Approximation by Operators of Probabilistic Type

L. HAHN

Lehrstuhl A für Mathematik, Technological University of Aachen, Aachen, West Germany

Communicated by P. L. Butzer

Received October 2, 1978

This paper is concerned with sets $\mathfrak W$ of sequences $(W)_{n=1}^{\infty} \in \mathfrak W$ of positive linear operators which are of certain probabilistic type and act on certain function classes K. Necessary and sufficient conditions upon $\mathfrak W$ are determined such that each element $(W_n)_{n=1}^{\infty} \in \mathfrak W$ approximates U with a given order of approximation $\psi(n)$ and a given function class K, the limiting operator U being either the identity I or an operator connected with the normal distribution. The saturation problem in this setting is also solved, now in a form giving the order of saturation $\psi(n)$ such that convergence of $(W_n)_{n=1}^{\infty}$ towards U of order $o(\psi(n))$, $n \to \infty$, is impossible unless $W_n = I$, $n \in \mathbb{N}$, and there exists a non-trivial element $(W_n)_{n=1}^{\infty}$ which approximates U with order $\mathcal{O}(\psi(n))$.

1. Introduction

In the classical saturation problem, one considers a fixed approximation process $(U_n)_{n=1}^{\infty}$ and asks for its saturation class

$$S = \{ f; || U_n f - f || = \mathcal{O}(\psi(n)) \}, \tag{1.1}$$

where ψ is to be determined in such a way that $||U_n f - f|| = o(\psi(n))$, $n \to \infty$, implies f to be trivial. Several mathematicians such as Zhuk and Natanson [9] consider what they call a converse problem of saturation theory: they start with a class of functions K and try to find *one* approximation process for which K is the saturation class.

In this paper we also consider a fixed function class K, but we are interested in *all* sequences of operators (of a certain probabilistic type) that approximate each $f \in K$ with a prescribed order of approximation ψ , not necessarily the order of saturation. Moreover, looking at (1.1), we consider convergence of $(U_n)_{n=1}^{\infty}$ to not merely to the identity operator I but also to some other operators, and, in contrast to (1.1), not (only) in the norm but even in the pointwise sense.

More exactly, let K be a fixed function class defined on some interval $J \subseteq \mathbb{R}$, and consider \mathfrak{U} , the set of sequences of positive linear operators $(U_n)_{n=1}^{\infty}$

with $K \subseteq D(U_n)$, the domain of U_n , as well as some positive linear operator, namely the limiting operator U with $K \subseteq D(U)$. Given K, U, ψ , the problem is to determine the subclass $\mathfrak{U}_0 \subseteq \mathfrak{U}$ such that

$$|U_n f(x) - Uf(x)| = \mathcal{O}(\psi(n)) \quad (\text{each } f \in K)$$
 (1.2)

if and only if $(U_n)_{n=1}^{\infty} \in \mathfrak{U}_0$, where $x \in J$ is fixed and ψ describes the order of approximation, i.e., to find

$$\mathfrak{U}_0 = \{ (U_n)_{n=1}^\infty \in \mathfrak{U}; \mid U_n f(x) - U f(x) \mid = \mathcal{O}(\psi(n)), \text{ each } f \in K \}. \quad (1.3)$$

 \mathfrak{U}_0 is the saturation class in this setting, now called satisfy class, if (i) \mathfrak{U}_0 is not the trivial class, i.e., we do not have, for all $f \in K$, $n \in \mathbb{N}$,

$$U_n f(x) = U f(x), (1.4)$$

but, $f \in K$ and

(ii)
$$|U_n f(x) - Uf(x)| = \varrho(\psi(n))$$
 (1.5)

imply (1.4); in this case $\mathfrak U$ is said to be satiated with satiety order $\psi(n)$. Notice that the large " $\mathscr O$ " and small " $\mathscr O$ " terms may depend on the fixed point $x \in J$, the operator U and the function $f \in K$. Of course, one could detach the matter from a fixed $x \in J$ and consider the problem for all $x \in J$, i.e., add " $x \in J$ " in formulas (1.2)–(1.5), the problem nevertheless still being pointwise. Recall that assertions such as (1.2) or (1.5) for, e.g., the Szàsz-Mirakjan or γ operators do not hold uniformly, so that we have to consider them in the pointwise sense.

In the following, as indicated above, we do not consider the whole class $\mathfrak U$ but only a subclass $\mathfrak B \subset \mathfrak U$ such that the elements $(W_n)_{n=1}^{\infty} \in \mathfrak B$ are of the probabilistic form

$$W_n(\varphi(n))f(x) = E(f \circ [\varphi(n)(X_1 + X_2 + \dots + X_n)])$$

$$= E(f \circ \varphi(n) S_n) = \int_{-\infty}^{\infty} f(u) dF_{\varphi(n)S_n}(u) \qquad \text{(each } f \in K),$$

where $(X_i)_{i=1}^{\infty}$ are real independent, identically distributed (i.i.d.) random variables (r.v.) which will be constructed in dependence upon $x \in J$, S_n is the nth partial sum of the r.v. X_i , and φ is a normalizing function $\varphi \colon \mathbb{N} \to \mathbb{R}$, tending to zero for $n \to \infty$; in the sequel we consider two concrete specializations of φ , namely $\varphi(n) = n^{-1}$ and $\varphi(n) = n^{-1/2}$. Examples of operators that belong to \mathfrak{W} are the Bernstein, Szàsz-Mirakjan and Baskakov operators in the discrete case, and the Weierstrass and γ operators (in Feller's version) in the continuous case. The probabilistic structure of the operators guarantee at once that $W_n(\varphi(n))1 = 1$.

The precise formulations of the main results are to be found in Theorems 1, 2 and 3; they will be proved by elementary stochastic methods.

2. NOTATIONS AND PRELIMINARIES

Let $(X_i)_{i=1}^{\infty}$ be a sequence of real i.i.d.r.v., all distributed as a r.v. X, and assume the limiting r.v. Z to be φ -decomposable, i.e., for each $n \in \mathbb{N}$ there exist n independent r.v. $Z_{1,n},...,Z_{n,n}$ such that

$$F_Z = F_{\varphi(n)} \sum_{i=1}^n Z_{i,n},$$

where F_Y denotes the distribution function (d.f.) of a r.v. Y (the decomposability concept was introduced in [3] to prove general limit theorems in probability theory).

Further, let C(J) denote the vector space of all real-valued continuous functions on an interval $J \subseteq \mathbb{R}$, and $C_b(J)$ the subset of all bounded and uniformly continuous functions on J. For $r, j \in \mathbb{P} := \mathbb{N} \cup \{0\}$ set

$$C_b{}^r\!(J):=\{f\!\in\!C(J);f',f'',...,f^{(r-1)}\!\in\!C(J),f^{(r)}\!\in\!C_b(J)\},$$

and $f_i(x) := x^i$. Finally, let $]\alpha[$ be the greatest integer less than or equal to $\alpha > 0$, and $[\alpha]$ be the smallest integer greater than or equal to $\alpha > 0$. Two general inverse theorems on r.v. will be needed (see [6]).

THEOREM. Let the rth moment $E(Z^r)$ be finite for some $r \in \mathbb{N}$.

A. If for some $\beta > 0$

$$\int_{-\infty}^{\infty} f(u) d[F_{\sigma(n)S_n}(u) - F_{Z}(u)] = \mathcal{O}(n\varphi(n)^{\beta})$$

for each $f \in C_b^r(\mathbb{R})$, then

- (i) $E(X^r) < \infty$,
- (ii) $E(X^j) = E([Z/\varphi(1)]^j)$ $(0 \leqslant j \leqslant [\min(r, \beta)] 1)$.

B. If

$$\int_{-\infty}^{\infty} f(u) d[F_{\sigma(n)}S_n(u) - F_Z(u)] = o(n\varphi(n)^r) \qquad (n \to \infty)$$

for each $f \in C_b^r(\mathbb{R})$, then

- (i) $E(X^r) < \infty$,
- (ii) $E(X^j) = E([Z/\varphi(1)]^j)$ $(0 \leqslant j \leqslant r)$.

In view of $C_b^r(\mathbb{R}) \subseteq C_b^r(J)$ one may replace $C_b^r(\mathbb{R})$ by $C_b^r(J)$ for some interval $J \subseteq \mathbb{R}$.

3. THEOREMS CONNECTED WITH THE LIMITING OPERATOR I

In this section we consider the case $\varphi(n) = 1/n$ and U = I, so that the operators U_n are of the form (cf. (1.6))

$$W_n(1/n)f(x) = E(f \circ S_n/n) \qquad (n \in \mathbb{N}). \tag{3.1}$$

THEOREM 1. Let $x \in J$ be fixed.

(a) *If*

$$|W_n(1/n)f(x) - f(x)| = o(1) \quad (n \to \infty)$$
 (3.2)

for each $f \in C_h^{-1}(J)$, then

- (i) $E(X) < \infty$,
- (ii) $W_n(1/n)f_1(x) = f_1(x) = x \ (n \in \mathbb{N}).$
- (b) *If*

$$|W_n(1/n)f(x) - f(x)| = \theta(n^{-\alpha})$$
(3.3)

for each $f \in C_b^2(J)$ and some $\alpha > 0$, then

- (i) $E(X^2) < \infty$,
- (ii) $W_n(1/n)f_1(x) = f_1(x) = x \ (n \in \mathbb{N}).$
- (c) If

$$W_n(1/n)f(x) = f(x) = o(n^{-1}) \qquad (n \to \infty)$$
 (3.4)

for each $f \in C_b^2(J)$, then

- (i) $E(X^2) < \infty$,
- (ii) $W_n(1/n)f(x) = f(x)$ (each $f \in C(J)$; $n \in \mathbb{N}$), i.e., $W_n(1/n)$ interpolates at x.

Proof. To utilize the theorems given above for $\varphi(n) = n^{-1}$, we have to determine the limiting r.v. Z and the associated components. For fixed $x \in \mathbb{R}$ let Z be distributed as X_x , i.e.,

$$F_Z(u) := F_{X_x}(u) := 0, \qquad -\infty < u < x$$

: = 1, $x \le u < \infty.$

In other words, $P({X_x = x}) = 1$. Notice that here the limiting r.v. Z also depends on $x \in J$. It is easy to see that X_x is 1/n-decomposable with components $Z_{i,n}$ all distributed as X_x ; hence

$$F_{X_r} := F_{(1/n)\sum_{i=1}^n Z_{i,n}}$$
.

For the moments of X_x we have $E(X_x^r) = x^r$. In view of (1.6) and

$$\int_{-\infty}^{\infty} f(u) dF_{X_{\varepsilon}}(u) = f(x)$$
(3.5)

(3.2) can therefore be written in the form

$$\int_{-\infty}^{\infty} f(u) \, d[F_{S_n/n}(u) - F_{X_2}(u)] = o(1) \qquad (\text{each } f \in C_b^{-1}(J); \, n \to \infty).$$

To prove part (a), use Theorem B with r = 1 to yield $E(X) < \infty$ and $E(X) = E(X_x) = x$. This leads to

$$W_n(1/n) f_1(x) = E(S_n/n) = \frac{1}{n} \sum_{i=1}^n E(X) = x.$$

Concerning part (b), Theorem A is applicable with r=2, $\beta=1+\alpha$, so that $[\min(r,\beta)]=2$.

Finally, for part (c) use Theorem B again, but now for r = 2, so that $E(X^2) < \infty$ follows immediately, and additionally

$$E(X^j) := E(X_{s^j}) \qquad (0 \leqslant j \leqslant 2).$$

Hence E(X) = x and $E(X^2) = x^2$. Then

$$Var(X) := E([X - E(X)]^2) = E(X^2) - x^2 = 0,$$
 (3.6)

leading to X=x P-a.s. Therefore $F_X=F_{X_x}$, which implies $F_{S_n/n}=F_{X_x}$. Finally,

$$W(1/n) f(x) = \int_{-\infty}^{\infty} f(u) dF_{S_n/n}(u)$$

$$= \int_{-\infty}^{\infty} f(u) dF_{X_n}(u) = f(x),$$
(3.7)

even for each measurable and bounded f since $W_n(1/n)$ is defined for such f. This completes the proof of Theorem 1.

It seems worthwhile to combine the inverse Theorem 1 with direct theorems of [7]. So one has the following results for $\mathfrak{B}(1/n)$, the set of sequences of operators of the form (1.6) with $\varphi(n)=n^{-1}$ and $x\in J$. Below, for example, $\{\mid W_n(1/n)f(x)-f(x)\mid=\varrho(1),\ f\in C_b^{-1}(J)\}$ means the set of all sequences $(W_n(1/n))_{n=1}^\infty\in\mathfrak{B}(1/n)$ such that $\mid W_n(1/n)f(x)-f(x)\mid=\varrho(1),\ n\to\infty$, for each $f\in C_b^{-1}(J)$.

THEOREM 2. Let $W_n(1/n)f(x)$ be defined by (3.1). Then one has

(a)
$$\{|W_n(1/n)f(x) - f(x)| = o(1), f \in C_b^{-1}(J)\} = \{W_n(1/n)f_1(x) = f_1(x), n \in \mathbb{N}, and E(X) < \infty\},$$

(b)
$$\{|W_n(1/n)f(x) - f(x)| = \mathcal{O}(n^{-1}), f \in C_b^2(J)\} = \{W_n(1/n)f_1(x) = f_1(x), n \in \mathbb{N}, and E(X^2) < \infty\},$$

(c)
$$\{|W_n(1/n)f(x) - f(x)| = o(n^{-1}), f \in C_b^2(J)\} = \{W_n(1/n)f(x) = f(x), n \in \mathbb{N}, f \in C_b(J), and E(X^2) < \infty\}.$$

To discuss assertions (a), (b) and (c), if one assumes directly that $E(X^2) < \infty$, i.e., $C_b^2(J)$ belongs to the domain of $W_n(1/n)$, $n \in \mathbb{N}$, then a comparison of (a) and (b) shows that a pointwise approximation process of $\mathfrak{W}(1/n)$ automatically approximates f(x) with $\mathcal{C}(1/n)$ (for smooth f). Moreover, (b) and (c) together show that $\mathfrak{W}(1/n)$ is satiated with satiety order 1/n, i.e., the rate o(1/n) is impossible, unless $W_n(1/n)$ is the identity I.

Notice that the last assertion gives a partial answer to a problem posed by P. L. Butzer in 1963 and recalled by a number of authors (compare, e.g., [1, 5]): Is it possible to construct a sequence of algebraic polynomials, defined on [0, 1], which are of the same structure as the Bernstein polynomials and which approximate the associated function f with an order $\mathcal{O}(n^{-2})$ on [0, 1] provided $f'' \in C[0, 1]$? As mentioned above, the sequence of the Bernstein operators belongs to $\mathfrak{W}(1/n)$, and Theorem 2 states in this respect that as long as one modifies the Bernstein operators in such a fashion that the new operators still belong to $\mathfrak{W}(1/n)$, then a better rate than $\mathcal{O}(n^{-1})$ is impossible. So in this case the answer to the problem just mentioned is negative.

Recalling the well-known Bohman–Korovkin theorem, we know that for the class $\mathfrak U$ the (test) functions f_0 , f_1 and f_2 already guarantee that $(U_n)_{n=1}^{\infty} \in \mathfrak U$ is an approximation process. If one now considers the smaller class $\mathfrak B(1/n) \subset \mathfrak U$, then (a) shows that the two functions f_0 , f_1 already guarantee $(W_n(1/n))_{n=1}^{\infty} \in \mathfrak U(1/n)$ to be an approximation process. But (a) gives even more information. If one has a sequence of positive linear operators U_n that defines an approximation process, and $U_n f_1(x) = f_1(x)$ does not hold (but of course approximatively for $n \to \infty$), then U_n cannot have the probabilistic form (3.1), i.e., there cannot exist r.v. X_i such that $U_n f(x) = E(f \circ [S_n/n])$.

4. Results for the Limiting U Connected with the Normal Distribution

For a second application of Theorems A and B we choose $\varphi(n) = 1/\sqrt{n}$ and $F_Z = F_{X^*}$, where X^* is a normally distributed r.v. with mean 0 and variance 1, i.e.,

$$F_{X^*}(u) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt.$$

Concerning (1.6) we now consider operators $(W_n(1/\sqrt{n}))_{n=1}^{\infty} \in \mathfrak{W}(1/\sqrt{n})$ of the form

$$W_n(1/\sqrt{n})f(x) = E(f \circ [S_n/\sqrt{n}]) \qquad (f \in K), \tag{4.1}$$

the limiting operator now being

$$Uf = U_1 f := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-t^2/2} dt, \tag{4.2}$$

which is independent of $x \in \mathbb{R}$. Then we have the following theorem (the proof follows from [6])

THEOREM 3. Let $x \in R$ be fixed.

(a) If

$$|W_n(1/\sqrt{n})f(x) - U_1f| = \mathcal{C}(n^{-\alpha}) \qquad (each f \in C_b^{12\alpha + 2[(\mathbb{R})]})$$
 (4.3)

for some $\alpha > 0$, then

(i)
$$E(X^{12\alpha+2\delta}) < \infty$$
,
(ii) $E(X^j) = E(X^{*j})$ $(0 \le j \le [2\alpha+2]-1)$.

(b) *If*

$$|W_n(1/\sqrt{n})f(x) - U_1f| = o(n^{-r/2}) \quad (each f \in C_b^{r+2}(\mathbb{R}); n \to \infty)$$
(4.5)

for some $r \in \mathbb{P}$, then

(i)
$$E(X^{r+2}) < \infty$$
,
(ii) $E(X^j) = E(X^{*j})$ $(0 \le j \le r+2)$.

With corresponding direct results of [4], conditions (4.4) and (4.6) also imply (4.3) and (4.5), respectively, the counterpart of Theorem 2 could therefore be formulated, characterizing the operator classes $\mathfrak{W}_{\alpha}(1/\sqrt{n}) \subset \mathfrak{W}(1/\sqrt{n})$.

In contrast to the previous case (with the limiting operator 1), $\mathfrak{W}(1/\sqrt{n})$ is not polynomially satiated since each fixed arbitrary high order of approximation $n^{-\alpha}$, $\alpha > 0$, of $(W_n(1/\sqrt{n}))_{n=1}^{\infty} \in \mathfrak{W}(1/\sqrt{n})$ towards U_1 is possible (cf. (4.4)) without implying $F_X = F_{X^*}$, or equivalently $W_n(1/\sqrt{n}) = U_1$, $n \in \mathbb{N}$. But if one would require exponential approximation order e^{-an} , some a > 0, implying

$$|W_n(1/\sqrt{n})f(x)-U_1f|=\mathcal{C}(n^{-\alpha})$$

for each $\alpha > 0$, then (4.4) would yield that all moments of the r.v. X coincide with those of X^* , thus $F_X = F_{X^*}$, since the moments of X^* are known to generate a well-defined Hamburger moment problem.

On the other hand, given an arbitrary sequence of r.v. X_i that determine a corresponding operator, one may transform this sequence by setting $X_i' := [X_i - E(X_i)]/(\text{Var}(X_i))^{1/2}$ if $0 < \text{Var}(X_i) < \infty$, so that $E(X^j) = E(X^{*j})$, j = 0, 1, 2. [Of course, in general $E(X^3) \neq E(X^{*3})$ (= 0)]. This means that for these operators the best possible rate of convergence is given by

$$|W_n(1/\sqrt{n})f(x) - U_1f| = \ell(n^{-1/2}) \quad (f \in C_b^3(\mathbb{R})).$$

This is the situation for binomially and exponentially distributed r.v. X leading to Bernstein and Szàsz-Mirakjan operators. An example of r.v. which yield higher order of approximation is given in [2]. For each $x \in (0, \frac{1}{2}]$, the corresponding r.v. X(x) are defined via their distributions by $P(\{X(x) - (2x)^{-1/2}\}) = P(\{X(x) = (2x)^{-1/2}\}) = x$ and $P(\{X(x) = 0\}) = 1 - 2x$. Some calculations yield that the corresponding operator $B_n(1/\sqrt{n})f(x) = E(f \in [S_n/\sqrt{n}])$, $x \in (0, \frac{1}{2}]$, is of the form

$$B_n(1/\sqrt{n}) f(x) = \sum_{k=0}^n \sum_{k=1}^n {n \choose k} {k \choose j} x^k (1 - 2x)^{n-k} f(\frac{k-2j}{\sqrt{2nx}}).$$

For this operator, which is a polynomial of degree n, we have as best possible rate for $x \in (0, \frac{1}{2}] \setminus \{\frac{1}{6}\}$

$$|B_n(1/\sqrt{n})f(x) - U_1f| = \ell(n^{-1})$$
 (each $f \in C_b^4(\mathbb{R})$)

and for $x = \frac{1}{6}$

$$|B_n(1/\sqrt{n})f(\frac{1}{6})-U_1f|+\mathcal{C}(n^{-2})$$
 (each $f\in C_b^6(\mathbb{R})$).

Let us conclude this paper with the following remark: An important result of A. Y. Khintchine (cf., e.g., [8]) states that if $\varphi(n)$ S_n converges in distribu-

tion to a r.v. Z, then Z must be stably distributed; convergence in distribution means that

$$\lim_{n\to\infty} E(f\circ [\varphi(n)\,S_n]) = \lim_{n\to\infty} W_n(\varphi(n))\,f(x)$$

$$= E(f\circ Z) \qquad \text{(each } f\in C_h^r(\mathbb{R}))$$

for any $r \in \mathbb{P}$. Thus all possible limiting operators U of sequences of operators of probabilistic type (1.6) are determined by the stably distributed r.v. Z, namely $Uf = E(f \circ Z)$. If one now assumes additionally that $f_2 \in D(U)$ —which is natural when considering rates of convergence—or, equivalently that the second moment of Z is finite, then the only remaining stably distributed r.v. Z are X_x and X^* . Furthermore, the φ -decomposability of these limiting r.v. is needed, which is exactly true for $\varphi(n) = n^{-1}$ and $\varphi(n) = n^{-1/2}$, respectively. This is the reason why the probabilistic approach of this paper only allows one to consider the operators $W_n(1/n)$, $W_n(1/\sqrt{n})$ with the corresponding limiting operators I and U_1 .

ACKNOWLEDGMENTS

The author would like to thank Professors Butzer and Nessel and Dr. Splettstößer for their critical reading of the manuscript.

This work was supported by research Grant II B2-FA 7233 of "Der Minister für Wissenschaft und Forschung des Landes Nordrhein-Westfalen."

REFERENCES

- 1. P. L. BUTZER, On a problem of algebraic approximation, posed in 1963, *in* "Linear Operators and Approximation II" (P. L. Butzer and B. Sz.-Ngay, Eds.), ISNM Vol. 25, pp. 580–581, Birkäuser, Basel, 1974.
- 2. P. L. BUTZER, L. HAHN, AND U. WESTPHAL, On the rate of approximation in the central limit theorem, *J. Approximation Theory* **13** (1975), 327–340.
- 3. P. L. BUTZER AND L. HAHN, General theorems on rates of convergence in distribution of random variables. I. General limit theorems, J. Multivar. Anal. 8 (1978), 181–201.
- 4. P. L. BUTZER AND L. HAHN, General theorems on rates of convergence in distribution of random variables. II. Applications to the stable limit laws and weak law of larg numbers, *J. Multivar. Anal.* 8 (1978), 202–221.
- P. L. BUTZER, R. L. STENS, AND M. WEHRENS, Approximation by algebraic convolution integrals, Proceedings, Conference on Approximation Theory and Functional Analysis, Campinas, Brasil 1978 (J. B. Prolla, Ed.), North-Holland, Amsterdam, 1978.
- L. Hahn, Inverse theorems on the rate of approximation for certain limit theorems in probability theory, Proceedings, Conference on Linear Spaces and Approximation, Oberwolfach 20–27 August 1977 (P. L. BUTZER and B. Sz.-Nagy, Eds.), ISNM Vol. 40, pp. 583–601, Birkhäuser, Basel, 1978.

- 7. L. Hahn, Stochastic methods in connection with approximation theorems for positive linear operators, 1979.
- 8. B. V. GNEDENKO AND A. N. KOLMOGOROV, "Limit Distributions for Sums of Independent Random Variables," Addison-Wesley, Reading, Mass. 1954.
- 9. V. V. Zhuk and G. I. Natanson, Saturation theory converse problem [in Russian], *Mat. Zametki* 6 (1967), 583–590; English translation in *Math. Notes* 6 (1969), 811–815.