Math. Z. 122, 63–70 (1971)

A Generalization of Bohman-Korovkin's Theorem

GÜNTER FELBECKER and WALTER SCHEMPP

1. Introduction

Let $I=[0,1]$ be the compact unit interval of the real line $\mathbb{R}$ and let us denote by $\mathscr{C}_{\mathbb{R}}(I)$ the real algebra of continuous real-valued functions on I. We provide $\mathscr{C}_{\mathbb{R}}(I)$ with the topology of uniform convergence on I induced by the Čebyšev norm

$$\mathscr{C}_{\mathbb{R}}(I)\ni f\mapsto \|f\|_{\infty}=\sup_{x\in I}|f(x)|.$$

Then the Weierstrass approximation theorem states that the monomials (restricted to I)

$$\{I\ni x\mapsto x^{\nu}\,|\,\nu\in\mathbb{N}\}$$

form a *total* subset of the real Banach space $\mathscr{C}_{\mathbb{R}}(I)$.

Nowadays numerous proofs of this fundamental theorem are available. In addition to general functional analytic methods (for instance, methods making use of results from the theory of Laplace transform and from complex function theory combined with the extension principle of Helly-Hahn-Banach or ideas which are related to the generalization of the Weierstrass approximation theorem by M.H. Stone) there are various *constructive* methods. To these latter ones we should reckon the most general known method for generating approximations, i.e. the technique of smoothing a given function by convolving it with a suitable kernel ("*regularization methods*"). For a detailed account of this circle of ideas we refer to Shapiro [14]. A concise survey including further references will be found in Todd [15].

Among the constructive approaches to the Weierstrass approximation theorem the most elementary one employs the *Bernstein polynomials*

$$B_n f=\sum_{0\leqq k\leqq n} f\left(\frac{k}{n}\right) b_{nk}, \qquad (n\geqq 1) \tag{1}$$

constructed for any given $f\in\mathbb{R}^I$ by means of the positive functions

$$b_{nk}\colon I\ni x\mapsto \binom{n}{k} x^k(1-x)^{n-k}, \qquad (0\leqq k\leqq n). \tag{2}$$

It is based upon the well-known fact that for any prescribed function $f\in\mathscr{C}_{\mathbb{R}}(I)$ the sequence $(B_n f)_{n\geqq 1}$ of polynomials converges uniformly on I to f. See Bernstein [1] and, for example, the monograph by Lorentz [9].

As pointed out by Bohman [2] and Korovkin [7, 8], a thorough examination of the Bernstein approximation process exhibits the remarkable fact that the *linearity* of the mappings

$$f \mapsto B_n f, \quad (n \geqq 1)$$

of $\mathscr{C}_{\mathbb{R}}(I)$ into itself, together with their *positivity* relative to the natural lattice structure of the real algebra $\mathscr{C}_{\mathbb{R}}(I)$, and the uniform convergence on I of the sequence $(B_n f)_{n \geqq 1}$ towards f *merely* for the following three monomials

$$f: I \ni x \mapsto x^\nu, \quad (\nu = 0, 1, 2)$$

are sufficient to ensure the uniform convergence on I of the sequence $(B_n f)_{n \geqq 1}$ to f for *any* given function $f \in \mathscr{C}_{\mathbb{R}}(I)$. Employing the identical bijection $\mathrm{id}_I: I \ni x \mapsto x$, the Bohman-Korovkin result can be formally stated as follows:

Theorem 1. *Let $(L_n)_{n \geqq 1}$ be a sequence of positive linear mappings of the ordered Banach algebra $\mathscr{C}_{\mathbb{R}}(I)$ into itself. Suppose that*

$$\lim_{n \to \infty} \| L_n(\mathrm{id}_I^\nu) - \mathrm{id}_I^\nu \|_\infty = 0$$

for $\nu = 0, 1, 2$. Then we have

$$\lim_{n \to \infty} \| L_n(f) - f \|_\infty = 0$$

for all functions $f \in \mathscr{C}_{\mathbb{R}}(I)$.

Taking this statement for granted, Bernstein's proof of the Weierstrass approximation theorem requires only a short calculation. See, for instance, Cheney [6] or Meinardus [10]. In addition, Theorem 1 has various applications to approximation problems. For example, it yields a simplified proof of Fejér's result concerning approximation by means of Hermite interpolation with derivatives controlled at the Čebyšev nodes. Moreover, by the positivity of Fejér's kernel, the analogue of Bohman-Korovkin's theorem for the one-dimensional torus group $\mathbb{T}$ implies as an immediate consequence the well-known result on Cesàro summability for the Fourier expansion of functions $f \in \mathscr{C}_{\mathbb{R}}(\mathbb{T})$. See Cheney [6] and particularly Korovkin [8], where further applications along these lines will be found.

An obvious generalization of Bernstein polynomials can be obtained by forming for any two multi-indices $k = (k_j)_{1 \leqq j \leqq m} \in \mathbb{N}^m$, $n = (n_j)_{1 \leqq j \leqq m} \in \mathbb{N}^m$, $1 \leqq n_j$, $k \leqq n$, the tensor product

$$b_{nk} = \bigotimes_{1 \leqq j \leqq m} b_{n_j k_j}$$

of the functions (2) and modelling by complete analogy with (1) the definition of the polynomials $B_n f$ for any real-valued function f on the compact unit hypercube I^m of $\mathbb{R}^m$. Furthermore, for an adaptation of the Bernstein procedure to vector-valued functions we refer to Tucker [16].

However, the definition of Bernstein polynomials for real-valued functions on the compact standard *simplex* of $\mathbb{R}^m$ sketched by Lorentz in Chapter II of his book [9] seems to be more powerful. It has the advantage of being closely related to those elementary considerations of probability (concerning the binomial distribution) which serve usually to motivate the definition (1) of classical Bernstein polynomials. See Todd [15]. Moreover, it is the natural starting-point for an interesting generalization of Bernstein polynomials to spaces of probability measures on compact topological spaces which has been given and extensively discussed in Schnabl [11–13]. The convergence theorem proved by Schnabl [11] represents a translation of the familiar technique into the framework of his general theory and disregards the type of argumentation due to Bohman and Korovkin.

It is the aim of the present paper to establish an extension of Theorem 1 which covers the generalized Bernstein polynomials in the sense of Schnabl. This will be done in Theorem 2 of Section 3 after having set down in Section 2 the concepts we shall deal with. Finally, in Section 4 a convergence theorem will be deduced from Theorem 2 which implies as a special case Schnabl's approximation theorem mentioned above.

Concerning the general notions and terminology needed from the theory of topological vector spaces and functional analytic integration theory, we refer to Bourbaki [3–5].

2. Bernstein-Schnabl Functions

In the sequel let us denote by T a *compact* topological space. As usual in analysis, we assume that the notion of compactness includes the Hausdorff separation axiom. We endow the real vector space $\mathscr{C}_{\mathbb{R}}(T)$ with the *topology of uniform convergence* on T and denote by

$$\mathscr{M}_{\mathbb{R}}(T)=\mathscr{C}'_{\mathbb{R}}(T)$$

the topological dual of the locally convex topological vector space $\mathscr{C}_{\mathbb{R}}(T)$. It will be convenient to denote the canonical bilinear form associated with the topological duality $(\mathscr{C}_{\mathbb{R}}(T), \mathscr{M}_{\mathbb{R}}(T))$ by

$$\mathscr{C}_{\mathbb{R}}(T)\times\mathscr{M}_{\mathbb{R}}(T)\ni(f,\mu)\mapsto\langle f,\mu\rangle=\int_T f\,d\mu.$$

Thus $\mathscr{M}_{\mathbb{R}}(T)$ represents the vector space of the *real Radon measures* on T. We shall assume that $\mathscr{M}_{\mathbb{R}}(T)$ carries the *vague* topology, which means that the vector space $\mathscr{M}_{\mathbb{R}}(T)$ is equipped with the weak topology $\sigma(\mathscr{M}_{\mathbb{R}}(T), \mathscr{C}_{\mathbb{R}}(T))$. The canonical (algebraic) vector space isomorphism of $\mathscr{C}_{\mathbb{R}}(T)$ onto the topological dual $\mathscr{M}'_{\mathbb{R}}(T)$ of its weak dual $\mathscr{M}_{\mathbb{R}}(T)$ will be denoted by

$$f\mapsto\tilde{f}\colon\mu\mapsto\langle f,\mu\rangle.$$

If we introduce the pointed convex cone $\mathscr{M}_+(T)$ in $\mathscr{M}_{\mathbb{R}}(T)$ consisting of the *positive* Radon measures on T, then

$$\mathscr{M}_+^1(T)=\{\mu\in\mathscr{M}_+(T)\,\wr\,\langle 1_T,\mu\rangle=1\}$$

stands for the set of all Radon *probability measures* on T. We shall provide $\mathscr{M}_+^1(T)$ with the relative topology induced by the vague topology $\sigma(\mathscr{M}_{\mathbb{R}}(T), \mathscr{C}_{\mathbb{R}}(T))$. The *compact* topological space obtained in this way will be designated henceforth by X. It is well known that X represents a (vaguely compact) *base* of the convex cone $\mathscr{M}_+(T)$. With respect to the Čebyšev norm $\|\cdot\|_\infty$, $\mathscr{C}_{\mathbb{R}}(X)$ will be considered as a real Banach algebra.

For any natural number $n \geqq 1$ let us denote by T^n the n-fold cartesian product performed with T itself and equipped with the product topology:

$$T^n = \underset{1 \leqq j \leqq n}{\times} T_j, \quad (T_j = T, 1 \leqq j \leqq n).$$

Furthermore, let $P = (p_{nj})_{n \geqq 1, j \geqq 1}$ be a *lower triangular stochastic matrix*, i.e. an infinite real matrix the elements of which have the following properties:

$$\begin{aligned} p_{nj} &\geqq 0, \quad (n \geqq 1, j \geqq 1) \\ p_{nj} &= 0, \quad (j > n) \\ \sum_{j \geqq 1} p_{nj} &= 1, \quad (n \geqq 1). \end{aligned} \tag{3}$$

If we denote for any $t \in T$ by $\varepsilon_t \in X$ the Dirac measure on T placed at the point t, the application

$$\pi_{n,P}\colon T^n \ni (t_j)_{1 \leqq j \leqq n} \mapsto \sum_{j \geqq 1} p_{nj}\, \varepsilon_{t_j} \in X \tag{4}$$

defines a vaguely continuous *diffusion* of T^n into T having the norm $\|\pi_{n,P}\| = 1$. Thus, the mapping $\pi_{n,P}$ assigns to each n-tuple of elements of the compact space T a discrete probability measure on T, i.e. an element of X belonging to the convex hull of the set of extreme points of the base X.

For any Radon measure $\mu \in \mathscr{M}_{\mathbb{R}}(T)$ let us denote by

$$\mu^{\otimes n} = \bigotimes_{1 \leqq j \leqq n} \mu_j, \quad (\mu_j = \mu, 1 \leqq j \leqq n)$$

the n-fold tensor product measure on T^n performed with μ itself. Notice that the mapping

$$\mathscr{M}_+(T) \ni \mu \mapsto \mu^{\otimes n} \in \mathscr{M}_+(T^n)$$

is continuous with respect to the vague topologies. Finally, we shall denote by $\pi_{n,P}(\mu^{\otimes n})$ the *image* of the measure $\mu^{\otimes n}$ under the continuous mapping $\pi_{n,P}\colon T^n \to X$ defined in (4).

Definition. Let the function $F \in \mathscr{C}_{\mathbb{R}}(X)$ be given. For any natural number $n \geqq 1$ the continuous mapping

$$B_{n,P}(F)\colon X \ni \mu \mapsto \langle F, \pi_{n,P}(\mu^{\otimes n}) \rangle = \int_X F \, d\pi_{n,P}(\mu^{\otimes n}) \in \mathbb{R}$$

defined on the compact base X of the convex cone $\mathscr{M}_+(T)$ is called the n-th *Bernstein-Schnabl function* of F with respect to the matrix P of masses.

In the special case that all nonzero elements in the n-th row of P are supposed to be equal, i.e. if

$$p_{nj}=\frac{1}{n}, \quad (n\geqq 1, 1\leqq j\leqq n) \tag{5}$$

we obtain Schnabl's original definition [11].

The mapping $F\mapsto B_{n,P}(F)$ from $\mathscr{C}_{\mathbb{R}}(X)$ into itself is plainly *linear*. If we endow the real Banach algebra $\mathscr{C}_{\mathbb{R}}(X)$ with its natural lattice structure then $B_{n,P}$: $\mathscr{C}_{\mathbb{R}}(X)\to\mathscr{C}_{\mathbb{R}}(X)$ becomes a *positive* linear map. For this reason it is automatically continuous.

In the case that T reduces to the space $\{a, b\}$ consisting of two (distinct) points and T is equipped with the *discrete* topology, then $\mathscr{M}_{\mathbb{R}}(T)$ is the real vector space spanned by the unit point masses $\{\varepsilon_a, \varepsilon_b\}$. Consequently, $\mathscr{M}_{\mathbb{R}}(T)$ is topologically isomorphic to $\mathbb{R}^2$. Therefore the base X of the cone $\mathscr{M}_+(T)$ is homeomorphic to the interval I and the Banach algebras $\mathscr{C}_{\mathbb{R}}(X)$ and $\mathscr{C}_{\mathbb{R}}(I)$ can be identified.

3. The Generalized Theorem of Bohman-Korovkin

We shall adhere to the notations introduced in the previous section. In addition, for each function $f\in\mathscr{C}_{\mathbb{R}}(T)$ let us denote by

$$\hat{f}=\tilde{f}\,|\,X\in\mathscr{C}_{\mathbb{R}}(X)$$

the restriction of the continuous linear form $\tilde{f}\in\mathscr{M}'_{\mathbb{R}}(T)$ to the base X of the convex cone $\mathscr{M}_+(T)$. Then our generalization of Theorem 1 reads as follows:

Theorem 2. *Let $(L_n)_{n\geqq 1}$ be a sequence of positive linear mappings of the ordered Banach algebra $\mathscr{C}_{\mathbb{R}}(X)$ into itself. Suppose that*

$$\lim_{n\to\infty}\|L_n(\hat{f}^\nu)-\hat{f}^\nu\|_\infty=0$$

for each $f\in\mathscr{C}_{\mathbb{R}}(T)$ and $\nu=0,1,2$. Then we have

$$\lim_{n\to\infty}\|L_n(F)-F\|_\infty=0$$

for every function $F\in\mathscr{C}_{\mathbb{R}}(X)$.

Proof. Each continuous mapping F: $X\to\mathbb{R}$ is uniformly continuous with respect to the natural uniform structure on X inducing the relative vague topology and the canonical uniformity on $\mathbb{R}$. The vague topology, i.e. the topology on $\mathscr{M}_{\mathbb{R}}(T)$ of pointwise convergence, arises from the spectrum $\{|\tilde{f}|\,\wr\, f\in\mathscr{C}_{\mathbb{R}}(T)\}$ of seminorms on $\mathscr{M}_{\mathbb{R}}(T)$. Consequently the uniform structure on X will be generated by the spectrum

$$\{X\times X\ni(\mu,\lambda)\mapsto|\langle f,\mu-\lambda\rangle|\,\wr\, f\in\mathscr{C}_{\mathbb{R}}(T)\}$$

of pseudo-metrics.

Let $\eta>0$ be given. It is possible to select a family of k functions $(f_j)_{1\leqq j\leqq k}$ in the space $\mathscr{C}_{\mathbb{R}}(T)$ having the property that the relations $(\mu_1,\mu_2)\in X\times X$ and $\sum_{1\leqq j\leqq k}\langle f_j,\mu_1-\mu_2\rangle^2<1$ imply the inequality

$$|F(\mu_1)-F(\mu_2)|<\eta.$$

From this fact we infer that for all pairs $(\mu,\lambda)\in X\times X$ the following estimate

$$|F(\mu)-F(\lambda)|<\eta+2\cdot\|F\|_\infty\cdot\sum_{1\leqq j\leqq k}\langle f_j,\mu-\lambda\rangle^2$$

is valid. If $1_X=\mathrm{id}_X^0$: $x\mapsto 1$ denotes the unit element of the Banach algebra $\mathscr{C}_{\mathbb{R}}(X)$, then for all measures $\lambda\in X$ the inequality

$$|F-F(\lambda)\cdot 1_X|<\eta\cdot 1_X+2\,\|F\|_\infty\cdot\sum_{1\leqq j\leqq k}(\hat{f}_j^2-2\hat{f}_j(\lambda)\cdot\hat{f}_j+\hat{f}_j^2(\lambda)\cdot 1_X)$$

holds. Taking into account that $|L_n(G)|\leqq L_n(|G|)$ for each function $G\in\mathscr{C}_{\mathbb{R}}(X)$, we obtain for all natural numbers $n\geqq 1$ the inequality

$$|L_n(F)-F\cdot L_n(1_X)|\leqq\eta\cdot L_n(1_X)+2\,\|F\|_\infty\cdot\sum_{1\leqq j\leqq k}(L_n(\hat{f}_j^2)-2\hat{f}_j L_n(\hat{f}_j)+\hat{f}_j^2\,L_n(1_X))$$

so that

$$\|L_n(F)-F\cdot L_n(1_X)\|_\infty\leqq\eta\cdot\|L_n(1_X)\|_\infty+2\,\|F\|_\infty\cdot\sum_{1\leqq j\leqq k}N_{nj}.$$

For brevity we have introduced the following quantities with indices $n\geqq 1$ and $1\leqq j\leqq k$:

$$\begin{aligned}N_{nj}&=\|L_n(\hat{f}_j^2)-2\hat{f}_j L_n(\hat{f}_j)+\hat{f}_j^2\,L_n(1_X)\|_\infty\\&\leqq\|L_n(\hat{f}_j^2)-\hat{f}_j^2\|_\infty+2\,\|\hat{f}_j\|_\infty\cdot\|L_n(\hat{f}_j)-\hat{f}_j\|_\infty+\|\hat{f}_j^2\|_\infty\cdot\|L_n(1_X)-1_X\|_\infty.\end{aligned}$$

From the hypotheses we infer that we have $\lim_{n\to\infty}\sum_{1\leqq j\leqq k}N_{nj}=0$. Hence, by virtue of the fact that η was an arbitrary strictly positive real number, the estimates

$$\begin{aligned}\|L_n(F)-F\|_\infty&\leqq\eta\cdot\|L_n(1_X)\|_\infty+2\,\|F\|_\infty\cdot\sum_{1\leqq j\leqq k}N_{nj}+\|F\|_\infty\cdot\|L_n(1_X)-1_X\|_\infty\\&\leqq\eta+(\eta+\|F\|_\infty)\cdot\|L_n(1_X)-1_X\|_\infty+2\,\|F\|_\infty\cdot\sum_{1\leqq j\leqq k}N_{nj}\end{aligned}$$

combined with the hypotheses establish the statement. –

4. A Convergence Theorem

A straightforward application of the foregoing theorem yields the following convergence theorem valid for Bernstein-Schnabl functions:

Theorem 3. *Let $P=(p_{nj})_{n\geqq 1,\,j\geqq 1}$ be an infinite lower triangular stochastic matrix. Suppose that the elements of P satisfy the condition*

$$\lim_{n\to\infty}\sum_{j\geqq 1}p_{nj}^2=0.$$

Then we have

$$\lim_{n\to\infty} \|B_{n,P}(F)-F\|_\infty = 0$$

for each function $F\in\mathscr{C}_{\mathbb{R}}(X)$.

Proof. According to the properties of the mappings $(B_{n,P})_{n\geqq 1}$ of $\mathscr{C}_{\mathbb{R}}(X)$ into itself, which we have mentioned in Section 2, and in view of the fact that

$$B_{n,P}(\hat{f}^\nu)=\hat{f}^\nu, \qquad (\nu=0,1)$$

for $n\geqq 1$ and all functions $f\in\mathscr{C}_{\mathbb{R}}(T)$, it is sufficient to consider the case $\nu=2$. It can be easily found

$$\hat{f}^2\circ\pi_{n,P}\colon\ T^n\ni(t_j)_{1\leq j\leq n}\mapsto \sum_{j\geqq 1} p_{nj}^2 f^2(t_j)+\sum_{\substack{j\geqq 1,\,k\geqq 1\\ j\neq k}} p_{nj}p_{nk} f(t_j)f(t_k).$$

Thus we obtain the following representation:

$$B_{n,P}(\hat{f}^2)\colon\ X\ni\mu\mapsto\langle\hat{f}^2,\pi_{n,P}(\mu^{\otimes n})\rangle=\sum_{j\geqq 1} p_{nj}^2\langle f^2,\mu\rangle+\sum_{\substack{j\geqq 1,\,k\geqq 1\\ j\neq k}} p_{nj}p_{nk}\langle f,\mu\rangle^2.$$

By the properties of the elements of P listed in (3), we conclude that

$$\|B_{n,P}(\hat{f}^2)-\hat{f}^2\|_\infty=\Big(\sum_{j\geqq 1} p_{nj}^2\Big)\cdot\|\widehat{f^2}-\hat{f}^2\|_\infty$$

holds for all integers $n\geqq 1$. This equality establishes the statement. –

In the special case when the nonzero elements of the matrix P are chosen as indicated in (5), Theorem 3 reduces to the approximation theorem due to Schnabl [11].

References

1. Bernstein, S.N.: Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités. Comm. Soc. Math. Kharkov **13**, 1–2 (1913).
2. Bohman, H.: On approximation of continuous and of analytic functions. Ark. Mat. **2**, 43–56 (1954).
3. Bourbaki, N.: Espaces vectoriels topologiques, Chap. I, II, 2e éd. Paris: Hermann 1966.
4. – Intégration, Chap. I–IV, 2e éd. Paris: Hermann 1965.
5. – Intégration, Chap. V, 2e éd. Paris: Hermann 1967.
6. Cheney, E.W.: Introduction to approximation theory. New York: McGraw-Hill Book Company 1966.
7. Korovkin, P.P.: On convergence of linear positive operators in the space of continuous functions. Doklady Akad. Nauk SSSR (N.S.) **90**, 961–964 (1953).
8. – Linear operators and approximation theory. Delhi: Hindustan Publishing Corp. (India) 1960.
9. Lorentz, G.G.: Bernstein polynomials. Toronto: University of Toronto Press 1953.
10. Meinardus, G.: Approximation von Funktionen und ihre numerische Behandlung. Berlin-Göttingen-Heidelberg-New York: Springer 1964.

11. Schnabl, R.: Eine Verallgemeinerung der Bernsteinpolynome. Math. Ann. **179**, 74–82 (1968).
12. – Zum Saturationsproblem der verallgemeinerten Bernsteinoperatoren. Proceedings of the Conference on "Abstract Spaces and Approximation" held at the Mathematical Research Institute at Oberwolfach, July 18–27, 1968. Ed. by P. L. Butzer and B. Sz.-Nagy. Basel: Birkhäuser 1969.
13. – Zur Approximation durch Bernsteinpolynome auf gewissen Räumen von Wahrscheinlichkeitsmaßen. Math. Ann. **180**, 326–330 (1969).
14. Shapiro, H. S.: Smoothing and approximation of functions. New York: Van Nostrand Reinhold Company 1969.
15. Todd, J.: Introduction to the constructive theory of functions. Basel: Birkhäuser 1963.
16. Tucker, D. H.: A note on Bernstein polynomial type approximations. Proc. Amer. Math. Soc. **18**, 492–494 (1967).

Günter Felbecker
Dozent Dr. Walter Schempp
Institut für Mathematik
der Ruhr-Universität Bochum
BRD-4630 Bochum, Buscheystraße NA
Germany

(Received December 12, 1970)