2)}(t) \qquad \forall p \in P, \quad \forall t \in [0, 1].$$

Now let us consider a subspace $\Sigma_0 \subseteq \Sigma$ and a map $L: \Sigma_0 \to \Sigma$. We say that L is *linear* if $L(\alpha\sigma^{(1)} + \beta\sigma^{(2)}) = \alpha L\sigma^{(1)} + \beta L\sigma^{(2)}$ and *positive* if $\sigma^{(1)} \prec \sigma^{(2)} \Rightarrow L\sigma^{(1)} \prec L\sigma^{(2)}$.

For convenience, we call Σ_0 *full* if the following conditions hold

(i) For $i = 0, 1, 2$, ${}_i\sigma \in \Sigma_0$, where ${}_i\sigma_p(t) = t^i \; \forall p \in P$, $\forall t \in [0, 1]$ (note that, since Σ_0 is a subspace, this implies that ${}_{[x]}\sigma \in \Sigma_0$, where ${}_{[x]}\sigma_p(t) = (t - x)^2 \; \forall p \in P, \forall t \in [0, 1], x \in [0, 1]$ fixed).

(ii) If $\sigma = \langle \sigma_p \rangle \in \Sigma_0$, then for each fixed $x \in [0, 1]$, ${}_{(x)}\sigma \in \Sigma_0$, where ${}_{(x)}\sigma_p(t) = \sigma_p(x) \;\; \forall p \in P, \;\; \forall t \in [0, 1]$, (loosely, Σ_0 must contain enough constants).

Finally, we define

$$\gamma(\sigma, L) = \sup_t \sup_p |[L\, {}_{(t)}\sigma]_p(t) - \sigma_p(t)|. \tag{4.1}$$

We are now prepared to state a uniform bound.

PROPOSITION. *Let Σ_0 be full and let $L\colon \Sigma_0 \to \Sigma$ be a positive, linear map. Then for each $\sigma \in \Sigma_0$*

$$\| \sigma - L\sigma \| \leqslant \omega_\sigma(\mu)[\| L_0\sigma \| + 1] + \gamma(\sigma, L),$$

where

$$\mu^2 = \sup_x \sup_p |[L\, {}_{[x]}\sigma]_p(x)|.$$

Proof. We follow an argument of Shisha and Mond [13], who have developed a similar quantitative estimate in the case where P has a single element.

Fix $\sigma = \langle \sigma_p \rangle \in \Sigma_0$. Then, for each $p \in P$ and $\delta > 0$,

$$| \sigma_p(t) - \sigma_p(x)| \leqslant \omega_\sigma(\delta) \left[1 + \frac{(t - x)^2}{\delta^2}\right].$$

Consider one of the two associated inequalities, for example,

$$\sigma_p(t) \leqslant \omega_\sigma(\delta) \left[1 + \frac{(t - x)^2}{\delta^2}\right] + \sigma_p(x).$$

Regarding x as fixed, we see that this is equivalent to

$$\sigma \prec \omega_\sigma(\delta) \left[{}_0\sigma + \frac{1}{\delta^2}\, {}_{[x]}\sigma\right] + {}_{(x)}\sigma$$

and hence

$$L\sigma \prec \omega_\sigma(\delta) \left[L\, {}_0\sigma + \frac{1}{\delta^2} L\, {}_{[x]}\sigma\right] + L\, {}_{(x)}\sigma.$$

The opposite ordering is similar. We take pth components evaluated at x and combine the two resulting inequalities to get

$$|[L\sigma]_p(x) - [L\,{}_{(x)}\sigma]_p(x)| \leqslant \omega_\sigma(\delta) \left| [L\,{}_0\sigma]_p(x) + \frac{1}{\delta^2} [L\,{}_{[x]}\sigma]_p(x) \right|.$$

By assumption,

$$|[L\,{}_{(x)}\sigma]_p(x) - \sigma_p(x)| \leqslant \gamma(\sigma, L)$$

and two applications of the triangle inequality yield

$$|[L\sigma]_p(x) - \sigma_p(x)| \leqslant \omega_\sigma(\delta) \left[|[L\,{}_0\sigma]_p(x)| + \frac{1}{\delta^2} |[L\,{}_{[x]}\sigma]_p(x)| \right] + \gamma(\sigma, L).$$

Taking $\sup_x \sup_p$ on each side yields

$$\| L\sigma - \sigma \| \leqslant \omega_\sigma(\delta) \left[\| L\,{}_0\sigma \| + \frac{1}{\delta^2} \mu^2 \right] + \gamma(\sigma, L).$$

If $\mu > 0$, we take $\delta = \mu$ and are done. If $\mu = 0$, then a limiting argument (see [13]) similarly yields the assertion. ▮

COROLLARY. *Let Σ_0 be full and, for each $n = 1, 2,\ldots$, let L_n be a positive, linear map taking Σ_0 into Σ. If $L_n\,{}_i\sigma \to {}_i\sigma$ for $i = 0, 1, 2$, then, for each $\sigma \in \Sigma_0$, $\gamma(\sigma, L_n) \to 0$ implies $L_n\sigma \to \sigma$.*

Proof. In view of the proposition, we only have to show that ${\mu_n}^2 = \sup_x \sup_p |[L_n\,{}_{[x]}\sigma]_p(x)| \to 0$. Note that each component of ${}_{[x]}\sigma$ is $t^2 - 2xt + x^2$ (here t is the free variable) or equivalently

$${}_{[x]}\sigma = {}_2\sigma - 2x\,{}_1\sigma + x^2\,{}_0\sigma.$$

We apply L_n, take pth components, and evaluate at x:

$$[L_n\,{}_{[x]}\sigma]_p(x) = [L_n\,{}_2\sigma]_p(x) - 2x[L_n\,{}_1\sigma]_p(x) + x^2[L_n\,{}_0\sigma]_p(x).$$

Adding and subtracting $2x^2$ appropriately on the right and taking absolute values yields the bound

$$\begin{aligned} |[L_n\,{}_{[x]}\sigma]_p(x)| \leqslant{}& |[L_n\,{}_2\sigma]_p(x) - x^2| + 2\,|[L_n\,{}_1\sigma]_p(x) - x| \\ &+ |[L_n\,{}_0\sigma]_p(x) - 1|. \end{aligned} \tag{4.2}$$

Operating with $\sup_x \sup_p$ on each side of (4.2) then gives

$$\mu_n^2 \leqslant \| L_n\, {}_2\sigma - {}_2\sigma \| + 2 \| L_n\, {}_1\sigma - {}_1\sigma \| + \| L_n\, {}_0\sigma - {}_0\sigma \|$$

and, by assumption, each of the three terms on the right tends to zero. ∎

We now adapt these considerations to the proof of sufficiency in Theorem 3. Recalling the identification between a set $K \in \mathbb{K}_c$ and its support function $s(p, K)$, we see that an $F \in \mathbb{C}[\mathbb{K}_c]$ can similarly be identified with its family of support functions

$$F \leftrightarrow s(F) = \langle s(p, F(\cdot)) \rangle;$$

where the indexing set is $P = \{p \mid \| p \| = 1\}$. Now

$$\sup_p | s(p, F(t))| = \| F(t) \| \leqslant H(F) < \infty$$

and

$$\sup_{|t-x| \leqslant \delta} \sup_p | s(p, F(t)) - s(p, F(x))| = \omega_F(\delta),$$

where ω_F is the modulus of continuity of F defined in the obvious way. Hence $s(F) \in \Sigma$. In fact, the collection of all $s(F)$ form a positive, convex cone C in Σ by virtue of the identification

$$\alpha F + \beta G \leftrightarrow \alpha s(F) + \beta s(G), \qquad \alpha, \beta \geqslant 0.$$

A $\mathbb{K}_c$-linear operator $T: \mathbb{C}[\mathbb{K}_c] \to \mathbb{C}[\mathbb{K}_c]$ induces a natural map $L: C \to C$ via

$$Ls(F) \equiv s(TF)$$

which obviously satisfies

$$L[\alpha s(F) + \beta s(G)] = \alpha Ls(F) + \beta Ls(G), \qquad \alpha, \beta \geqslant 0.$$

In order to apply the proposition and corollary, we need to extend the domain of L to a *subspace* of Σ. Accordingly, let Σ_0 be the span of C (i.e., all finite linear combinations of the form $\sum \alpha_i s(F_i)$) and define, for any $s(F)$,

$$L[-s(F)] = -Ls(F) \qquad (= -s(TF)).$$

With this done, it is straightforward to verify that $L: \Sigma_0 \to \Sigma$ is linear. Moreover, since $F \subseteq G \Leftrightarrow s(F) < s(G)$, it follows directly that if T is $\mathbb{K}_c$-positive, then L is positive.

Next we show that Σ_0 is full. Obviously, $F^{(i)} \leftrightarrow s(F^{(i)}) = {}_i\sigma \in \Sigma_0$, $i = 0, 1, 2$. Further, ${}_{(x)}\sigma$ is of the form ${}_{(x)}\sigma_p(t) = \sum \alpha_i s(p, F_i(x))$. But each $s(\cdot, F_i(x))$, regarded as a *function of t*, corresponds to a constant set-valued function $F(t) \equiv F_i(x)$. Hence each ${}_{(x)}\sigma \in \Sigma_0$.

Note also that, since for any $F, G \in \mathbb{C}[\mathbb{K}_c]$, $H(F, G) = \| s(F) - s(G)\|$, we have $T_n F^{(i)} \to F^{(i)} \Rightarrow L_n\, {}_i\sigma \to {}_i\sigma$, $i = 0, 1, 2$.

It remains to show that $\gamma \to 0$. If $\sigma = \sum \alpha_i s(F_i)$, then we have (see (4.1))

$$\gamma(\sigma, L_n) \leqslant \sum | \alpha_i | \gamma(s(F_i), L_n)$$

so that it is sufficient to consider $\sigma = s(F)$. We have

$$\gamma(s(F), L_n) = \sup_t \sup_p |[L_n\langle s(p, F(t)) \cdot {}_0\sigma_p\rangle]_p(t) - s(p, F(t))|. \tag{4.3}$$

Let us regard t as fixed and consider the constant set-valued function $F_t(x) \equiv F(t)$, $0 \leqslant x \leqslant 1$. Then

$$\begin{aligned} H(T_n F_t, F_t) &= \| L_n\langle s(p, F(t)) \cdot {}_0\sigma_p\rangle - \langle s(p, F(t)) \cdot {}_0\sigma_p\rangle\| \\ &= \sup_x \sup_p |[L_n\langle s(p, F(t)) \cdot {}_0\sigma_p\rangle]_p(x) - s(p, F(t))|. \end{aligned} \tag{4.4}$$

Using (4.3) and (4.4), we see that

$$\gamma(s(F), L_n) \leqslant \sup_t H(T_n F_t, F_t).$$

By the second assumption of Theorem 3 and the bound $\| F(t)\| \leqslant H(F) < \infty$, the right-hand side tends to zero. This completes the proof of the theorem.

The convergence of Bernstein approximation is easily established in this context. $\mathbb{K}_c$-linearity and -positivity obviously hold. Moreover, for $i = 0, 1, 2$, $B_n(F^{(i)}; t) = B_n(t^i; t) \cdot B$ which establishes convergence for the $F^{(i)}$. Finally, given any constant $F(t) \equiv K$, $B_n(F; t) \equiv F(t)$ so that the derived γ in each case is zero.

It is equally straightforward to establish convergence of the piecewise linear scheme (3.2).

5. Aspects of Bernstein Approximation

In this section we discuss some features of Bernstein approximation which complement the uniform convergence result. Some have been alluded to before and are true for similar approximation schemes.

We begin with some properties which follow directly from the support function embedding and properties of Bernstein approximation in the real-valued case.

PROPOSITION. (i) $K_1 \subseteq F(t) \subseteq K_2$, $\forall t \Rightarrow K_1 \subseteq B_n(F; t) \subseteq K_2$, $\forall t$. *In particular*, $\bigcap_t F(t) \subseteq B_n(F; t) \subseteq \mathrm{con}[\bigcup_t F(t)]\ \forall t$.

(ii) $F(s) \subseteq (\supseteq) F(t)\ \forall s \leqslant t \Rightarrow B_n(F; s) \subseteq (\supseteq) B_n(F; t)\ \forall s \leqslant t$.

(iii) $$F\left(\frac{s+t}{2}\right) \subseteq (\supseteq) \frac{1}{2}[F(s) + F(t)] \qquad \forall s, t$$

$$\Rightarrow B_n\left(F; \frac{s+t}{2}\right) \subseteq (\supseteq) \frac{1}{2}[B_n(F; s) + B_n(F; t)] \qquad \forall s, t.$$

Property (i) is, of course, a special instance of the positivity property. As we have seen, this is a natural extension of the real-valued case. One might, however, try to argue another type of extension. If $f(t) > g(t)\ \forall t$ then the Bernstein approximants of these functions share the same ordering. Alternatively, one could say that nonintersection of graphs is preserved. Accordingly, in the set-valued case, we might ask whether nonintersection—$F(t) \cap G(t) = \varnothing\ \forall t$—is maintained for approximants. The following example, however, shows that this is not generally the case. In the complex plane. let $F(t) = \{e^{2\pi i t}\}$ and let $G \equiv \{z \mid \| z \| \leqslant \epsilon\}$. Then, for each t, $B_n(F; t)$ is a point, which for $t \neq 0, 1$ is of modulus less than 1. Hence if ϵ is sufficiently close to (but smaller than) unity, $F(t) \cap G(t) = \varnothing\ \forall t$ whereas this property fails for the approximants. It is possible, however, to show that nonintersection is *ultimately* preserved in general.

PROPOSITION. *Let* $F(t) \cap G(t) = \varnothing\ \forall t$. *Then, for n sufficiently large*, $B_n(F; t) \cap B_n(G; t) = \varnothing\ \forall t$.

Proof. Let $\epsilon = \inf_t \inf\{\| f - g \| \mid f \in F(t), g \in G(t)\}$. Compactness and continuity ensure that ϵ is strictly positive. The assertion then holds for n such that $\epsilon/2 > \max\{H(F_n, B_n(F_n;\ )), H(G, B_n(G;\ ))\}$. ∎

We turn now to the behavior of approximants when juxtaposed with mappings of the "background space" $\mathbb{R}^d$. If M is a $d \times d$ matrix, we can define a map taking $\mathbb{K}_c$ into $\mathbb{K}_c$ by $K \mapsto MK = \{Mk \mid k \in K\}$. The following easy result is typical.

PROPOSITION. $B_n(MF; t) = MB_n(F; t)$.

In particular, Bernstein approximation commutes with projections. Alternatively, let us consider a continuous one-parameter family of matrices M_t, $0 \leqslant t \leqslant 1$ (continuity can be assumed in any reasonable sense, e.g., in the Euclidean norm). Then $F(t) = M_t K$, for a fixed $K \in \mathbb{K}_c$ is an element of $\mathbb{C}[\mathbb{K}_c]$ (one might think, for instance, of a continuous rotation of a fixed figure). As well as uniform convergence, we have the following.

PROPOSITION. $B_n(M_t; t)K \subseteq B_n(M_t K; t)$.

Proof. As a consequence of the general inclusion $(M_1 + M_2)K \subseteq M_1K + M_2K$ (matrix addition on the left, set addition on the right), we have

$$B_n(M_t\,;t)K = \left(\sum_{j=0}^{n} b_{jn}(t)\, M_{j/n}\right) K \subseteq \sum_{j=0}^{n} b_{jn}(t)\, M_{j/n}K = B_n(M_tK; t). \quad \blacksquare$$

Note incidentally that $B_n(M_t\,;\ )K$ also converges uniformly to M_tK which suggests further comparisons with the convergence of the Bernstein approximants.

An area of particular interest is the behavior of geometric functionals under Bernstein approximation. For instance, given any functional $\varphi\colon \mathbb{K}_c \to \mathbb{R}^1$ which satisfies $\varphi(\alpha K_1 + \beta K_2) = \alpha\varphi(K_1) + \beta\varphi(K_2)$, $\alpha, \beta \geqslant 0$, we have the obvious relation $\varphi(B_n(F; t)) = B_n(\varphi \circ F; t)$. Examples are, for fixed p, $\varphi(K) = s(p, K)$, *the extent of K in the direction p* and $\varphi(K) = s(p, K) + s(-p, K)$, *the width of K in the direction p*. In the plane, $\varphi(K) = \operatorname{per}(K) =$ perimeter of K is another example. Here a convenient parameterization takes $p = (\cos\theta, \sin\theta)$ so that the support function may be regarded as a function of the angle θ. Then $\operatorname{per}(K) = \int_0^{2\pi} s(\theta, K)\, d\theta$ (see, e.g., [14]) and

$$\operatorname{per}(B_n(F; t)) = \sum_{j=0}^{n} b_{jn}(t) \int_0^{2\pi} s(\theta, F(j/n))\, d\theta.$$

Nonlinear functionals naturally require individual attention. It may be possible to invoke classical considerations, as in the following bound for the volume of $B_n(F; t)$, which is a straightforward consequence of the Brunn–Minkowski inequality (see, for instance, [7]).

PROPOSITION. $\operatorname{vol} B_n(F; t) \geqslant [B_n((\operatorname{vol} F)^{1/d}; t)]^d$.

In the plane an explicit expression can be displayed for the area functional. With sufficient smoothness of the support functions,

$$\operatorname{area} B_n(F; t) = \tfrac{1}{2} \sum_{j=0}^{n} \sum_{k=0}^{n} b_{jn}(t)\, b_{kn}(t) \int_0^{2\pi} s(\theta, F(j/n)) \cdot r(\theta, F(k/n))\, d\theta,$$

where $r(\theta, K) = ((\partial^2/\partial\theta^2) + I)\, s(\theta, K)$ (see the discussion in [11] on mixed areas).

6. NOTES

Section 1. The present study was motivated in part by earlier work of the author and colleagues in related areas—in particular, approximation of plane convex sets [6, 9], random sets [3, 17], computational considerations [15], and modeling of tumor growth [16].

Section 2. The Hausdorff metric is evidently a distance of L_∞ type (in the space of support functions). It would be of considerable interest to find an appropriate L_2 formulation.

Section 3. Theorem 1 has many variants along traditional lines. For instance, uniform convergence in a subinterval can be asserted under weaker conditions. Moreover, pointwise convergence rates can be derived for each of the component $s(p, F(t))$.

Theorem 2 suggests that linear methods may not be natural for approximation in $\mathbb{C}[\mathbb{K}]$.

Section 4. The extended development in the text for families of functions can perhaps be avoided by an appeal to the abstract machinery of Banach lattices (see, for instance, [12, esp. Vol. 2]). We have not seen a clear way to do this, and in any case the quantitative formulation given may be of particular use. Although it is not explicitly given in the text, the following bound seems best possible

$$H(TF, F) \leqslant \omega_F(\mu)[H(F^{(0)}) + 1] + \gamma(F, T),$$

where

$$\mu^2 = \sup_x \|[T[(t-x)^2 B)](x)\|$$

and

$$\gamma(F, T) = \sup\{H(TG, G) \mid G(x) \equiv F(t), 0 \leqslant x \leqslant 1, t \text{ fixed}\}.$$

Condition (ii) of Theorem 3 has the equivalent (but apparently weaker) formulation in the plane ($d = 2$):

$$\sup\{H(T_nF, F) \mid F(t) \equiv K, \|K\| = 1, K = \text{a point, line segment, or triangle}\} \to 0.$$

We indicate why this is so. Given any K, we can approximate it in the Hausdorff metric arbitrarily well with a finite sum of the form

$$K_\epsilon = q + \sum \alpha_i \Delta_i \qquad (H(K, K_\epsilon) = \epsilon),$$

where q is a point, the Δ_i are either line segments or triangles containing $\{0\}$, $\|\Delta_i\| = 1$ in each case, and $\alpha_i > 0$ (see, for instance, Yaglom and Boltyanskii [19]). Then

$$H(TK, K) \leqslant H(TK, TK_\epsilon) + H(TK_\epsilon, K_\epsilon) + H(K_\epsilon, K).$$

The last term equals ϵ and the first is bounded above by $\epsilon \| TB \|$. As for the second term, we have the bound

$$\begin{aligned} H(TK_\epsilon, K_\epsilon) &\leqslant H(Tq, q) + \sum \alpha_i H(T\Delta_i, \Delta_i) \\ &\leqslant H(Tq, q) + \sup\{H(T\Delta, \Delta) \mid \| \Delta \| = 1\} \cdot \sum \alpha_i . \end{aligned}$$

Now comparing perimeters of K and K_ϵ we have

$$\operatorname{per}(K_\epsilon) = \sum \alpha_i \operatorname{per}(\Delta_i) \leqslant \operatorname{per}(K) + 2\pi\epsilon.$$

Since $0 \in \Delta_i$ and $\| \Delta_i \| = 1$, we have $\operatorname{per}(\Delta_i) \geqslant 2 \| \Delta_i \| = 2$ (achieved when Δ_i looks like a unit vector) and so

$$\sum \alpha_i \leqslant \tfrac{1}{2}[\operatorname{per}(K) + 2\pi\epsilon] \leqslant \tfrac{1}{2}[2\pi \| K \| + 2\pi\epsilon] = \pi[\| K \| + \epsilon].$$

As for the first term, $0 \in K - q$ so that $\| q \| \leqslant \| K \|$ and hence for $\| q \| > 0$, we have

$$\begin{aligned} H(Tq, q) &= \| q \| \, H(Tq/\| q \|, q/\| q \|) \\ &\leqslant \| K \| \sup\{H(T\tilde{q}, \tilde{q}) \mid \| \tilde{q} \| = 1\}, \end{aligned}$$

(the trivial case $\| q \| = 0$ is, of course, included in the final inequality). Passing to the limit as $\epsilon \searrow 0$, we then have

$$\begin{aligned} H(TK, K) \leqslant{} & [1 + \pi] \| K \| \\ & \cdot \sup\{H(T\Delta, \Delta) \mid \| \Delta \| = 1, \Delta = \text{point, line segment, or triangle}\}. \end{aligned}$$

ACKNOWLEDGMENT

The author wishes to thank Professor Carl de Boor of the Mathematics Research Center and Professor Philip Davis of Brown University for providing comments on a preliminary version of this paper.

REFERENCES

1. K. J. ARROW AND F. H. HAHN, "General Competitive Analysis," Holden–Day, San Francisco, 1971.
2. Z. ARTSTEIN, On the calculus of set-valued functions, *Indiana Univ. Math. J.* **24** (1974), 433–441.
3. Z. ARTSTEIN AND R. A. VITALE, A strong law of large numbers for random compact sets, *Ann. Prob.* **3** (1975), 879–882.

4. R. J. AUMANN, Integrals of set-valued functions, *J. Math. Anal. Appl.* **12** (1965), 1–12.
5. P. J. DAVIS, "Interpolation and Approximation," Blaisdell, New York, 1963.
6. P. J. DAVIS, R. A. VITALE AND E. BEN-SABAR, On the deterministic and stochastic approximation of regions, *J. Approximation Theory* **21** (1977), 60–88.
7. H. G. EGGLESTON, "Convexity," Cambridge Univ. Press, London/New York, 1969.
8. G. MATHERON, "Random Sets and Integral Geometry," Wiley, New York, 1975.
9. D. E. MCCLURE AND R. A. VITALE, Polygonal approximation of plane convex sets, *J. Math. Anal. Appl.* **51** (1975), 326–358.
10. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, N.J., 1970.
11. L. SANTALÓ, "Integral Geometry and Geometric Probability," Addison–Wesley, Reading, Mass., 1976.
12. H. H. SCHAEFER, "Banach Lattices and Positive Operators," Springer–Verlag, New York, 1974.
13. O. SHISHA AND B. MOND, The degree of convergence of linear positive operators, *Proc. Nat. Acad. Sci. U.S.A.* **60** (1968), 1196–1200.
14. F. A. VALENTINE, "Convex Sets," McGraw–Hill, New York, 1964.
15. R. A. VITALE AND A. TARR, An APL package for convex geometry, *in* "Proc. APL-75 Congress, Pisa, 1975," pp. 376–383.
16. R. A. VITALE AND A. TARR, Mathematical models of tumor growth, DAM Report, Brown University, July 1975.
17. R. A. VITALE, Asymptotic area and perimeter of sums of random plane convex sets, MRC Technical Summary Report No. 1770, University of Wisconsin, July 1977.
18. D. H. WAGNER, Survey of measurable selection theorems, *SIAM J. Control Optimization* **15** (1977), 859–903.
19. I. M. YAGLOM AND V. G. BOLTYANSKII, V. G. "Convex Figures," Holt, Rinehart & Winston, New York, 1961.