Journal of Approximation Theory **92**, 22–37 (1998)
Article No. AT963113

On the Monotonicity of Positive Linear Operators

M. Kazim Khan

Department of Mathematics and Computer Science, Kent State University, Kent, Ohio 44242

B. Della Vecchia

Instituto per Applicazioni della Matematica del C.N.R. di Napoli, via P. Castellino 111, 80131 Naples, Italy

and

A. Fassih

Department of Mathematics and Computer Science, Kent State University, Kent, Ohio 44242

Communicated by Ranko Bojanic

Received December 19, 1995; accepted December 3, 1996

We provide sufficient conditions for a sequence of positive linear approximation operators, $L_n(f, x)$, converging to $f(x)$ from above to imply the convexity of f. We show that, for the convolution operators of Feller type, $K_n(f, x)$, generated by a sequence of iid random variables taking values in an interval I, having a finite moment generating function, the inequalities $K_n(f, x) \geqslant f(x)$ $(x \in I, n \geqslant 1)$ are necessary and sufficient conditions for f to be convex. This provides a converse of a well-known result of R. A. Khan (*Acta. Math. Acad. Sci. Hungar.* **39** (1980), 193–203). It contains, as a special case, the corresponding result for the Bernstein polynomials and extends two results obtained for bounded continuous functions by Horova for Szász and Baskakov operators. As examples, similar results are also provided for the beta, Meyer-König Zeller, Picard, and Bleiman, Butzer, and Hahn operators.

1. INTRODUCTION

Let I be an interval and, for each $x \in I$, let $\{\mu_{n,x}, n = 1, 2, ...\}$ be a sequence of finite measures concentrating on I. Let $g(t)$ be a non-negative continuous function increasing to infinity as $t \to \pm\infty$. We define $D_g(I)$ to

0021-9045/98 $25.00

be the linear space of continuous functions over I so that $f \in D_g(I)$ if and only if for each $\alpha > 0$

$$\sup_{t \in I} \frac{|f(t)|}{e^{\alpha g(t)}} < \infty.$$

When I is unbounded we assume that g grows faster than $\log(|t|+1)$ to avoid trivialities. For all our applications $g(t)$ will grow at least as $|t|$. When $g(t) = |t|$, we will denote $D_g(I)$ by $D(I)$. Throughout we will assume that for some given fixed g, each $f \in D_g(I)$ is integrable with respect to each $\mu_{n,x}$, $n = 1, 2, ..., x \in I$ and denote the integral by

$$L_n(f, x) := \int_I f(t)\, \mu_{n,x}(dt), \qquad n = 1, 2, ...; \quad x \in I.$$

In the next section we provide conditions under which $L_n(f, x) \geqslant f(x)$ for all $n \geqslant 1$ and each $x \in I$ implies that f is convex.

An important special case of the above sequence of positive linear operators which reproduce constant functions can be stated using probability distributions. Let M_n be a sequence of positive linear approximation operators of the type

$$M_n(f, x) := Ef(T_{n,x}) = \int_I f(t)\, dG_{n,x}(t); \qquad n = 1, 2, ...,$$

where the random variables $T_{n,x}$, taking values in the interval I, have respective distributions $G_{n,x}$, provided $E\,|f(T_{n,x})| < \infty$ for each $x \in I$ and $n = 1, 2, ...$ and $f \in D_g(I)$. This class contains, besides the following large group of convolution operators, the beta, Bleiman, Butzer, and Hahn, Meyer-König Zeller, Picard, and many other classical operators.

A further special case of M_n is obtained when $T_{n,x} = (X_{1,x} + X_{2,x} + \cdots + X_{n,x})/n$ where $X_{i,x}$ are iid random variables taking values in I. The resulting sequence of operators will be called the convolution operators (of Feller type) due to the fact that the distribution $G_{n,x}$ is a scaled version of the convolution of the individual distributions of $X_{i,x}$, $i = 1, 2, ..., n$. We will denote the resulting sequence of operators by K_n.

For a number of positive linear approximation operators of the type M_n, it is known that for convex functions the convergence is monotone from above. For this result, the reader may see [1, 3, 4, 12, 14, 16, 18] and a list of further references in these papers. In particular, [14] gave a probabilistic method of showing that, under the condition of existence of first moment, K_n have to be monotone from above for convex functions. A different proof of this result and some modifications were recently provided by [13].

In this paper we will provide some converse results along the lines of [18]. However, our main difference is that we do not assume that L_n is defined over a compact interval nor do we necessarily assume that L_n reproduce constant and linear functions. We will provide a number of examples of operators of the type M_n as special cases. Somewhat as a converse of the result of R. A. Khan for K_n, we prove that the existence of the moment generating function (mgf) of $X_{1,x}$, having mean x, positive variance, and $K_n(f, x) \geqslant f(x)$ $(x \in I, n \geqslant 1)$ implies that f must be convex over I. As further special cases, this result contains two well-known results for the Bernstein operator [15, 17] and extensions of the results in [9, 10] for Baskakov and Szász operators and a few results in [18].

2. MONOTONICITY OF FUNCTIONALS

For the operator $L_n(f, x)$ as defined in the Introduction we will use the notation

$$\tau_{n,x}(\delta) = \int_{\{t:\,|t-x| \geqslant \delta\}} \mu_{n,x}(dt); \qquad n = 1, 2, \ldots; \quad \delta > 0.$$

THEOREM 1. *Let $L_n(f, x)$ be the sequence of positive linear operators as defined in the Introduction so that $L_n(f, x) \to f(x)$ for each $x \in I$ and each $f \in D_g(I)$. Let L_n reproduce the functions $f(t) = 1$ and $f(t) = t$. Let for each x in the interior of I and any $\delta > 0$ there exist $0 < \delta_1(x) < \delta_2(x) < \delta$ and a constant $v = v(\delta_1, \delta_2, x) \in (0, 1)$ and positive integers $n_i = n_i(\delta_1, \delta_2, x)$, $n_i < n_{i+1}$ such that*

$$\tau_{n_i,x}(\delta_1) > \tau_{n_i,x}(\delta_2); \qquad i = 1, 2, \ldots, \tag{1.1}$$

$$(\tau_{n,x}(\delta_2))^v = O(\tau_{n,x}(\delta_1)); \qquad as \quad n = n_i \to \infty. \tag{1.2}$$

Then for any $f \in D_g(I)$, the inequalities $L_n(f, x) \geqslant f(x)$; $x \in I, n = 1, 2, \ldots$ hold if and only if f is convex over I. When I is compact, we may take $v = 1$ and replace (1.2) *with*

$$\tau_{n,x}(\delta_2) = o(\tau_{n,x}(\delta_1)) \qquad as \quad n = n_i \to \infty. \tag{1.3}$$

Proof. If f is not convex then (cf. [15]) there exist an element x in the interior of I, an interval $(x-\delta, x+\delta) \subseteq I$, and a line $\ell(t)$ such that

$$f(x) = \ell(x),\ f(t) < \ell(t) \qquad \text{for all} \quad t \in (x-\delta, x+\delta) \setminus \{x\}.$$

Extend ℓ over I. Now select $0<\delta_1<\delta_2<\delta$ so that the conditions of the theorem hold and define

$$A:=\{t\in I: |t-x|\geqslant\delta_2\},\ B:=\{t\in I: \delta_1\leqslant|t-x|<\delta_2\},$$
$$C:=\{t\in I: |t-x|<\delta_1\}.$$

Let χ_A be the characteristic function of the set A and let $L_n(f-\ell, x)=U_n+V_n+W_n$, where

$$U_n:=L_n((f-\ell)\chi_A, x),\quad V_n:=L_n((f-\ell)\,\chi_B, x),\quad W_n:=L_n((f-\ell)\,\chi_C, x).$$

Clearly, both V_n and W_n are non-positive terms. Over the closure of the set B, $\ell(t)-f(t)$ attains its minimum value, say m, where $m>0$ and gives

$$|V_n|=L_n((\ell-f)\,\chi_B, x)\geqslant mL_n(\chi_B, x).$$

Using $q=1/v(\delta_1, \delta_2, x)$ and $1/p+1/q=1$, the Hölder inequality provides

$$|U_n|\leqslant\left(\int_I |f(t)-\ell(t)|^p\,\mu_{n,x}(dt)\right)^{1/p}(L_n(\chi_A, x))^{1/q}$$
$$=o(1)(L_n(\chi_A, x))^{v(\delta_1, \delta_2, x)}.$$

Now condition (1.1) gives $\{n_i\}$, $n_i=n_i(\delta_1, \delta_2, x)$, so that $L_{n_i}(\chi_B, x)>0$ for all i. For $n=n_i$, we have

$$a_n:=\frac{L_n(\chi_{\{|t-x|\geqslant\delta_2\}}, x)}{L_n(\chi_{\{|t-x|\geqslant\delta_1\}}, x)}\leqslant\frac{(\tau_{n,x}(\delta_2))^{1-v}\,(\tau_{n,x}(\delta_2))^{v}}{\tau_{n,x}(\delta_1)}\to 0.$$

Therefore, for $n=n_i$, we have

$$\frac{|U_n|}{|V_n|}\leqslant\frac{o(1)}{m}\,\frac{\{L_n(\chi_A, x)\}^{v}}{L_n(\chi_B, x)}=\frac{o(1)}{m(1-a_n)}\,\frac{\{\tau_{n,x}(\delta_2)\}^{v}}{\tau_{n,x}(\delta_1)}\to 0.$$

Hence, for some sufficiently large value of $n=n_i$,

$$0\leqslant L_n(f, x)-f(x)=L_n(f-\ell, x)<0.$$

(Note that we have used the reproduction of linear functions only at the above step.) This contradiction gives the first part of the theorem. When I is compact, the same argument of the proof goes through, except that now $D_g(I)=C(I)$ and let M be a bound of f. Then over $n=n_i$,

$$\frac{|U_n|}{|V_n|}\leqslant\frac{M\,\{L_n(\chi_A, x)\}}{m\,L_n(\chi_B, x)}=\frac{M}{m(1-a_n)}\,\frac{\tau_{n,x}(\delta_2)}{\tau_{n,x}(\delta_1)}\to 0.$$

Conversely, if f is convex then Jensen's inequality gives that $L_n(f, x) \geqslant f(x)$ for all $n \geqslant 1$ and each $x \in I$. This finishes the proof. ∎

The following extension is needed when proving the converse results for those operators which do not reproduce the function $f(t) = t$.

THEOREM 1(a). *Let $L_n(f, x)$ be the sequence of positive linear operators as defined in the Introduction. Let $b_n(x) = L_n(t, x) - x$, which need not be zero. For each x in the interior of I and any $\delta > 0$ let there exist $0 < c(x) < \delta_1(x) < \delta_2(x) < \delta$ and a constant $v = v(\delta_1, \delta_2, x) \in (0, 1)$ and positive integers $n_i = n_i(\delta_1, \delta_2, x)$, $n_i < n_{i+1}$ such that conditions* (1.2) *and* (1.1) *hold and*

$$\tau_{n_i, x}(c) > \tau_{n_i, x}(\delta_1), \qquad b_n(x) = o(\tau_{n_i, x}(c) - \tau_{n_i, x}(\delta_1)) \qquad \text{as} \quad n = n_i \to \infty.$$

Then for any $f \in D(I)$, $L_n(f, x) \geqslant f(x)$, $x \in I$, $n = 1, 2, \ldots$ imply that f is convex over I. When I is compact, we may again take $v = 1$ and replace (1.2) *by* (1.3).

Proof. The argument of the proof of Theorem 1 gives that for all large $n = n_i$,

$$U_n + V_n + W_n = L_n(f, x) - f(x) + b_n(x) < 0.$$

We need to show that $W_n - b_n(x) < 0$ for large $n = n_i$. The only extra step that needs to be added in the proof of Theorem 1 is that

$$\frac{|b_n(x)|}{|W_n|} \leqslant \frac{|b_n(x)|}{s\{\tau_{n, x}(x) - \tau_{n, x}(\delta_1)\}} \to 0; \qquad \text{for} \quad n = n_i \to \infty,$$

where $0 < s$ is the minimum of $\ell(t) - f(t)$ over the set $\{t : c(x) \leqslant |t - x| \leqslant \delta_1(x)\}$. ∎

Remark. It is now clear how we can further split the W_n term in the proof of Theorem 1 to accommodate those operators which do not reproduce even the constant functions. We omit the details, however. The above results are modifications of those of Ziegler [18]. Although his results were applicable to generalized convexity, the two main restrictions were that the interval I was assumed to be compact and that L_n were assumed to reproduce the first two polynomials of an extended Chebyshev system. To see a negative result, the reader is referred to [11].

3. CONVOLUTION OPERATORS (FELLER TYPE)

In 1980, R. A. Khan showed that for the convolution operator $K_n(f, x)$, the convexity of f implies that $K_n(f, x) \geqslant f(x)$. We now provide a result in the opposite direction. For this we will need the following lemma.

Lemma 1. *Let X be a non-degenerate random variable taking values in an interval I, having finite mgf $\phi(\theta)$ with mean x. If $g_\delta(\theta) = (x+\delta)\theta - \log\phi(\theta)$ then there exists a $\delta > 0$ so that for any choice of constants $0 < \delta_1 < \delta_2 < \delta$, we have $g_{\delta_1}(\gamma(\delta_1)) < g_{\delta_2}(\gamma(\delta_2))$, where $\theta^* = \gamma(\delta)$ is a unique solution of the equation $\phi'(\theta) = \phi(\theta)(x+\delta)$. Furthermore, θ^* is the point of maxima of $g_\delta(\theta)$.*

Proof. Since X is non-degenerate, it is straightforward to show that $\phi'(\theta)/\phi(\theta)$ is a strictly increasing function of θ in a small enough interval $[0, u]$. Therefore, we can find a $\delta > 0$, sufficiently small so that there exists a unique $\theta = \gamma(\delta)$ so that $\phi'(\theta) = \phi(\theta)(x+\delta)$, and $0 < \gamma(\delta) < u < \varepsilon$. Furthermore, $\gamma(t)$ is a strictly increasing function of $t \in (0, \phi'(u)/\phi(u))$. The function

$$g_\delta(\theta) := (x+\delta)\theta - \log\phi(\theta); \qquad 0 < \theta < u$$

is strictly concave over $0 \leqslant \theta \leqslant u$ and the value $\gamma(\delta)$ is the unique point of maxima of $g_\delta(\theta)$ over $[0, u]$. For any choice of $0 < \delta_1 < \delta_2 < \delta$,

$$\begin{aligned} g_{\delta_1}(\gamma(\delta_1)) &< (x+\delta_2)\,\gamma(\delta_1) - \log\phi(\gamma(\delta_1)) = g_{\delta_2}(\gamma(\delta_1)) \\ &\leqslant (x+\delta_2)\,\gamma(\delta_2) - \log\phi(\gamma(\delta_2)) = g_{\delta_2} = g_{\delta_2}(\gamma(\delta_2)). \end{aligned}$$

That is, the constants $g_{\delta_i}(\delta_i))$ are increasing with $i = 1, 2$. ∎

Theorem 2. *Let $K_n(f, x)$ be the convolution operators (of Feller type) so that, $K_n(f, x) \to f(x)$ for each $f \in D_g(I)$ and each $x \in I$. For each x in the interior of I, let $X_{1,x}$ have finite mgf, $E(X_{1,x}) = x$ and* $\mathrm{Var}(X_{1,x}) > 0$. *For any $f \in D_g(I)$, the following statements are equivalent:*

(1) *f is convex*

(2) $K_n(f, x) \geqslant K_{n+1}(f, x)$ $(x \in I, n \geqslant 1)$,

(3) $K_n(f, x) \geqslant f(x)$ $(x \in I, n \geqslant 1)$.

Proof. The result of R. A. Khan [14] implies that (1) implies (2). By the given hypothesis, (2) implies (3). To show that (3) implies (1), we will show that K_n satisfy the conditions of Theorem 1. Let x be a point in the interior of I. By two applications of Lemma 1 (one for $X_{1,x}$ and the other for $Y = -X_{1,x}$), we may select an $\varepsilon > 0$ sufficiently small so that for any choice of $0 < \delta_1 < \delta_2 < \varepsilon$ we have $g_{\delta_1}(\gamma(\delta_1)) < g_{\delta_2}(\gamma(\delta_2))$ and for $h_\delta(\theta) = (y+\delta)\theta - \log\psi(\theta)$, we have $h_{\delta_1}(\mu(\delta_1)) < h_{\delta_2}(\mu(\delta_2))$, where $\mu(\delta)$ is the solution of the equation

$$\frac{\psi'(\theta)}{\psi(\theta)} = \frac{-\phi'(-\theta)}{\phi(-\theta)} = -x + \delta.$$

By the rate of convergence in the law of large numbers (see, for instance, [5, 6]) for any $0<\delta_1<\varepsilon$,

$$P(S_n \geqslant n(x+\delta_1)) \approx \exp\{-ng_{\delta_1}(\gamma(\delta_1))\}.$$

Hence, for any $0<\delta_1<\delta_2<\varepsilon$ and all large values of n, $\tau_{n,x}(\delta_1)>\tau_{n,x}(\delta_2)$. Take

$$1<q<\min\left\{\frac{g_{\delta_2}(\gamma(\delta_2))}{g_{\delta_1}(\gamma(\delta_1))}, \frac{h_{\delta_2}(\mu(\delta_2))}{h_{\delta_1}(\mu(\delta_1))}\right\},$$

and let $v=v(\delta_1, \delta_2, x)=1/q$. By using the rate of convergence in the law of large numbers once again, for all large values of n, we have

$$\frac{\{\tau_{n,x}(\delta_2)\}^{v}}{\tau_{n,x}(\delta_1)} \leqslant \frac{\{P(S_n \geqslant n(x+\delta_2))\}^{1/q}}{P(S_n \geqslant n(x+\delta_1))} + \frac{\{P(-S_n \geqslant n(-x+\delta_2))\}^{1/q}}{P(-S_n \geqslant n(-x+\delta_1))} \to 0.$$

This finishes the proof. ∎

To avoid the assumption $K_n(f, x) \to f(x)$ in Theorem 2, we may restrict $g(t)=|t|$. In this case for any $f \in D(I)$, the Lebesgue dominated convergence theorem gives that $K_n(f, x) \to f(x)$ for each $x \in I$. This gives the following corollary.

Corollary 1. *Let $K_n(f, x)$ be the convolution operators (of Feller type) so that, for each x in the interior of I, $X_{1,x}$ has finite mgf, $E(X_{1,x})=x$ and* $\mathrm{Var}(X_{1,x})>0$. *Over the space $D(I)$, the following statements are equivalent*:

(1) *f is convex,*

(2) $K_n(f, x) \geqslant K_{n+1}(f, x)$ $(x \in I, n \geqslant 1)$,

(3) $K_n(f, x) \geqslant f(x)$ $(x \in I, n \geqslant 1)$.

Example 1. If we take $P(X_{1,x}=1)=x=1-P(X_{1,x}=0)$, for $x \in [0, 1]$, in Corollary 1, then $K_n(f, x)$ reduces to the usual Bernstein polynomials and gives the results in [15, 17, 18].

Example 2. Let $P(X_{1,x}=j)=\exp(-x)\, x^j/j!$; $j=0, 1, 2, \ldots$ and $x>0$. Now $K_n(f, x)$ reduces to the Szász operator. Horova [10] provided the converse result for the Szász operators under the assumption that f is bounded. Theorem 2 shows that the converse result of convexity holds for a more general class of functions $D(I)$. In fact, it is known (cf. [8]) that the Szász operators can approximate functions lying in the larger space $D_g(I)$, where $g(t)=t\ln(t+1)$ for $t \in [0, \infty)$. Hence Theorem 2 is applicable over this space.

EXAMPLE 3. Let $P(X_{1,x}=j)=(1+x)^{-1}x^j(1+x)^{-j}; j=0, 1, 2, ...$ and $x>0$. For this case the convolution operator reduces to the Baskakov operator. Corollary 1 shows that the converse convexity result holds for any $f \in D(I)$. This extends a result of Horova [9] over a larger class of functions.

EXAMPLE 4. For $x>0$, the Gamma operator is defined as

$$G_n(f, x)=\frac{x^{-n}}{(n-1)!}\int_0^{\infty} f\left(\frac{y}{n}\right) y^{n-1}e^{-y/x}\,dy; \qquad n=1, 2, ...$$

This operator is generated by an iid sequence of exponential random variables with parameter $1/x$. Corollary 1 applies in this case.

EXAMPLE 5. Let $X_{1,x}$ have probability density function defined over $I=(-\infty, \infty)$ by

$$f_{X_{1,x}}(y)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y-x)^2\right),$$

where $x \in I$. By a change of variable we see that the Weierstrass operator

$$W_n(f, x)=\sqrt{\frac{n}{2\pi}}\int_{-\infty}^{\infty} f\left(x+\frac{u}{n}\right)\exp(-(nu^2)/2)\,du$$

is also a special case of K_n. Note that W_n approximates functions lying in $D_g(\mathbb{R})$ where $g(t)=t^2$ and Theorem 2 is applicable. This result was directly proved in [18].

EXAMPLE 6. If we take $Y_i \overset{\text{iid}}{\sim} G$ taking values in $[0, b], b>1$ with $E(Y_1)=1$ and $\text{Var}(Y_1)>0$, then for any $f \in [0, 1]$, Ziegler [18] considered the operator

$$U_n(f, x)=Ef\left(\frac{x\sum_{i=1}^n Y_i}{n}\right)=Ef\left(\frac{\sum_{i=1}^n X_{i,x}}{n}\right),$$

where $X_{i,x}=xY_i$, $x \in [0, 1]$ and $E(X_{1,x})=x$ and $\text{Var}(X_{i,x})=x^2$ $\text{Var}(Y_1)>0$ for $x>0$. Corollary 1 now gives a result of Ziegler. Similarly, some other results of Ziegler can be directly deduced from Corollary 1 by this approach.

Theorem 2 does not apply directly to non-Feller operators. In the following we present a few applications of Theorems 1 and 1(a).

4. MONOTONICITY OF BETA OPERATORS

When $I=[0,1]$, the beta operator is defined by

$$B_n(f,x)=\int_0^1 f(t)\,\frac{t^{nx-1}(1-t)^{n(1-x)-1}}{B(nx,n(1-x))}\,dt; \qquad n=1,2,...; \quad x\in(0,1),$$

where $B(a,b)$ is the beta function. When $x=0$ or 1, then $B_n(f,x):=f(x)$ for all n.

THEOREM 3. *Over the space $C[0,1]$ the following statements are equivalent*:

(1) *f is convex,*

(2) *$B_n(f,x)\geqslant B_{n+1}(f,x)$ for each $n\geqslant 1$ and each $x\in[0,1]$,*

(3) *$B_n(f,x)\geqslant f(x)$ for each $n\geqslant 1$ and each $x\in[0,1]$.*

Proof. It was conjectured in [12] that (1) implies (2). This conjecture was recently proved by Adell *et al.* [1]. And that (2) implies (3) is trivial. To prove that (3) implies (1), we need only check the conditions of Theorem 1. Let $x\in(0,1)$ and let $\delta>0$ be any small enough number so that $0<x-\delta<x+\delta<1$. Trivially,

$$\mu_{n,x}(\{t\colon \delta_1\leqslant|t-x|<\delta_2\})>0, \qquad \text{for all} \quad n=1,2,... .$$

Now note that

$$B_n(f,x)=Ef(T_n), \qquad \text{where} \quad T_n=\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i+\sum_{i=1}^n Y_i},$$

where $X_i \overset{\text{iid}}{\sim} G(1,x)$ and $Y_i \overset{\text{iid}}{\sim} G(1,1-x)$ and the two sequences are mutually independent. (Here $G(\lambda,\alpha)$ stands for the gamma distribution.) The beta operator reproduces the linear functions. Now note that for any $0<\varepsilon<\min\{x,1-x\}$,

$$P(T_n>x+\varepsilon)=P\left(\sum_{i=1}^n Z_i>0\right),$$

where $Z_i=X_i(1-x-\varepsilon)+Y_i(-x-\varepsilon)$.

Indeed, $E(Z_1)=-\varepsilon$ and its mgf equals

$$\phi(\theta)=\left(\frac{1}{1-(1-x-\varepsilon)\theta}\right)^x\left(\frac{1}{1+(x+\varepsilon)\theta}\right)^{1-x}; \qquad \frac{-1}{x+\varepsilon}<\theta<\frac{1}{1-x-\varepsilon}.$$

By the standard results on large deviations (cf. [5]),

$$P\left(\sum_{i=1}^{n} Z_i > 0\right) = P\left(\sum_{i=1}^{n} Z_i > n(-\varepsilon+\varepsilon)\right) \approx e^{-ng(\theta^*)}.$$

Direct calculations (simple details omitted) give that

$$\theta^* = \frac{\varepsilon}{(1-x-\varepsilon)(x+\varepsilon)}, \qquad g(\theta^*) = \log\left\{\left(\frac{x}{x+\varepsilon}\right)^{x}\left(\frac{1-x}{1-x-\varepsilon}\right)^{1-x}\right\}.$$

Hence,

$$P(T_n > x+\varepsilon) \approx \left\{\left(\frac{x+\varepsilon}{x}\right)^{nx}\left(\frac{1-x-\varepsilon}{1-x}\right)^{n(1-x)}\right\}.$$

A similar argument gives that

$$P(T_n < x-\varepsilon) \approx \left\{\left(\frac{1-x+\varepsilon}{1-x}\right)^{n(1-x)}\left(\frac{x-\varepsilon}{x}\right)^{nx}\right\}.$$

These two results imply that for any $0 < \varepsilon_1 < \varepsilon_2$ small enough,

$$P(|T_{n,x} - x| > \varepsilon_2) = o(P(|T_{n,x} - x| > \varepsilon_1)),$$

and Theorem 1 finishes the proof. ∎

5. BLEIMANN, BUTZER, AND HAHN OPERATOR

For any $f \in C[0,\infty)$, the BBH functionals are defined by

$$L_n(f,x) = (1+x)^{-n} \sum_{k=0}^{n} f\left(\frac{k}{n-k+1}\right)\binom{n}{k} x^k; \qquad n = 1, 2, \ldots.$$

We can write this as

$$L_n(f,x) = Ef\left(\frac{S_n}{n-S_n+1}\right), \qquad S_n \sim B(n,p), \qquad p = \frac{x}{1+x},$$

and $Y_1, Y_2, \ldots, \overset{\text{iid}}{\sim} B(1,p)$, with $S_n \sim Y_1 + Y_2 + \cdots + Y_n$. Since

$$L_n(t,x) = E\left(\frac{S_n}{n-S_n+1}\right) = x - xp^n \neq x;$$

we use Theorem 1(a) now.

THEOREM 4. *For the BBH operator, if for any $f \in D(I)$, $L_n(f, x) \geqslant f(x)$ for all $n \geqslant 1$ and each $x \in I$ then f is convex. Furthermore, for any bounded continuous function f over I, the following statements are equivalent*:

(1) *f is convex,*

(2) *$L_n(f, x) \geqslant L_{n+1}(f, x)$ for all $n \geqslant 1$ and $x \in I$,*

(3) *$L_n(f, x) \geqslant f(x)$ for all $n \geqslant 1$ and $x \in I$.*

Proof. Della Vecchia [4] showed that if a continuous function f is bounded and convex then (1) implies (2). Hermann [7] showed that $L_n(f, x) \to f(x)$ for any $f \in D(I)$. Thus, we need only show that, for $f \in D(I)$, if $L_n(f, x) \geqslant f(x)$ for all $n \geqslant 1$ and each $x \in I$ then f is convex. Let $T_{n,x} = S_n/(n+1-S_n)$, let $\delta > 0$ be a fixed number, and note that

$$P(T_{n,x} > x+\delta) = P\left(\frac{S_n}{n+1-S_n} > x+\delta\right) = P\left(\frac{S_n}{n} - p > \varepsilon + \varepsilon_n\right),$$

where

$$\varepsilon = \varepsilon(\delta) = \frac{\delta}{(1+x+\delta)(1+x)} > 0 \qquad \text{and} \qquad \varepsilon_n = \frac{x+\delta}{n(1+x+\delta)} > 0.$$

Since ε is an increasing function of $\delta > 0$, for any fixed δ^* which is arbitrarily close to δ and $\delta^* > \delta$, we have

$$P\left(\frac{S_n}{n} - p > \varepsilon(\delta^*)\right) \leqslant P(T_{n,x} > x+\delta) \leqslant P\left(\frac{S_n}{n} - p > \varepsilon(\delta)\right)$$

for all large values of n. So it will be sufficient for us to evaluate

$$P\left(\frac{S_n}{n} - p > \varepsilon\right); \qquad \varepsilon = \frac{\delta}{(1+x+\delta)(1+x)} > 0.$$

Once again, the standard results on large deviations can be applied (direct calculations omitted) to give

$$P\left(\frac{S_n}{n} - p > \varepsilon\right) \approx \exp\left\{-n\left(\frac{x+\delta}{1+x+\delta}\log\left(1+\frac{\delta}{x}\right) - \log\left(\frac{1+x+\delta}{1+x}\right)\right)\right\}.$$

Similarly,

$$P(T_{n,x} < x-\delta) = P(\bar{Z}_n - (-p) > \varepsilon - \varepsilon_n),$$

where

$$\varepsilon = \varepsilon(\delta) = \frac{\delta}{(1+x-\delta)(1+x)}, \qquad \varepsilon_n = \frac{x-\delta}{n(1+x-\delta)}, \qquad \text{and} \qquad \bar{Z}_n = \frac{-S_n}{n}.$$

We make sure that $0 < \delta < x$ so that $\varepsilon_n > 0$ for all n. Since $\varepsilon(\delta)$ is an increasing function of $\delta > 0$, by picking a δ^* arbitrarily close to δ and $\delta^* < \delta$ we have

$$P\left(\bar{Z}_n - (-p) > \varepsilon(\delta)\right) \leqslant P(T_{n,x} < x - \delta) \leqslant P\left(\bar{Z}_n - (-p) > \varepsilon(\delta^*\right)$$

for all large values of n. Hence, it will be sufficient for us to evaluate

$$P\left(\bar{Z}_n - (-p) > \varepsilon\right); \qquad \varepsilon = \frac{\delta}{(1+x-\delta)(1+x)} > 0.$$

Again, standard results on large deviations give that

$$P\left(\bar{Z}_n - (-p) > \varepsilon\right) \approx \exp\left\{-n\left(\frac{\delta - x}{1+x-\delta}\log\left(\frac{x}{x-\delta}\right) - \log\left(\frac{1+x-\delta}{1+x}\right)\right)\right\}.$$

The above results give that, for any $\delta_1 < \delta_2$ sufficiently small, we can find a $v \in (0, 1)$ such that

$$\frac{(P(T_{n,x} - x > \delta_2))^v}{P(T_{n,x} - x > \delta_1)} \to 0, \qquad \text{as} \quad n \to \infty.$$

For the BBH operator, $b_n(x) = L_n(t, x) - x = xp^n$. So, to end the proof, we need only verify that for $0 < c < \delta_1$,

$$\frac{p^n}{P(T_{n,x} - x > c)} \to 0.$$

This means that we need to prove that

$$\frac{x+c}{1+x+c}\log\left(1+\frac{c}{x}\right) - \log\left(\frac{1+x+c}{1+x}\right) + \log\left(\frac{x}{1+x}\right) < 0.$$

However, this is easily seen to be true since $x + c < 1 + x + c$. This finishes the proof. ∎

6. MEYER-KÖNIG ZELLER OPERATOR

For $f \in C[0, 1]$ and $v \in (0, 1)$, the Meyer-König Zeller (MKZ) functionals are defined by

$$M_n(f, v) := (1-v)^{n+1} \sum_{k=0}^{\infty} f\left(\frac{k}{n+k}\right)\binom{n+k}{k} v^k$$
$$= Ef(S_{n+1}/(n+S_{n+1})); \qquad n=1, 2, ...,$$

where S_{n+1} is a negative binomial random variable with parameters $n+1$ and $1-v$. For $v=0$ and $v=1$, $M_n(f, v)$ is taken to equal $f(v)$ for all values of n.

THEOREM 5. *For the MKZ operators, the following statements are equivalent*:

(1) *f is convex*,

(2) *$M_n(f, v) \geqslant M_{n+1}(f, x)$ for each $n \geqslant 1$ and each $v \in [0, 1]$*,

(3) *$M_n(f, v) \geqslant f(v)$ for each $n \geqslant 1$ and each $v \in [0, 1]$.*

Proof. Cheney and Sharma [2] showed that (1) implies (2). A simple probabilistic proof was given by R. A. Khan [14]. It is trivial that (2) implies (3). Now we show that (3) implies (1). Since this operator reproduces constant and linear functions, we need only check the two conditions of Theorem 1. For this let $v \in (0, 1)$ be a fixed number. Pick any $\delta > 0$ small enough so that $\delta < v(1-v)$. Denote by $T_{n, v} = S_{n+1}/(n+S_{n+1})$, $\bar{X}_n = S_n/n$ and $\mu = E(\bar{X}_n) = v/(1-v)$. Note that

$$P(T_{n, v} > v+\delta) = P(\bar{X}_{n+1} - \mu > \varepsilon - \varepsilon_n),$$

where

$$\varepsilon = \varepsilon(\delta) = \frac{\delta}{(1-v)(1-v-\delta)}, \qquad \varepsilon_n = \frac{(v+\delta)}{(n+1)(1-v-\delta)}.$$

We can select a δ^* arbitrarily close to δ with $\delta < \delta^* < v(1-v)$ so that

$$P(\bar{X}_{n+1} - \mu > \varepsilon(\delta)) \leqslant P(T_{n, v} > v+\delta) \leqslant P(\bar{X}_{n+1} - \mu > \varepsilon(\delta^*))$$

for all large values of n. Hence it will be enough for us to find an approximation of the left term. By the usual results on large deviations we have

$$P\left(\bar{X}_{n+1} - \mu > \frac{\delta}{(1-v)(1-v-\delta)}\right) \approx e^{-(n+1)\, g(\theta^*)},$$

where

$$g(\theta^*) = \frac{v+\delta}{1-v-\delta} \ln\left(\frac{v+\delta}{v}\right) - \log\left(\frac{1-v}{1-v-\delta}\right).$$

Similarly, if $\bar{Z}_n = -\bar{X}_n$, we have

$$P(T_{n,v} < v - \delta) = P(\bar{Z}_{n+1} - (-\mu) > \varepsilon + \varepsilon_n)$$

where

$$\varepsilon = \varepsilon(\delta) = \frac{\delta}{(1-v)(1-v+\delta)}; \qquad \varepsilon_n = \frac{(-\delta)}{(n+1)(1-v+\delta)}.$$

Again, it will be sufficient to find a sharp approximation of

$$P\left(\bar{Z}_{n+1} - (-\mu) > \frac{\delta}{(1-v)(1-v+\delta)}\right).$$

And for this we use the usual results on large deviations to get

$$P\left(\bar{Z}_{n+1} - (-\mu) > \frac{\delta}{(1-v)(1-v+\delta)}\right) \approx e^{-(n+1)\,h(\theta^*)},$$

where

$$h(\theta^*) = \frac{\delta - v}{1-v+\delta} \ln\left(\frac{v}{v-\delta}\right) - \log\left(\frac{1-v}{1-v+\delta}\right).$$

These results imply that for any $0 < \delta_1 < \delta_2$ sufficiently small,

$$\frac{P(|T_{n,v} - v| \geqslant \delta_2)}{P(|T_{n,v} - v| \geqslant \delta_1)} \to 0.$$

Now Theorem 1 gives the result. ∎

7. PICARD OPERATOR

For any $f \in D(I)$ where $I = \mathbb{R}$, the Picard operator is defined by

$$P_n(f, x) = \frac{n}{2} \int_{-\infty}^{\infty} f(t)\, e^{-n\,|t-x|}\, dt = Ef(T_{n,x}), \qquad n = 1, 2, \ldots,$$

where $T_{n,x}$ has the above double exponential density. We give one last application of Theorem 1.

THEOREM 6. *For the Picard operator, over the space* $D(\mathbb{R})$ *the following are equivalent*:

(1) f *is convex*,

(2) $P_n(f, x) \geqslant f(x)$ *for each* $x \in \mathbb{R}$ *and* $n \geqslant 1$.

Proof. The Lebesgue dominated convergence theorem implies that $P_n(f, x) \to f(x)$ for each $x \in \mathbb{R}$ and each $f \in D(I)$. This operator reproduces the linear functions. Since

$$P(|T_{n,x} - x| > \delta) = 2e^{-n\delta}, \qquad n = 1, 2, ...,$$

for any $0 < \delta_1 < \delta_2$, we may take $v = (\delta_1 + \delta_2)/(2\delta_2) < 1$ to have

$$\frac{(P(|T_{n,x} - x| > \delta_2))^v}{P(|T_{n,x} - x| > \delta_1)} = \frac{2e^{-n(\delta_1 + \delta_2)/2}}{2e^{-n\delta_1}} = e^{-n(\delta_2 - \delta_1)/2} \to 0.$$

Theorem 1 finishes the proof. ▮

Remark. In the end we should remark that the above results cannot be extended to multivariate approximation operators of the type discussed (without either strengthening the assumptions or weakening some conclusions) as one can easily construct positive linear operators $L_n(f, x, y)$ which converge to $f(x, y)$ from above for each (x, y) in a closed convex set and yet f is not convex.

REFERENCES

1. J. Adell, F. Badia, J. De la Cal, and F. Plo, On the property of monotonic convergence for beta operators (1995). [*Manuscript*]
2. E. W. Cheney and A. Charma, Bernstein power series, *Can. J. Math.* **16** (1964), 241–252.
3. B. Della Vecchia, On monotonicity of some linear positive operators, *in* "Numerical Methods and Approximation Theory III" (G. V. Milovanović, Ed.), pp. 165–178, Nis, 1988.
4. B. Della Vecchia, Some properties of a rational operator of Bernstein-type, *in* "Progress in Approximation Theory" (P. Nevai and A. Pinkus, Eds.), pp. 177–185, Academic Press, 1991.
5. R. Durrett, "Probability: Theory and Examples," Wadsworth & Brooks/Cole, Belmont, California, 1991.
6. W. Feller, "An Introduction to the Theory of Probability and Its Applications," 2nd ed., Vol. II, Wiley, New York, 1971.
7. T. Hermann, On the operator of Bleimann, Butzer and Hahn, *Colloq. Math. Soc. János Bolyai* **58** (1990), 355–360.

8. T. Hermann, Approximation of unbounded functions on unbounded interval, *Acta. Math. Acad. Sci. Hung*. **29** (1977), 393–398.
9. I. Horova, Buskakov operators and convex functions, *Mathematica* **19**, No. 2 (1990), 133–136.
10. I. Horova, Linear positive operators of convex functions, *Mathematica 10* **33**, No. 2 (1968), 275–283.
11. I. Horova and M. Budíková, A note on D. D. Stancu operators, *Richerche Mat*., in press.
12. M. K. Khan, Approximation properties of beta operators, *in* “Progress in Approx. Theory” (P. Nevai and A. Pinkus, Eds.), pp. 483–495, Academic Press, 1991.
13. M. K. Khan and M. Hanif, On some refinements of Jensen’s inequality, *J. Approx. Theory*, (in press) (1996).
14. R. A. Khan, Some probabilistic methods in the theory of approximation operators, *Acta. Math. Sci. Hung*. **39** (1980), 193–203.
15. L. Kosmák, Les polynomes de Bernstein des fonctions convexes, *Mathematica 9*, **32**, No. 1 (1967), 71–75.
16. G. G. Lorentz, “Bernstein Polynomials,” 2nd ed., Chelsea, New York, 1986.
17. E. Moldovan, Observations sur la suite des polynomes de S. N. Bernstein d’une fonction continue, *Mathematica 4*, **27**, No. 2 (1962), 289–292.
18. Z. Ziegler, Linear approximation and generalized convexity, *J. Approx. Theory* **1** (1968), 420–433.