IMPERIAL COLLEGE OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF MATHEMATICS

THE LIMITING SEMIGROUP OF THE BERNSTEIN ITERATES: PROPERTIES AND APPLICATIONS

by

Manuel Rogério de Jesus da Silva

Thesis submitted

for the degree of

DOCTOR OF PHILOSOPHY

in the University of London

June, 1978

ACKNOWLEDGEMENTS

I owe thanks to many people in the writing of this thesis.

To Dr. E. ORTIZ of Imperial College, for suggesting the subject of this research work and for his continual guidance and kind supervision.

To Prof. R. NUNES of OPorto University. Without his advice, encouragement, and assistance this research could not have been possible.

To Mr. G. ARCE of Imperial College, for his help on computing matters.

To the London Mathematical Society, for a grant that made it possible for me to read some of the results contained in this thesis at their Durham symposium on multivariate approximation.

To INIC (Instituto Nacional de Investigação Científica - Lisboa - Portugal), for their financial support.

Finally, to my wife and children, who have suffered through long hours of neglect.

ABSTRACT

Our main purpose in this thesis is to study the properties of the limiting semigroup of the iterates of the Bernstein polynomials as well as to give some applications.

Chapter 1 reviews some well-known results on the Bernstein approximation theory: connection with the translation semigroup, smoothing effects, variation diminishing properties, etc.; gives some apparently new interpretations of less well-known results, namely, the recursive calculation of the Bernstein polynomial and its derivatives; and extends the Bernstein construction to the approximation of continuous, multivariate, real-valued and vector-valued functions.

Chapter 2 offers a new approach to the numerical condensation of a given multivariate polynomial P as a natural extension of Lanczos' telescoping technique; gives sufficient conditions for the existence of condensed forms of P, and an algorithm for their step by step computation; applies these considerations to Bernstein-Bézier approximants, and gives several examples on the shape approximation problem in one and two dimensions.

In Chapter 3 the Bernstein operator is regarded as a linear transformation onto the space of algebraic polynomials with real coefficients and degree at most n, and the properties of its iterates of nonnegative

order are studied from a fairly elementary matrix analysis standpoint. These iterates are shown to be contractive, variation diminishing, convexity preserving, and convergent to a limiting operator which is explicitly given and shown to be totally positive.

Chapter 4 re-interprets the limiting results of Chapter 3 in the context of the operator semigroup theory as an alternative approach to Karlin-Ziegler's identification and representation of the limiting semigroup of Bernstein iterates of nonnegative order. We give here some new applications of this semigroup, namely, the approximation properties of two new operators of de Leeuw's and Micchelli's type and the characterization of the linear operators commuting with the bivariate Bernstein polynomials.

Finally, Chapter 5 parallels, for the Szász operators, the analysis carried out in Chapters 1, 3, and 4 for the Bernstein polynomials. We show that they are totally positive and give their saturation theory as another application of the operator semigroup method.

CONTENTS

SYMBOLS AND NOTATION

SYMBOL	MEANING
$\in$	is, belongs to
$\Longrightarrow$	implies
$\Longleftrightarrow$	iff, if and only if
$\longrightarrow$	approaches
$\downarrow$	approaches from above
$\uparrow$	approaches from below
$[\ ]$	largest integer $\leqslant$
$\mathbb{N}$	natural numbers
$\mathbb{R}$	real numbers
$\mathbb{R}_N$	N-dimensional Euclidean space
C	$C[0,1]$, continuous on that range
$C^{(n)}$	continuous with derivatives of order $\leqslant n$
I	identity operator
$\mathcal{P}$	algebraic polynomials with real coefficients
$\mathcal{P}_n$	algebraic polynomials with real coefficients and degree $\leqslant n$
$\mathcal{P}_{n_1,\dots,n_N}$	multivariate algebraic polynomials with real coefficients and degree n_i in the variable x_i, $i = 1(1)N$
TP	totally positive
STP	strictly TP
ΔTP	triangular and TP
ΔSTP	triangular and STP

INTRODUCTION

Ever since S. N. Bernstein introduced in 1912 his celebrated polynomials to give a constructive proof of the Weierstrass uniform approximation theorem, they have been the starting point of many investigations.

The fascinating approximation properties of the Bernstein polynomials and the fundamental works of S. Karlin on total positivity, of G. Lorentz and C. Micchelli on the Bernstein saturation problem, of I. Schoenberg on variation diminishing approximation methods, and, above all, the pioneer work of R. Kelisky and T. Rivlin on the iterates of the Bernstein polynomials have had a decisive influence on this thesis.

Chapter 1 deals on the whole with properties and applications of the Bernstein approximation to continuous, real-valued and vector-valued functions.

The trivial observation that the Bernstein polynomial $B_n f$ of a given f in C can be written as the product of n averagings or n forward linear shiftings (see Lemma 1.1) leads immediately to:

(i) The well-known conventional polynomial form of $B_n f$ and its interpolation properties (Section 1).

(ii) The recursive calculation of $B_n f$ and its derivatives $B_n^{(j)} f$ (Sections 1 and 4).

(iii) The uniform convergence results that $B_n^{(j)}(f;x) \longrightarrow f^{(j)}(x)$, $j = 0,1,\ldots$ as $n \longrightarrow \infty$ (Section 2).

(iv) An immediate extension of the foregoing results to the N-dimensional Bernstein polynomials (Section 3).

(v) An elementary and straightforward construction of the Bernstein approximation theory: Smoothing effects, variation diminishing properties, etc. (Section 4).

(vi) An easy extension to cover the Bézier methods, namely, the recursive construction of Bézier curves and their derivatives, variation diminishing properties, Bézier iterates, etc. (Section 5).

Apparently, Lemma 1.1 does not appear published anywhere in the vast literature on Bernstein polynomials. Its main interest lies in the ease with which the semiclassical Bernstein approximation theory is generalized to the multivariate and parametric cases.

There is no new material in Section 2, which connects $B_n f$ with the translation semigroup in $C[0,\infty]$.

In Section 3, the recursive generation of multivariate Bernstein polynomials and their convex hull and interpolation properties are essentially contained in Gordon and Riesenfeld (1974 a)).

Since derivatives of Bernstein polynomials are also Bernstein polynomials (of another function!), then their recursive calculation and geometric representation, which Section 4 deals with, may also be seen as essentially contained in Gordon and Riesenfeld (1974 a)).

The recursive construction of the matrix $A_{N+1}(n)$ representing B_n

acting on $\mathcal{P}_N$ and the observation that, thanks to the smoothing effects of B_n, any interpolating sequence at equidistant nodes can always be made uniformly convergent appear to be new.

The recursive construction of Bézier curves and surfaces, their convex hull and interpolation properties, the recursive calculation and geometric representation of derivatives of Bézier polynomials, which are dealt with in Section 5, are all, once again, essentially contained in Gordon and Riesenfeld (1974 a), b)).

Finally, following G.-Bonne and Sablonnière (1976), we extend to the Bézier operator the variation diminishing properties of the Bernstein polynomial, which are due to Pólya and Schoenberg (1958).

In Chapter 2 we offer a new approach to the numerical problem of condensing (telescoping) a given multivariate polynomial $P = P(x_1,x_2,\ldots,x_N)$ defined on the unit hypercube of R_N, leading to a considerable simplification of the work required to perform it (cf. E. Ortiz (1977) and E. Ortiz and M. da Silva (1978)). In particular, we try to avoid polynomial basis transformations, and practical apriori tests for the existence of a condensed representation of P appear naturally as immediate extensions of the univariate case.

Section 1 offers a new algorithm for step by step computation of a condensed representation of a given $P \in \mathcal{P}_n$.

Section 2 extends this algorithm to $P \in \mathcal{P}_{I_1,I_2,\ldots,I_N}$ (the linear space of multivariate polynomials with real coefficients and degree I_k in x_k, $k = 1(1)N$); uses the numbers s_k of condensation steps to measure the smoothness of P in the x_k-directions, and to define the principal variables of P; and deals with the problem of approximating a given multivariate polynomial by another polynomial of fewer variables.

Section 3 applies the above considerations to Bernstein-Bézier approximants and gives several numerical examples on the shape approximation problem in one and two dimensions.

The greater part of the material of Chapter 3 is seemingly new. It has been largely inspired by Kelisky and Rivlin (1967), who were the first to study the convergence of the iterates of Bernstein polynomials $B_n^r(f;x)$ as $r \longrightarrow \infty$, both in the case that r is independent of n and, for polynomial f, when r is a function of n. They have treated only these convergence problems, leaving, therefore, scope for more work, namely on properties and applications. We deduce here the properties of the Bernstein iterates of all orders using only elementary matrix methods. We show that the operators B_n^r, $r > 0$, are contractive, variation diminishing, norm not increasing, and convergent to a limiting operator, which, in each of the following cases:

i) n fixed, $r \longrightarrow \infty$

ii) $n \longrightarrow \infty$, $r_m \longrightarrow t \in \mathbb{R}$, fixed, as $m \longrightarrow \infty$ independently of n

iii) $r = r_n \longrightarrow \infty$,

is explicitly given and shown to be totally positive.

Section 1 puts the Bernstein generalized iteration problem in the context of matrix-valued functions. Arbitrary iterates are immediate extensions of those of natural order, and these are simply reduced to matrix multiplications.

Section 2 offers the apparently new results that the matrix representation of B_n acting on $\mathcal{P}_N$ and the triangle of Stirling numbers of the second kind are both totally positive.

Section 3 deals with the positivity of B_n^r , $r>0$, and shows that its matrix representation is column-stochastic for all sufficiently large r (Theorem 3.1).

Section 4 gives a neater and richer theory of the limiting behaviour of B_n^r , $r>0$, than that in Kelisky and Rivlin (1967). In particular, Theorem 4.1 throws light into the structural limiting properties of the matrix representation of B_n^r , $r>0$, and Theorem 4.2 enlarges and gives more insight into the meaning of certain seemingly nontrivial identities first observed by those authors.

Finally, we discuss in Section 5 the convexity preserving properties of the arbitrary Bernstein iterates. We show that B_2^r is convexity preserving for each real r and that B_n^r is convexity preserving or nearly so for all $r>0$ and $n>2$.

Sections 1 and 2 of Chapter 4 are essentially of conceptual value. We offer an alternative approach to Karlin and Ziegler (1970)'s identification and representation of the limiting semigroup $\{\mathcal{B}_t;\ t\geqslant 0\}$ of the Bernstein iterates of nonnegative order. Our approach does not rely on diffusion theory arguments as Karlin and Ziegler's but re-interprets the limiting theory of Chapter 3 in the context of the operator semigroup theory. Some results reproduce Karlin and Ziegler's, although in a more straightforward fashion, and some extensions are shown possible with our approach.

The existence of the limit

$$\mathcal{B}_t P(x) = \lim_{n\to\infty} B_n^{[nt]} P(x) , \qquad P\in\mathcal{P}_N , \qquad N\leqslant n ,$$

its total positivity and semigroup properties, and the infinitesimal generator

$$D P(x) = \lim_{n\to\infty} n(B_n - I) P(x) = \tfrac{1}{2} x(1-x) d^2P/dx^2 ,$$

all follow from the limiting theory of Chapter 3.

To extend $\mathcal{B}_t$ to C we define $W(t,x) = \mathcal{B}_t(f;x)$, are naturally led to the classical diffusion problem

$$\partial W/\partial t = D W , \qquad W = W(t,x) ,$$
$$W(0,x) = f(x) ,$$

and we find for $\mathcal{B}_t(f;x)$ the integral representation

$$\mathcal{B}_t(f;x) = \int_0^1 G(t;x,y) f(y) dy \qquad (f(0)=f(1)=0)$$

with the kernel G expressed in terms of the shifted Jacobi orthogonal polynomials of parameters (1,1).

A fundamental property of G is its total positivity, which implies that $\mathcal{B}_t$ inherits from B_n its shape preserving properties. This appears to be new, though essentially contained in Karlin and McGregor (1960) and Butzer and Nessel (1971).

Section 3 reviews some known applications of $\mathcal{B}_t$, namely, the Lorentz-Micchelli's treatment of the Bernstein saturation problem, the Karlin-Ziegler-Micchelli's characterizations of convexity, and the Karlin-Ziegler's identification of the linear operators commuting with B_n, and offers some new applications of $\mathcal{B}_t$, these are:

i) The saturation theory for the de Leeuw-like operators

$$K_n f = \sum_{k=0}^{n} l_{nk}^{*}(f)\, q_k ,$$

where $l_{nk}^{*}(f)$ are some linear functionals on f and $\{q_k\}_{k=0}^{n}$ is the Bernstein basis for $\mathcal{P}_n$. We show that $K_n^{[nt]} f \longrightarrow \mathcal{B}_t f$ strongly for all f in C, and that K_n and B_n have exactly the same saturation properties.

ii) The approximation properties of a new Micchelli's type operator which takes into account the spectral characteristics of B_n and leaves $\mathcal{P}_n$ intact.

iii) The generalization of our construction to the multivariate setting to receive a number of results as natural extensions of the univariate case, e.g., the identification of the linear operators commuting with the bivariate Bernstein polynomials.

Lastly, an example of the applicability of our technique is afforded by the Szász operators. We show in Chapter 5 their total positivity by working on the lines set out in Chapters 1, 3, and 4. Their saturation properties are also given as an application of the iteration method, reproducing, however, results already given by Suzuki (1967) by a different method.

CHAPTER 1

BERNSTEIN APPROXIMATION TO CONTINUOUS, REAL-VALUED, AND VECTOR-VALUED FUNCTIONS

1. BERNSTEIN APPROXIMATION TO REAL-VALUED FUNCTIONS

The n^{th} degree Bernstein polynomial approximation to a real $f(x)$ defined on $[0,1]$ is given by

$$B_n f(x) \equiv B_n(f;x) = \sum_{k=0}^{n} f_k \binom{n}{k} x^k (1-x)^{n-k}, \qquad n \geqslant 1, \qquad (1.1)$$

$$B_0 f(x) = f_0 ,$$

where $f_k = f(k/n)$, $k=0(1)n$. The polynomials

$$q_k = q_k(n,x) = \binom{n}{k} x^k (1-x)^{n-k}, \qquad k=0(1)n, \qquad (1.2)$$

form the Bernstein basis for $\mathcal{P}_n$ and are well-known to enjoy the following properties:

a) $\quad q_k \geqslant 0 , \qquad k=0(1)n ,$

b) $\quad \sum_{k=0}^{n} q_k = 1 ,$ $\qquad (1.3)$

c) $\quad \sum_{k=0}^{n} \frac{k}{n} q_k \equiv x .$

1.1. The convex hull property.

Owing to the properties (1.3), the graph of $B_n f$ develops within the convex hull of the points $\{(k/n, f_k)\}_{k=0}^{n}$. To be more precise, regarding the basic polynomials q_k as masses attached to the points $(k/n, f_k)$, the center of mass of those mass points describes the graph of $B_n f$ as x traverses $[0,1]$.

This elegant interpretation of the Bernstein construction is due to Gordon and Riesenfeld (1974 a)).

1.2. Recursive generation.

Making use of the fundamental operator in Finite Difference Calculus, namely, the forward shifting operator E defined from

$$Ef_k = f_{k+1}, \qquad k=0,1,\ldots,$$

and the forward difference operator Δ given by

$$\Delta f_k = f_{k+1} - f_k = (E - I)f_k, \qquad k=0,1,\ldots,$$

we may replace in (1.1) f_k with $E^k f_0$ to obtain

LEMMA 1.1. The Bernstein approximation to any given real-valued function f taking on the values f_k at the nodes k/n, $k=0(1)n$, is given by

a) $\quad B_n(f;x) = ((1-x)I + xE)^n f_0$

b) $B_n(f;x) = (I + x\Delta)^n f_0$.

From a), B_n is the product of n averagings:

$$\mathbb{B}_x f_k = (1-x)f_k + xf_{k+1} , \qquad k=0(1)n-1 ; \tag{1.4}$$

while b) shows that B_n is the product of n forward linear shiftings:

$$\mathbb{B}_x f_k = f_k + x(f_{k+1} - f_k) , \qquad k=0(1)n-1 . \tag{1.5}$$

Clearly,

$$B_n(f;x) = \mathbb{B}_x^n \, f_0 . \tag{1.6}$$

As immediate consequences of Lemma 1.1 we have

i) a numerical procedure for the recursive generation of $B_n(f;x)$. Indeed, given the table $\{k/n, f_k\}_{k=0}^{n}$, we construct the numerical triangle

$$\begin{array}{lllll}
f_0 & & & & \\
 & f_0^{(1)} & & & \\
f_1 & & f_0^{(2)} & & \\
 & f_1^{(1)} & \vdots & \ddots & \\
f_2 & \vdots & \vdots & & \\
\vdots & \vdots & \vdots & & f_0^{(n)} \\
 & \vdots & \vdots & ⋰ & \\
f_{n-2} & \vdots & \vdots & & \\
 & f_{n-2}^{(1)} & \vdots & & \\
f_{n-1} & & f_{n-2}^{(2)} & & \\
 & f_{n-1}^{(1)} & & & \\
f_n & & & &
\end{array} \tag{1.7}$$

with column entries $f_i^{(j)}$ given by

$$f_i^{(j)} = \mathbb{B}_x f_i^{(j-1)} \ , \quad j=1(1)n \ ; \ i=0(1)n-j \ ,$$
$$f_i^{(0)} = f_i \quad , \quad i=0(1)n \ , \tag{1.8}$$

and whose vertex $f_0^{(n)}$ is $B_n(f;x)$ by (1.8) and (1.6);

ii) the conventional polynomial form

$$B_n(f;x) = \sum_{k=0}^{n} \binom{n}{k} \Delta^k f_0 \, x^k \ , \tag{1.9}$$

which in turn implies that if $f \in \mathcal{P}_m$ then $B_n f \in \mathcal{P}_{\min\{m,n\}}$;

iii) since $\mathbb{B}_0 = I$ and $\mathbb{B}_1 = E$, then

$$B_n(f;0) = f_0 = f(0)$$
$$B_n(f;1) = E^n f_0 = f_n = f(1) \ , \tag{1.10}$$

the well-known result that $B_n f$ interpolates to f at the endpoints of $[0,1]$.

2. THE BERNSTEIN UNIFORM APPROXIMATION THEOREM AND THE TRANSLATION SEMIGROUP

Let X be a Banach space endowed with norm $\|\cdot\|$ and let $\mathcal{T} = \{T_t \ ; \ t \geq 0\}$ be a one-parameter family of linear bounded transformations on X to itself with the property

$$T_{s+t} = T_s T_t \ , \quad s,t \geq 0 \ .$$

We then speak of $\mathcal{T}$ as an operator semigroup.

THEOREM 2.1 (Kendall). If $\mathcal{T}$ is continuous in the strong operator topology for $t \geq 0$, then

$$\lim_{n \to \infty} \left\| ((1-t)I + tT_{1/n})^n f - T_t f \right\| = 0 \tag{2.1}$$

for each f in X and each t in $[0,1]$, uniformly in t.

PROOF. See D.G. Kendall (1954).

Let $X = C[0,\infty]$, the Banach space of real-valued, continuous, bounded functions on $[0,\infty]$ normed by

$$\|f\| = \sup_{0 \leq x \leq \infty} |f(x)| ,$$

and let $\mathcal{T}$ denote the semigroup of translations in $C[0,\infty]$:

$$T_t f(x) = f(x+t) \ .$$

In this case, (2.1) shows that, for $0 \leqslant t \leqslant 1$,

$$f(x+t) = \lim_{n \to \infty} \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} f(x+\tfrac{k}{n}) , \tag{2.2}$$

where the limit exists uniformly with respect to x in $[0,\infty]$ and t in $[0,1]$. In particular, for x = 0, (2.2) gives

$$\lim_{n \to \infty} B_n(f;t) = f(t) , \quad 0 \leqslant t \leqslant 1 , \tag{2.3}$$

the well-known Bernstein uniform approximation theorem.

REMARK 1.1. For each fixed but arbitrary integer $j \geqslant 0$, it follows from Lemma 1.1 b) that

$$B_n^{(j)}(f;x) = \lambda_j (n\Delta_{1/n})^j (I + x\Delta_{1/n})^{n-j} f(0)$$

with

$$\lambda_0 = 1$$

and

$$\lambda_j = (1-\tfrac{1}{n})(1-\tfrac{2}{n})\ldots(1-\tfrac{j-1}{n}) = 1 - O(1/n) .$$

We recall that $n\Delta_{1/n} = d/dx + O(1/n)$ and observe that $\Delta_{1/n} \sim \Delta_{1/N}$, $N = n-j$, in the sense that their effects on f, assumed differentiable in (0,1), have the same limit as $n \to \infty$. Indeed,

$$\begin{aligned} (\Delta_{1/N} - \Delta_{1/n})f(x) &= f(x + 1/N) - f(x + 1/n) \\ &= f(x + 1/n + j/(nN)) - f(x + 1/n) \\ &= \frac{j}{nN} f'(y) , \quad x + 1/n < y < x + 1/N , \\ &= O(1/n^2) . \end{aligned}$$

Therefore,

$$(I + x\Delta_{1/n})^{n-j} f(0) \sim (I + x\Delta_{1/N})^{N} f(0) = B_N(f;x) \longrightarrow f(x) ,$$

and we have the well-known result that

$$B_n^{(j)} (f;x) \longrightarrow f^{(j)}(x) \quad \text{as} \quad n \longrightarrow \infty,$$

(cf. Lorentz (1953, p.12)) at each point $x \in [0,1]$ where $f^{(j)}(x)$ exists, the convergence being uniform provided $f^{(j)}$ is continuous.

3. BERNSTEIN APPROXIMATION TO CONTINUOUS, MULTIVARIATE FUNCTIONS

The results in the preceding sections afford a straightforward generalization to continuous functions of more than one variable. Let $f \in C[S_2]$, $S_2 = \{(x,y) \in \mathbb{R}_2 : 0 \leqslant x, y \leqslant 1\}$, be given. Then $B_{n,m}f(x,y) \equiv B_{n,m}(f;x,y)$, the Bernstein polynomial of n^{th} degree in x and m^{th} degree in y associated with $f(x,y)$, may be obtained by applying twice the well-known univariate Bernstein polynomial approximation formula. Regarding, for the moment, x as a parameter and y as the operational variable, we have

$$f(x,y) \approx B_m f(x,y) = \mathbb{B}_y^m f(x,0) ,$$

with

$$\mathbb{B}_y = I + y\Delta_k , \quad k = 1/m ,$$

$$= (1-y)I + yE_k .$$

The same approximation formula, applied this time to the variable x, gives

$$f(x,0) \approx B_n f(x,0) = \mathbb{B}_x^n f(0,0) ,$$

with

$$\mathbb{B}_x = I + x\Delta_h , \quad h = 1/n ,$$

$$= (1-x)I + xE_h .$$

Thus we have

$$f(x,y) \approx B_{n,m} f(x,y) = \mathbb{B}_y^m \mathbb{B}_x^n f(0,0) .$$

Had we started with y as a parameter and x as the operational variable, we would have ended up with

$$f(x,y) \approx B_{n,m}f(x,y) = \mathbb{B}_x^n \mathbb{B}_y^m f(0,0) \ .$$

Therefore

$$B_{n,m}f(x,y) = \mathbb{B}_x^n \mathbb{B}_y^m f(0,0) = \mathbb{B}_y^m \mathbb{B}_x^n f(0,0) \ , \qquad (3.1)$$

and we conclude that the bivariate Bernstein operator $B_{n,m}$ is simply the product of the commutative univariate operators B_n and B_m (cf. Gordon and Riesenfeld (1974 a))).

In order to extend this to higher dimensions we associate with each f in $C\left[S_N\right]$, S_N the unit hypercube of $\mathbb{R}_N$, $N \geqslant 1$, the N-dimensional Bernstein polynomial of degree n_i in x_i, $i=1(1)N$,

$$B_{\underline{n}}f(\underline{x}) = B_{\underline{n}}(f;S_N;\underline{x}) = B_{n_1,n_2,\ldots,n_N}(f;S_N;x_1,x_2,\ldots,x_N)$$

$$= \prod_{i=1}^{N} \mathbb{B}_{x_i}^{n_i} f(\underline{0}) \ , \qquad (3.2)$$

where

$$\mathbb{B}_{x_i} = I + x_i \underset{i}{\Delta} \qquad (3.3)$$

with

$$\underset{i}{\Delta} f(\underline{x}) = f(x_1,\ldots,x_{i-1},x_i+\tfrac{1}{n},x_{i+1},\ldots,x_N) - f(x_1,\ldots,x_i,\ldots,x_N) \ ,$$

or, equivalently,

$$\mathbb{B}_{x_i} = (1-x_i)I + x_i \underset{i}{\mathbb{E}} \qquad (3.4)$$

with

$$\underset{i}{\mathbb{E}} f(\underline{x}) = f(x_1,\ldots,x_{i-1},x_i+\frac{1}{n},x_{i+1},\ldots,x_N) \ .$$

It follows from (3.2) and (3.3) that

$$B_{\underline{n}}f(\underline{x}) = \prod_{i=1}^{N} \left\{ \sum_{j_i=0}^{n_i} \binom{n_i}{j_i} x_i^{j_i} \Delta_i^{j_i} f(\underline{0}) \right\}$$

$$= \sum_{j_1=0}^{n_1} \cdots \sum_{j_N=0}^{n_N} \left\{ \binom{n_1}{j_1} \cdots \binom{n_N}{j_N} \Delta_1^{j_1} \cdots \Delta_N^{j_N} f(\underline{0}) \right\} x_1^{j_1} \cdots x_N^{j_N}, \qquad (3.5)$$

generalizing (1.9) and implying that if

$$f \in \mathcal{P}_{\underline{r}} \equiv \text{span} \left\{ x_1^{i_1}, x_2^{i_2}, \ldots, x_N^{i_N} \right\}_{i_1=0,\ i_2=0,\ \ldots,\ i_N=0}^{r_1,\ \ r_2,\ \ \ldots,\ r_N}$$

then $B_{\underline{n}} f \in \mathcal{P}_{\underline{m}}$, $\underline{m}$ standing for $\min\{n_1, r_1\}, \min\{n_2, r_2\}, \ldots, \min\{n_N, r_N\}$.

Similarly, using (3.2) and (3.4), we obtain

$$B_{\underline{n}}f(\underline{x}) = \prod_{i=1}^{N} \left\{ \sum_{j_i=0}^{n_i} \binom{n_i}{j_i} (1-x_i)^{n_i-j_i} x_i^{j_i} E_i^{j_i} f(\underline{0}) \right\}$$

$$= \sum_{j_1=0}^{n_1} \cdots \sum_{j_N=0}^{n_N} f\left(\frac{j_1}{n_1}, \ldots, \frac{j_N}{n_N}\right) \left\{ \prod_{i=1}^{N} \binom{n_i}{j_i} (1-x_i)^{n_i-j_i} \right\} x_1^{j_1} \cdots x_N^{j_N} \qquad (3.6)$$

$$= \sum_{j_1=0}^{n_1} \cdots \sum_{j_N=0}^{n_N} f\left(\frac{j_1}{n_1}, \ldots, \frac{j_N}{n_N}\right) q_{j_1}(n_1, x_1) \cdots q_{j_N}(n_N, x_N) \qquad (3.7)$$

extending (1.1).

<u>REMARK 3.1</u>. Gordon's mechanical interpretation of the univariate Bernstein polynomial (see Subsection 1.1) affords an easy extension to the multivariate setting. Indeed, setting

$$M_{\underline{i}} = \prod_{k=1}^{N} q_{i_k}(n_k, x_k)$$

and

$$f_{\underline{i}} = f\left(\frac{i_1}{n_1}, \ldots, \frac{i_N}{n_N}\right)$$

then, in view of the easily verified properties

$$M_{\underline{i}} \geqslant 0 \, , \quad \sum M_{\underline{i}} = 1 \, ,$$

$$\sum \frac{i_k}{n_k} M_{\underline{i}} = x_k \, , \qquad k=1(1)N \, ,$$

$$\sum f_{\underline{i}} \, M_{\underline{i}} = B_{\underline{n}} f(\underline{x}) \, ,$$

the summations being assumed over $i_1=0(1)n_1$, $i_2=0(1)n_2$, ... , $i_N=0(1)n_N$, the center of mass of the points $(i_1/n_1, \ldots, i_N/n_N, f_{\underline{i}})$ with masses $M_{\underline{i}}$ describes the graph of $B_{\underline{n}}f(\underline{x})$ as $\underline{x}$ runs over S_N.

<u>REMARK 3.2</u>. Being the product of $n_1 + n_2 + \ldots + n_N$ averagings, $B_{\underline{n}}f(\underline{x})$ can be generated recursively by means of N triangular schemes similar to (1.7).

By way of example, we take $N = 2$ and construct $B_{n,m}f(x,y)$. Two numerical triangles have to be formed. The first,

$$\begin{array}{llll}
f(0,y) & & & \\
 & f^{(1)}(0,y) & & \\
f(\frac{1}{n},y) & \vdots & \cdot \, \cdot \, \cdot & \\
\vdots & \vdots & & f^{(n)}(0,y) \\
f(\frac{n-1}{n},y) & \vdots & \cdot \, \cdot \, \cdot & \\
 & f^{(1)}(\frac{n-1}{n},y) & & \\
f(1,y) & & &
\end{array}$$

with column entries $f^{(j)}(\frac{i}{n},y)$ given by

$$f^{(j)}\left(\frac{i}{n},y\right) = \mathbb{B}_x f^{(j-1)}\left(\frac{i}{n},y\right) , \quad j=1(1)n \; ; \; i=0(1)n-j ,$$

$$f^{(0)}\left(\frac{i}{n},y\right) = f\left(\frac{i}{n},y\right) , \quad i=0(1)n ,$$

having the vertex

$$P(x,y) = f^{(n)}(0,y) = \mathbb{B}_x^n\, f(0,y) = B_n f(x,y) ,$$

and the second,

$$\begin{array}{lll} P(x,0) & & \\ & P^{(1)}(x,0) & \\ P(x,\frac{1}{m}) & \vdots & \ddots \\ \vdots & \vdots & \quad P^{(m)}(x,0) , \\ P(x,\frac{m-1}{m}) & \vdots & \\ & P^{(1)}(x,\frac{m-1}{m}) & \\ P(x,1) & & \end{array}$$

with

$$P^{(j)}\left(x,\frac{i}{m}\right) = \mathbb{B}_y P^{(j-1)}\left(x,\frac{i}{m}\right) , \quad j=1(1)m \; ; \; i=0(1)m-j ,$$

$$P^{(0)}\left(x,\frac{i}{m}\right) = P\left(x,\frac{i}{m}\right) , \quad i=0(1)m ,$$

and

$$P^{(m)}(x,0) = B_{n,m} f(x,y) .$$

REMARK 3.3. The N-dimensional Bernstein polynomial $B_{\underline{n}} f(\underline{x})$ interpolates to f at the vertices of S_N.

Taking again N = 2, we now have, from (3.1) and in correspondence with the interpolation results (1.10):

$$B_{n,m}f(0,0) = f(0,0)$$

$$B_{n,m}f(1,0) = E_h^n\, f(0,0) = f(1,0)$$

$$B_{n,m}f(0,1) = E_k^m\, f(0,0) = f(0,1)$$

$$B_{n,m}f(1,1) = E_h^n\, E_k^m\, f(0,0) = f(1,1)\ ,$$

i.e., $B_{n,m}f$ interpolates to f at the four corners of S_2. Moreover,

$$B_{n,m}f(x,0) = \mathbb{B}_x^n\, f(0,0) = B_n f(x,0)$$

$$B_{n,m}f(x,1) = \mathbb{B}_x^n\, E_k^m\, f(0,0) = \mathbb{B}_x^n\, f(0,1) = B_n f(x,1)$$

$$B_{n,m}f(0,y) = \mathbb{B}_y^m\, f(0,0) = B_m f(0,y)$$

$$B_{n,m}f(1,y) = \mathbb{B}_y^m\, E_h^n\, f(0,0) = \mathbb{B}_y^m\, f(1,0) = B_m f(1,y)\ ,$$

i.e., the bivariate Bernstein polynomial approximation to f(x,y) reduces to the appropriate univariate one on each side of S_2 (cf. Gordon and Riesenfeld (1974 a))).

REMARK 3.4. Since $B_{\underline{n}}(f;S_N;\underline{x})$ can be factored into N univariate Bernstein polynomials, each of which converging uniformly in the unit interval, then

$$B_{\underline{n}}(f;S_N;\underline{x}) \longrightarrow f(\underline{x}) \quad \text{as} \quad \underline{n} \longrightarrow \infty\ ,$$

i.e., as $n_i \longrightarrow \infty$ for i=1(1)N, uniformly in $\underline{x}$ in S_N.

REMARK 3.5. The operators $B_{\underline{n}}(f;S_N;\underline{x})$ are a particular case of the following linear positive operators $L_{\underline{n}}(f;K_N;\underline{x})$ introduced by Schurer (1963) for the approximation of multivariate functions, continuous on the region K_N of the first hyperquadrant of $\mathbb{R}_N$:

$$L_{\underline{n}}(f;K_N;\underline{x}) = \sum_{j_1=0}^{\infty} \cdots \sum_{j_N=0}^{\infty} f\left(\frac{j_1}{n_1},\ldots,\frac{j_N}{n_N}\right)\left\{\frac{(-1)^{j_1+\ldots+j_N}}{j_1!\ldots j_N!}\phi_{\underline{n}}^{(\underline{j})}(\underline{x})\right\} x_1^{j_1}\ldots x_N^{j_N}, \tag{3.8}$$

$$\longrightarrow f(\underline{x}) \quad \text{as} \quad \underline{n}\longrightarrow\infty \text{ uniformly in } \underline{x} \text{ in } K_N ,$$

where

$$\phi_{\underline{n}}^{(\underline{j})}(\underline{x}) = \frac{\partial^{j_1+\ldots+j_N}}{\partial x_1^{j_1}\ldots\partial x_N^{j_N}}\phi_{\underline{n}}(\underline{x})$$

and $\phi_{\underline{n}}(\underline{x})$, called the generating function, is such that

a) $\phi_{\underline{n}}(\underline{x}) \in C^{\infty}(K_N)$,

b) $\phi_{\underline{n}}(\underline{0}) = 1$,

c) $(-1)^{j_1+\ldots+j_N}\phi_{\underline{n}}^{(\underline{j})}(\underline{x}) \geqslant 0$, $j_1,\ldots,j_N = 0,1,\ldots$; $\underline{x}\in K_N$,

d) $-\phi_{\underline{n}}^{(\underline{j})}(\underline{x}) = n_i\,\phi_{\underline{n}}^{(\underline{j}-e_i)}(\underline{x})\left\{1 + \alpha_{n_i}(\underline{x})\right\}$,

where $\underline{j} - e_i$ stands for $j_1, \ldots , j_{i-1}, j_i - 1, j_{i+1}, \ldots ,j_N$ and, for $i = 1(1)N$, $\alpha_{n_i}(\underline{x}) \longrightarrow 0$ uniformly in $\underline{x}$ in K_N if $n_i \longrightarrow \infty$.

If we take

$$\phi_{\underline{n}}(\underline{x}) = \prod_{i=1}^{N}(1-x_i)^{n_i} , \qquad \underline{x}\in K_N ,$$

then

$$\phi_{\underline{n}}^{(\underline{j})}(\underline{x}) = (-1)^{j_1+\ldots+j_N}\, j_1!\,\ldots\, j_N!\,\prod_{i=1}^{N}\binom{n_i}{j_i}(1-x_i)^{n_i-j_i} ,$$

i.e.,

$$\prod_{i=1}^{N}\binom{n_i}{j_i}(1-x_i)^{n_i-j_i} = \frac{(-1)^{j_1+\ldots+j_N}}{j_1!\ldots j_N!}\phi_{\underline{n}}^{(\underline{j})}(\underline{x}) ,$$

and thus (3.6) is included in (3.8), i.e., the N-dimensional Bernstein polynomial operators $B_{\underline{n}}$ are a particular instance of Schurer's $L_{\underline{n}}$.

In the sequel we take, for simplicity, $n_i = n$, $i = 1(1)N$, and write

$$B_n^*(f;\underline{x}) = B_{\underline{n}}(f;S_N;\underline{x})$$

$$= \sum_{j_1=0}^{n}\cdots\sum_{j_N=0}^{n}\left\{\binom{n}{j_1}\cdots\binom{n}{j_N}\Delta_1^{j_1}\cdots\Delta_N^{j_N} f(\underline{0})\right\}x_1^{j_1}\cdots x_N^{j_N} \quad (3.9)$$

by (3.5),

$$= \sum_{j_1=0}^{n}\cdots\sum_{j_N=0}^{n} f\left(\frac{j_1}{n},\ldots,\frac{j_N}{n}\right) q_{j_1}(n,x_1)\ldots q_{j_N}(n,x_N) \quad (3.10)$$

by (3.7).

N-dimensional Bernstein polynomials may also be associated with multivariate functions $f(\underline{x})$ defined on other domains of $\mathbb{R}_N$, e.g., $T_N = \left\{\underline{x} = (x_1, x_2, \ldots, x_N) \in \mathbb{R}_N : x_i \geq 0,\ x_1 + x_2 + \ldots + x_N \leq 1,\ 1 \leq i \leq N\right\}$:

$$B_{\underline{n}}(f;T_N;\underline{x}) = \sum_{\substack{i_1+\ldots+i_N\leq n\\ i_k\geq 0}} f\left(\frac{i_1}{n_1},\ldots,\frac{i_N}{n_N}\right)\frac{n!}{i_1!\ldots i_N!(n-i_1-\ldots-i_N)!}\cdot x_1^{i_1}\ldots x_N^{i_N}(1-x_1-\ldots-x_N)^{n-i_1-\ldots-i_N} . \quad (3.11)$$

See Lorentz (1953, p.51) and Schurer (1962,1963). See also Stancu (1960 a), b)) for probabilistic interpretations of these generalized Bernstein polynomials. See also Stancu (1963 a)) for a particularly nice approach to defining bivariate Bernstein polynomials on domains given by the equations

$$x = 0,\ x = 1,\ y = u_1(x),\ \text{and}\ y = u_2(x)\ ,$$

where u_1 and u_2 are polynomials such that $0 \leq u_1(x) \leq u_2(x)$ for $0 \leq x \leq 1$, leading to the problem of rational Bernstein type approximation.

4. SMOOTHING PROPERTIES OF THE BERNSTEIN OPERATOR

4.1. Derivatives of Bernstein polynomials.

For each fixed but arbitrary integer $j \geq 0$ and any given $f(x)$ defined on $[0,1]$, it follows from Lemma 1.1 that

$$B_n^{(j)}(f;x) = \frac{n!}{(n-j)!}\Delta^j((1-x)I + xE)^{n-j}f_0$$

$$= \frac{n!}{(n-j)!}\sum_{k=0}^{n-j}\binom{n-j}{k}\Delta^j f_k\, x^k(1-x)^{n-j-k}\ , \qquad (4.1)$$

from which

$$\frac{(n-j)!}{n!}B_n^{(j)}(f;x) = \begin{cases} B_{n-j}(F^j;x)\ , & j=0(1)n \\ 0\ , & j>n\ , \end{cases} \qquad (4.2)$$

F^j being such that

$$\begin{aligned} F^j\left(\frac{k}{n-j}\right) &= \Delta^j f_k\ , \quad k=0(1)n-j>0\ , \\ F^{(n)}(0) &= \Delta^n f_0\ . \end{aligned} \qquad (4.3)$$

REMARK 4.1. Noting (4.2), the j^{th} derivative of the n^{th} degree Bernstein polynomial of f is, apart from the coefficient

$(n-j)!/n!$, the $(n-j)^{th}$ degree Bernstein polynomial of the function F^j derived from f according to (4.3). In view of this, not only $B_n f$ but all its derivatives as well afford a recursive generation and an easy geometric construction. For details see Remark 5.2 below.

The following assertions are easily seen to follow from (4.1):

$$\text{(i)} \qquad \begin{aligned} B_n^{(j)}(f;0) &= \frac{n!}{(n-j)!}\Delta^j f_0 \\ B_n^{(j)}(f;1) &= \frac{n!}{(n-j)!}\Delta^j f_{n-j} \end{aligned} \qquad j=0(1)n\ . \tag{4.4}$$

Incidentally, (1.9) follows also at once from (4.4) and Taylor's expansion of $B_n(f;x)$ about 0:

$$B_n(f;x) = \sum_{j=0}^{n} \frac{B_n^{(j)}(f;0)}{j!} x^j = \sum_{j=0}^{n} \binom{n}{j} \Delta^j f_0\, x^j\ .$$

(ii) B_n preserves most of the global characteristics of f, namely, positivity, monotonicity, convexity, etc. (see Pólya and Schoenberg (1958) and Schoenberg (1959)).

(iii) If f is <u>absolutely monotonic</u> in $[0,1]$, i.e.,

$$f^{(j)}(x) \geq 0\ , \quad j=0,1,\ldots\ ;\ 0 \leq x \leq 1\ , \tag{4.5}$$

then, from (4.1), so is $B_n f$. In particular, for any integer $j \geq 0$, the monomial x^j satisfies (4.5) and thus

$$B_n(t^j;x) = \sum_{i=0}^{j} a_{ij}\, x^i$$

where

$$a_{ij} = a_{ij}(n) = \binom{n}{i} \Delta^i_{1/n} t^j \Big|_{t=0} \geq 0 \tag{4.6}$$

$$a_{ij} = \begin{cases} 0 & , \quad i > j \\ \lambda_i n^{i-j} \sigma_{ij} , & 0 \leqslant i \leqslant j \leqslant n , \end{cases} \tag{4.7}$$

with

$$\lambda_i = \binom{n}{i} i!/n^i , \quad i=0(1)n , \tag{4.8}$$

$$\sigma_{00} = 1 ,$$

and

$$\sigma_{ij} = \begin{cases} 0 & , \quad i > j \\ \frac{1}{i!} \sum_{k=0}^{i} (-1)^{i-k} \binom{i}{k} k^j , & i \leqslant j = 1,2,\ldots \end{cases} \tag{4.9}$$

The numbers σ_{ij} are called Stirling numbers of the second kind. They are nonnegative (see (4.6)-(4.8)) and satisfy the following recurrence relation

$$\sigma_{i,j+1} = \sigma_{i-1,j} + i\,\sigma_{ij} . \tag{4.10}$$

The upper triangular matrices

$$A_N = (a_{ij}) , \quad 1 \leqslant i \leqslant j \leqslant N \leqslant n , \quad \text{and} \quad A_{N+1} = \begin{bmatrix} 1 & 0 \\ 0 & A_N \end{bmatrix} \tag{4.11}$$

given by (4.7) will play an important role in the sequel. It follows from (4.7), (4.8) and (4.10) that the a_{ij}'s may also be generated recursively:

$$a_{i,j+1} = \frac{i}{n} a_{ij} + \left(1 - \frac{i-1}{n}\right) a_{i-1,j} . \tag{4.12}$$

This, in turn, implies that the column sums of the matrices (4.11) are all equal to 1, which is also a trivial consequence of the fact that $B_n(x^j;1) \equiv 1$. A_N and A_{N+1} are said to be column-stochastic in the sense that their column entries are all nonnegative and sum to 1.

4.2. Variation diminishing properties.

Let $v(f)$ denote the number (finite or infinite) of sign changes of $f(x)$ as x traverses $[0,1]$, then, from (4.1),

$$v(B_n f) \leqslant v(\{f_k\}) \leqslant v(f) \tag{4.13}$$

$$v(B_n' f) \leqslant v(\{\Delta f_k\}) \leqslant v(f') \quad (\text{if } f \in C^1) \tag{4.14}$$

$$v(B_n'' f) \leqslant v(\{\Delta^2 f_k\}) \leqslant v(f'') \quad (\text{if } f \in C^2) \tag{4.15}$$

$$\vdots$$

describing, respectively, the so-called sign, monotonicity, and convexity variation diminishing properties of the Bernstein construction (see Pólya and Schoenberg (1958) and Schoenberg (1959)) - the gragh of $B_n f$ cannot have more zeros, maxima and minima, and points of inflexion than the corresponding numbers for the graph of f - and giving a good deal of information on the relative location and shape of the graphs of f and $B_n f$.

A more general description of the sign variation diminishing property (4.13), also contained in the references given above, is as follows.

let L denote any given straight line with equation $y = ax + b$, and $v_L(f)$ the number (finite or infinite) of intersections of f with L, i.e.,

$$v_L(f) = v(f(x) - ax-b) \; . \tag{4.16}$$

Then

$$v_L(B_n f) \leqslant v_L(f) \; , \tag{4.17}$$

Still another important smoothing effect of B_n is that, contrary to the commonly used interpolation and minimum norm approximation methods, B_n is a contraction operator on the space of continuous functions of bounded variation, as observed by Gordon and Riesenfeld (1974 a)), in the sense that

$$V(B_n f) \leqslant V(f) \ ,$$

where $V(f)$ denotes the total variation of f over $[0,1]$, the equality holding iff f is monotonic there (see Schoenberg (1959) and Karlin (1968)).

We end this Subsection with the well-known observation that $B_n f$ possesses all the nice shape preserving properties referred to above at the expense of having a notoriously slow rate of convergence (i.e. like $1/n$) and the following

<u>REMARK 4.2</u>. Being $B_{\underline{n}}(f;\underline{x})$, $f \in C\left[S_N\right]$, the product of N univariate Bernstein polynomials (see (3.2)), then the foregoing Schoenberg's results concerning the variation diminishing properties of B_n carry over into higher dimensions.

4.3. Polynomial interpolation at equidistant nodes.

Lagrange's ℓ_k and Bernstein's q_k basic polynomials for interpolation of a given f in C assuming the values f_k at the nodes k/n, k = O(1)n, are related by

$$B_n(\ell_k;x) = q_k(n,x) \ , \qquad k = O(1)n \ ,$$

from which it follows, on multiplying both sides by f_k and summing over k = O(1)n, that

$$B_n(L_n f;x) = B_n(f;x) \ .$$

It is a well-known fact that $L_n f = \sum_{k=0}^{n} f_k \ell_k$ does not converge uniformly to f for every f in C. However, the exceedingly good behaviour of $B_n f$ near the endpoints of [0,1] compensates the bad behaviour of $L_n f$ in such a way that

$$B_n(L_n f;x) \longrightarrow f(x) \qquad \text{as } n \longrightarrow \infty$$

uniformly in $0 \leq x \leq 1$ for all f in C. Thanks to these smoothing effects of B_n, every interpolating sequence at equidistant nodes can always be made uniformly convergent.

5. BERNSTEIN APPROXIMATION TO VECTOR-VALUED FUNCTIONS

5.1. Vector-valued Bernstein polynomials.

DEFINITION 5.1 (Gordon - Riesenfeld). The n^{th} degree vector-valued (parametric) Bernstein polynomial approximation to a given continuous vector-valued (parametric) function

$$F : [0,1] \longrightarrow \mathbb{R}_p \ , \quad F(s) = (X_1(s),\ldots,X_p(s))^T \ , \quad 0 \leqslant s \leqslant 1 \ , \quad p \geqslant 1 \tag{5.1}$$

is given, for $n \geqslant 0$, by

$$\mathcal{B}_0 F(s) = F(0)$$

$$\mathcal{B}_n F(s) = \sum_{k=0}^{n} F(\tfrac{k}{n})\, q_k(n,s) \ , \quad n \geqslant 1 \ , \text{ the } q_k\text{'s as in (1.2),} \tag{5.2}$$

$$= \sum_{k=0}^{n} \begin{bmatrix} X_1(\frac{k}{n}) \\ \vdots \\ X_p(\frac{k}{n}) \end{bmatrix} q_k(n,s) = \begin{bmatrix} B_n(X_1;s) \\ \vdots \\ B_n(X_p;x) \end{bmatrix} . \tag{5.3}$$

For $p=1$ $\mathcal{B}_n = B_n$. We take, therefore, $p>1$, the cases of principal practical interest being those of $p=2$ and $p=3$.

REMARK 5.1. If we take the forward shifting operators E and Δ to mean here that

$$EF_i = F_{i+1} \ , \quad \Delta F_i = F_{i+1} - F_i \ , \qquad i=0,1,2,\ldots,$$

then Lemma 1.1 is readily extended to cover $\mathcal{B}_n$:

$$\mathcal{B}_n(F;s) = ((1-s)I + sE)^n F_0 \ , \qquad F_k \equiv F(k/n) \ , \quad k=0(1)n \ , \tag{5.4}$$

$$\mathcal{B}_n(F;s) = (I + s\Delta)^n F_0 \tag{5.5}$$

$$= \sum_{k=0}^{n} \binom{n}{k} \Delta^k F_0 \, s^k ,$$

an n^{th} degree polynomial in the parameter s with point-valued coefficients.

Noting (5.4) and (5.5), $\mathcal{B}_n$ is, like B_n, the product of n linear averagings or n forward linear shiftings. Also, thanks to (5.3) and (2.3), $\mathcal{B}_n(f;s) \longrightarrow F(s)$ as $n \longrightarrow \infty$ uniformly in $0 \leqslant s \leqslant 1$.

5.2. Recursive generation and approximation properties of the Bernstein-Bézier operator.

DEFINITION 5.2 (Gordon - Riesenfeld). Let n+1 ordered points P_0, $P_1, \ldots, P_n$ in $\mathbb{R}_p$ be given and let $\mathbb{P} = \{P_k\}_{k=0}^n$ denote the (open) polygon formed by joining successive points. The Bézier curve associated with the n-sided Bézier polygon $\mathbb{P}$ is the parametric Bernstein polynomial

$$\mathcal{B}_n(P;s) = \sum_{k=0}^{n} P_k \, q_k(n,s) . \tag{5.6}$$

Here, the underlying vector-valued function is, of course, the polygonal function

$$F(s) = n\left[\left(\frac{k+1}{n} - s\right) P_k + \left(s - \frac{k}{n}\right) P_{k+1}\right], \quad \frac{k}{n} \leqslant s \leqslant \frac{k+1}{n}, \quad k=0(1)n-1 . \tag{5.7}$$

5.2.1. The convex hull property.

Subsection 1.1 is clearly applicable to $\mathcal{B}_n$; that is, the graph of $\mathcal{B}_n\, \mathbb{P}$ develops in the convex hull of the vertices of $\mathbb{P}$. In particular, the Bernstein-Bézier operator associates to a given point and a given line segment in $\mathbb{R}_p$ that point and that line segment themselves. It is also easily shown that the center of mass of the points P_k with masses q_k describes the graph of $\mathcal{B}_n(\mathbb{P};s)$ as s ranges from 0 to 1 (see Gordon and Riesenfeld (1974 a), b))).

5.2.2. Geometric construction of $\mathcal{B}_n\, \mathbb{P}$ and its derivatives.

Noting (5.5) and (5.7),

$$\mathcal{B}_n(\mathbb{P};s) = (I + s\Delta)^n\, P_0 = \mathbb{B}_s^n\, P_0 \,, \qquad (5.8)$$

and thus we can construct the Bézier curve (5.6) recursively. In correspondence with (1.7) and (1.8) we now have

$$\begin{array}{lllllll}
P_0 & & & & & & \\
 & P_0^1 & & & & & \\
P_1 & & P_0^2 & \cdots & & & \\
 & P_1^1 & \vdots & & P_0^{n-1} & & \\
P_2 & \vdots & \vdots & & & & P_0^n \equiv \mathcal{B}_n(\mathbb{P};s) \\
\vdots & \vdots & \vdots & & P_1^{n-1} & & \\
P_{n-2} & & \vdots & \cdots & & & \qquad (5.9) \\
 & P_{n-2}^1 & & & & & \\
P_{n-1} & & P_{n-2}^2 & & & & \\
 & P_{n-1}^1 & & & & & \\
P_n & & & & & &
\end{array}$$

where

$$P_i^j = \mathbb{B}_s\, P_i^{j-1}\,, \qquad j=1(1)n\ ;\quad i=0(1)n-j\,,$$
$$P_i^0 = P_i \qquad , \qquad i=0(1)n\ . \tag{5.10}$$

The points P_i^j in the j^{th} column of (5.9) are the vertices of a Bézier polygon $\mathbb{P}^j$ of order n-j. We arrive at the point on the Bézier curve (5.8) corresponding to the parameter value s by constructing successive Bézier polygons of lower and lower degree (cf. Bézier (1972), Gordon and Riesenfeld (1974 a))).

In correspondence with (4.1) and (4.2) we now have

$$\frac{(n-j)!}{n!}\mathcal{B}_n^{(j)}(\mathbb{P};s) = \sum_{k=0}^{n-j} \Delta^j P_k\, q_k(n-j,s) \tag{5.11}$$

$$= \begin{cases} \mathcal{B}_{n-j}(Q^j;s)\ , & j=0(1)n \\ 0 & ,\quad j>n \end{cases} \tag{5.12}$$

respectively, $Q^j = \left\{Q_0^j,\ Q_1^j,\ \dots\ ,\ Q_{n-j}^j\right\}$, j=0(1)n, being the (n-j)-sided Bézier polygon with vertices

$$Q_k^j = \Delta^j P_k\ , \qquad k=0(1)n-j\ ,$$
$$Q^0 = \mathbb{P}\ . \tag{5.13}$$

<u>REMARK 5.2</u>. Remark 4.1 is applicable to $\mathcal{B}_n$; that is, apart from the coefficient (n-j)!/n!, the j^{th} derivative of an n^{th} degree Bézier curve of Bézier polygon $\mathbb{P}$ is a Bézier curve of degree n-j whose Bézier polygon Q^j is derived from $\mathbb{P}$ according to (5.13). Consequently, a Bézier curve and all its derivatives can be calculated recursively and afford an easy geometric construction. Furthermore,

there is no need to construct a "triangle" similar to (5.9), based on Q^j. Indeed, since

$$\mathcal{B}_{n-j}(Q^j;s) = B_s^{n-j}(\Delta^j P_0) = \Delta^j(B_s^{n-j}P_0) \tag{5.14}$$
$$= \Delta^j P_0^{n-j} ,$$

differencing the entries in the $(n-j)^{th}$ column of (5.9) leads to $\mathcal{B}_n^{(j)}(\mathbb{P};s)$. In particular, there follows from (5.11) - (5.14), for j = 0 and s = 0,1:

$$\begin{aligned} \mathcal{B}_n(\mathbb{P};0) &= B_0^n\, P_0 = P_0 \\ \mathcal{B}_n(\mathbb{P};1) &= B_1^n\, P_0 = E^n P_0 = P_n , \end{aligned} \tag{5.15}$$

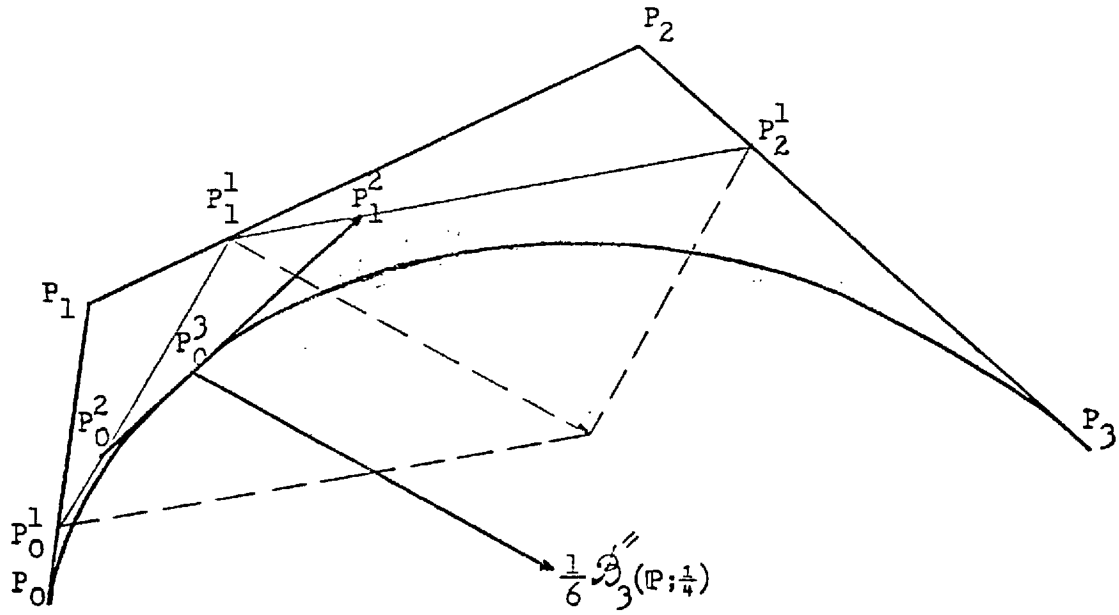
extending to $\mathcal{B}_n$ the interpolation properties (1.10) of B_n; and, for j = 1 and s = 0,1:

$$\begin{aligned} \frac{1}{n}\mathcal{B}_n'(\mathbb{P};0) &= \Delta P_0 = P_1 - P_0 \\ \frac{1}{n}\mathcal{B}_n'(\mathbb{P};1) &= \Delta(E^{n-1}P_0) = \Delta P_{n-1} = P_n - P_{n-1} . \end{aligned} \tag{5.16}$$

The relations (5.15) - (5.16) imply the tangency of the Bézier curve to the endsides at the endpoints of the corresponding Bézier polygon (cf. Gordon and Riesenfeld (1974 a))).

Figure 1 below illustrates the geometric construction of $\mathcal{B}_3^{(j)}(\mathbb{P};s)$, for j = 0,1,2, at the point s = 1/4.

Figure 1



$$\frac{1}{3}\mathcal{B}_3'(P;\tfrac{1}{4}) = \Delta P_0^2 = P_1^2 - P_0^2$$

$$\frac{1}{6}\mathcal{B}_3''(P;\tfrac{1}{4}) = \Delta^2 P_0^1 = (P_2^1 - P_1^1) - (P_1^1 - P_0^1)$$

5.2.3. Variation diminishing properties.

Noting (5.3) and (4.13) - (4.15), $\mathcal{B}_n$ can be said to be, like B_n, a variation diminishing operator in the sense that each component $B_n X_k$ of $\mathcal{B}_n F$ is at least as smooth as the corresponding component X_k of F, where smooth refers to the number of zeros, maxima and minima, points of inflexion, total variation, etc.

In analogy with (4.16) - (4.17) we may also describe the variation diminishing character of $\mathcal{B}_n$ with respect to a hyperplane H with equation

$$\sum_{i=1}^{p} h_i x_i = (h,x) = c \ , \tag{5.17}$$

where (.,.) denotes the inner product in $\mathbb{R}_p$. Defining

$$v_H(F) = v(\ (h,F) - c\ )\ , \tag{5.18}$$

the number (finite or infinite) of intersections of F with H, we have

THEOREM 5.1 (G.-Bonne - Sablonnière).

a) $v_H(\mathcal{B}_n F) \leqslant v_H(F)$

b) $v_H(\mathcal{B}_n' F) \leqslant v_H(F')$ for F in C^1, i.e., X_k in C^1, $k=0(1)p$.

PROOF. a)

$$v_H(\mathcal{B}_n F) = v(\ (h,\mathcal{B}_n F) - c\ )\ , \text{ by (5.18)},$$

$$= v(\ \sum_{i=1}^{p} h_i B_n X_i\ - c)\ , \text{ by (5.17) and (5.3)},$$

$$= v(B_n(\ \sum_{i=1}^{p} h_i X_i\ - c))\ , \text{ because } B_n \text{ is linear and preserves constants},$$

$$\leqslant v(\ (h,F) - c\ )\ , \text{ by (4.13) and (5.17)},$$

$$= v_H(F) \qquad , \text{ by (5.18)}.$$

Part b) follows in much the same way.

REMARK 5.3. Paralleling the analysis carried out above with F in C^2 and c = 0 in (5.17) we get

c) $v_H(\mathcal{B}_n'' F) \leqslant v_H(F'')$.

While a) and b) describe the sign and monotonicity variation diminishing properties of $\mathcal{B}_n$, c) does not mean, however, that $\mathcal{B}_n$ diminishes the convexity of F, since this depends in general on the vector product $F' \wedge F''$.

REMARK 5.4. If, instead of the polygon $\mathbb{P}$ with vertices at $F(k/n)$, $k=0(1)n$, we inscribe in the graph of F any other polygon $\mathcal{P}$ with vertices at $F(s_k)$, where $0 = s_0 < s_1 < \dots < s_n = 1$, then

$$v_H(\mathcal{B}_n \mathcal{P}) \leqslant v_H(\mathcal{P}),$$

implying that polygonal (piecewise linear) approximation to continuous parametric functions is a variation diminishing approximation method (cf. Marsden and Schoenberg (1966)). This suggests the application of Schoenberg (1967)'s variation diminishing splines to the approximation of continuous parametric curves in $\mathbb{R}_p$. In this connection, interesting results were given by Gordon and Riesenfeld (1974 b)) and, more recently, by Germain-Bonne and Sablonnière (1976, 1977).

5.3. Iterates of the Bernstein-Bézier operator.

Thanks to (5.3), the problem of iterating $\mathcal{B}_n$ may naturally be reduced to p problems involving ordinary Bernstein iterates:

$$\underbrace{\mathcal{B}_n(\mathcal{B}_n(\dots (\mathcal{B}_n}_{r} F(s)) \dots)) = \mathcal{B}_n^r F(s) = \begin{bmatrix} B_n^r X_1(s) \\ \cdot \\ \cdot \\ \cdot \\ B_n^r X_p(s) \end{bmatrix}.$$

We deal with them in Chapter 3.

5.4. Bernstein-Bézier methods for multivariate, vector-valued functions.

Instead of the univariate, vector-valued function (5.1) let us consider the mapping

$$F : \underline{x} \in S_N \subset \mathbb{R}_N \longrightarrow \mathbb{R}_p , \quad F(\underline{x}) = (X_1(\underline{x}), \ldots, X_p(\underline{x}))^T \in \mathbb{R}_p ,$$

with $\underline{x} = (x_1, \ldots, x_N)$, $S_N = \{ \underline{x} \in \mathbb{R}_N : 0 \leqslant x_i \leqslant 1, \ i=1(1)N \}$, and associate with F the Bernstein-Bézier operator

$$\mathcal{B}_{\underline{n}}(F;\underline{x}) = \sum_{i_1=0}^{n_1} \cdots \sum_{i_N=0}^{n_N} F_{\underline{i}} \, q_{i_1}(n_1,x_1) \ldots q_{i_N}(n_N,x_N) \qquad (5.19)$$

$$= (B_{\underline{n}}(X_1;\underline{x}), \ \ldots \ , \ B_{\underline{n}}(X_p;\underline{x}))^T \qquad (5.20)$$

with $\underline{n}$ denoting $n_1, n_2, \ldots, n_N$, $\underline{i}$ denoting $i_1, i_2, \ldots, i_N$, and $F_{\underline{i}} = F(i_1/n_1, \ \ldots, \ i_N/n_N)$.

Clearly, $\mathcal{B}_{\underline{n}}(F;\underline{x})$ is a vector-valued (parametric) Bernstein polynomial of degree n_i in x_i, $i=1(1)N$, and extends $\mathcal{B}_n$ to the N-dimensional setting. Incidentally, (5.19) gives, for $N = 2$ and $p = 3$, the basic formula in Bézier's free-form surface design technique.

Most of what has been said for $\mathcal{B}_n$ extends easily to $\mathcal{B}_{\underline{n}}$. Namely:

i) $\mathcal{B}_{\underline{n}}(F;\underline{x})$ represents a p-dimensional surface which develops within the convex hull of the points $F_{\underline{i}}$, the vertices of an $n_1 n_2 \ldots n_N$-faced net of line segments which plays here a role analogous to that of the Bézier polygon, and the center of mass of the points $F_{\underline{i}}$ with masses

$$M_{\underline{i}} = \prod_{k=1}^{N} q_{i_k}(n_k, x_k)$$

describes the graph of $\mathcal{B}_{\underline{n}}(F;\underline{x})$ as $\underline{x}$ runs over S_N.

ii) Since $B_{\underline{n}}$ is a smoothing operator, then, from (5.20), so is $\mathcal{B}_{\underline{n}}$.

Over the past several years, the methods of Bernstein-Bézier have attracted widespread attention, especially in connection with problems of computer-aided design and numerical control production of free-form curves and surfaces such as aeroplane fuselages, ship hulls, and automobile bodies. For detailed and intensely practical expositions we refer, e.g., to Bézier (1972), Barnhill and Riesenfeld (1974), and Forrest (1971, 1972).

CHAPTER 2

NUMERICAL CONDENSATION OF MULTIVARIATE POLYNOMIALS

Polynomial condensation, also referred to in the literature as telescoping or economization, is a numerical procedure aiming at reducing the computational effort required to evaluate a given polynomial P at a given point of its domain while allowing for a small oscillating error ε to be distributed over the domain where the condensed representation of P is sought. It was first conceived and applied to univariate problems of numerical mathematics by Lanczos (1938, 1952, 1956) and has recently been extended to the multivariate setting by Ortiz (1977).

1. CONDENSATION IN $\mathcal{P}_n$

Let $P = P(x) \in \mathcal{P}_n$, $n \geqslant 1$, and $x \in K_1$, a compact of $\mathbb{R}_1$. We assume, once and for all, that

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 , \quad a_n \neq 0 , \quad x \in K_1 , \tag{1.1}$$

is evaluated at $x \neq 0$ by Horner's nested multiplication, i.e., by means of the backward recurrence scheme

$$P = b_0 :$$

$$b_n = a_n$$

$$b_r = xb_{r+1} + a_r , \qquad r=n-1(-1)0 , \tag{1.2}$$

always requiring n multiplications regardless of the number of vanishing a_r's .

DEFINITION 1.1. We say that $P_\varepsilon = P_\varepsilon(x)$ is an ε-condensed polynomial representation of P in K_1 provided that

$$\text{(i)} \quad \partial P_\varepsilon < \partial P \qquad \text{and} \qquad \text{(ii)} \quad \| P - P_\varepsilon \| < \varepsilon ,$$

∂ standing for 'degree of' and $\|.\|$ for the uniform norm in K_1 .

1.1. A sufficient condition for polynomial condensation.

Sufficient conditions for the existence of $P_\varepsilon \in \mathcal{P}_{n-s-1}$, $0 \leqslant s < n$, in $K_1 = [0,1]$ may be obtained as follows.

We denote by $T_m^*(x)$ the m^{th} Chebyshev polynomial of the first kind shifted to $[0,1]$ and recall that

$$x^n = 2^{1-2n} \sum_{j=0}^{n}{}' \binom{2n}{n-j} T_j^*(x) , \tag{1.3}$$

where the prime indicates that the coefficient of T_0^* is to be halved. Using (1.3), P may be written in the form

$$P(x) = \sum_{j=0}^{n-1} a_j x^j + a_n x^n$$

$$= P^1(x) + 2^{1-2n} a_n T_n^*(x) \qquad (1.4)$$

with

$$P^1(x) = \sum_{j=0}^{n-1} a_j x^j + 2^{1-2n} a_n \sum_{j=0}^{n-1}{}' \binom{2n}{n-j} T_j^*(x) . \qquad (1.5)$$

It is clear from (1.4) that the condition

$$\varepsilon_0 = 2^{1-2n} |a_n| < \varepsilon \qquad (1.6)$$

implies that $P^1 \in \mathcal{P}_{n-1}$ and $\|P - P^1\| < \varepsilon$; that is, P^1 is an ε-condensed polynomial representation of P.

1.2. Implementation of the condensation process.

Assuming that (1.6) holds, recalling that

$$T_0^*(x) = 1$$

$$T_n^*(x) = (-1)^n + \sum_{k=1}^{n} (-1)^{n-k} \frac{n}{k} \binom{n+k-1}{n-k} 2^{2k-1} x^k , \qquad n \geqslant 1 , \quad (1.7)$$

and noting that

$$\sum_{k=0}^{n-1}{}' \binom{2n}{n-k} T_k^*(x) = (-1)^{n-1} + \sum_{j=1}^{n-1} 2^{2j-1} \sum_{k=j}^{n-1} (-1)^{k-j} \frac{k}{j} \binom{k+j-1}{k-j} \binom{2n}{n-k} x^j ,$$

we find

$$P^1(x) = \sum_{j=0}^{n-1} a_j^1 x^j \tag{1.8}$$

with

$$\begin{cases} a_0^1 = a_0 + (-1)^{n-1} 2^{1-2n} a_n \\ a_j^1 = a_j + 4^{j-n} a_n \sum_{k=j}^{n-1} (-1)^{k-j} \frac{k}{j} \binom{k+j-1}{k-j}\binom{2n}{n-k}, \quad j=1(1)n-1 . \end{cases} \tag{1.9}$$

We are now ready for the condensation of P^1, i.e., for the second step in the condensation of P. This will be possible provided that

$$\varepsilon_1 = 2^{1-2(n-1)} \left|a_{n-1}^1\right| < \varepsilon - \varepsilon_0 .$$

Assuming this true, we replace in (1.9) n with n-1 and a_j with a_j^1, j=0(1)n-2, to obtain a_j^2, j=0(1)n-2, the coefficient vector of a condensed polynomial representation of P with a new tolerance parameter ε equal to $\varepsilon_0 + \varepsilon_1$.

Setting $a_j^0 = a_j$, j=0(1)n, the algorithm

$$\begin{cases} a_0^{m+1} = a_0^m + (-1)^{n-m-1} 2^{1-2(n-m)} a_{n-m}^m \\ a_j^{m+1} = a_j^m + 4^{j-n+m} a_{n-m}^m \sum_{k=j}^{n-m-1} (-1)^{k-j} \frac{k}{j} \binom{k+j-1}{k-j}\binom{2(n-m)}{n-m-k}, \quad j=1(1)n-m-1, \end{cases} \tag{1.10}$$

is repeated for m=0(1)s, $0 \leq s < n$, as long as

$$\sum_{m=0}^{s} \varepsilon_m \quad , \quad \varepsilon_m = 2^{1-2(n-m)} \left| a_{n-m}^m \right| , \tag{1.11}$$

to give a_j^{s+1}, j=0(1)n-s-1, the coefficient vector of the condensed polynomial form of P.

REMARK 1.1. The number s+1 of condensation steps may be determined apriori if we carry out the basis transformation

$$P(x) = \sum_{r=0}^{n} a_{n-r}\, x^{n-r} = \sum_{m=0}^{n}{}' c_{n-m}\, T_m^*(x) \; ,$$

$$c_{n-m} = 2^{1-2n} \sum_{r=0}^{m} 4^r \binom{2(n-r)}{m-r} a_{n-r} \; , \qquad m=0(1)n \; ,$$

and observe that

$$\left| c_{n-m} \right| = \varepsilon_m \; , \qquad m=0(1)s \; .$$

Therefore, instead of (1.11), we can use

$$\sum_{r=0}^{s} 4^r \sum_{m=r}^{s} \binom{2(n-r)}{m-r} \left| a_{n-r} \right| < 2^{2n-1} \varepsilon \; , \qquad 0 \leqslant s < n \; , \qquad (1.12)$$

to predict the maximum degree reduction ε allows (see also Ortiz (1977)).

2. CONDENSATION OF MULTIVARIATE POLYNOMIALS

2.1. Bivariate polynomial evaluation schemes.

Let

$$P = P(x,y) = \sum_{i=0}^{I} \sum_{j=0}^{J} a_{ij}\, x^i\, y^j \in \mathcal{P}_{I,J}\ , \qquad a_{IJ} \neq 0, \qquad (x,y) \in K_2\ , \qquad (2.1)$$

where K_2 is a compact of $\mathbb{R}_2$. $A = (a_{ij})$, the $(J+1)x(I+1)$ coefficient matrix of P, is such that, for every i and j, a_{ij} is the coefficient of $x^i y^j$ in P.

The writng of P as

$$P = \sum_{j=0}^{J} p_j\, y^j\ , \qquad p_j = p_j(x) = \sum_{i=0}^{I} a_{ij}\, x^i\ , \qquad j=0(1)J\ , \qquad (2.2)$$

or

$$P = \sum_{i=0}^{I} p_i^*\, x^i\ , \qquad p_i^* = p_i^*(y) = \sum_{j=0}^{J} a_{ij}\, y^j\ , \qquad i=0(1)I\ , \qquad (2.3)$$

with p_j (p_i^*) uniquely associated with the j^{th} row (i^{th} column) of A, makes the evaluation and condensation problems for P(x,y) entirely analogous to those referred to in Section 1. Indeed, the evaluation of P at a point $(x,y) \in K_2$ with nonzero co-ordinates is reduced to J+2 or I+2 1-dimensional polynomial evaluation problems according to the representations (2.2) or (2.3) respectively, and carried out by means of two backward recurrence multiplication schemes similar to (1.2), these are:

$p_j = b_0 \quad (j=0(1)J)$:

$$b_I = a_{Ij}$$

$$b_r = xb_{r+1} + a_{rj} , \qquad r = I-1(-1)0 ,$$

$P = c_0$:

$$c_J = p_J$$

$$c_r = yc_{r+1} + p_r , \qquad r = J-1(-1)0 ;$$

or

$p_i^* = b_0 \quad (i=0(1)I)$:

$$b_J = a_{iJ}$$

$$b_r = yb_{r+1} + a_{ir} , \qquad r = J-1(-1)0 ,$$

$P = c_0$:

$$c_I = p_I^*$$

$$c_r = xc_{r+1} + p_r^* , \qquad r = I-1(-1)0 .$$

In either case, the number of multiplications required is always $(I+1)(J+1)-1$, no matter how sparse A may be. As for the ε-condensation problem for $P(x,y)$ in $K_2 = [0,1]x[0,1]$ we now have, in correspondence with (1.12), the following existence conditions

$$\sum_{r=0}^{s} 4^r \sum_{m=r}^{s} \binom{2(J-r)}{m-r} \left|p_{J-r}\right| < 2^{2J-1}\varepsilon , \qquad 0 \leqslant s < J , \tag{2.4}$$

or

$$\sum_{r=0}^{s} 4^r \sum_{m=r}^{s} \binom{2(I-r)}{m-r} \left|p_{I-r}^*\right| < 2^{2I-1}\varepsilon , \qquad 0 \leqslant s < I , \tag{2.5}$$

according as P is given by (2.2) or (2.3) respectively.

The above bidimensional evaluation and condensation problems afford an immediate extension to the multivariate setting. Given

$$P = P(x_1,x_2,\ldots,x_N) = \sum_{i_1=0}^{I_1} \sum_{i_2=0}^{I_2} \cdots \sum_{i_N=0}^{I_N} a_{i_1 i_2 \ldots i_N} x_1^{i_1} x_2^{i_2} \ldots x_N^{i_N} \qquad (2.6)$$

with $a_{I_1 I_2 \ldots I_N} \neq 0$ and $(x_1,x_2,\ldots,x_N) \in K_N$, a given compact of R_N, we write it in one of the following equivalent forms

$$P = \sum_{i_k=0}^{I_k} p_{i_k} x^{i_k} , \qquad k = 1(1)N , \qquad (2.7)$$

with

$$p_{i_k} = p_{i_k}(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_N)$$
$$= \sum_{i_1=0}^{I_1} \cdots \sum_{i_{k-1}=0}^{I_{k-1}} \sum_{i_{k+1}=0}^{I_{k+1}} \cdots \sum_{i_N=0}^{I_N} a_{i_1 i_2 \ldots i_N} x_1^{i_1} \ldots x_{k-1}^{i_{k-1}} x_{k+1}^{i_{k+1}} \ldots x_N^{i_N} .$$

To evaluate P at a point $(x_1,x_2,\ldots,x_N)$ in K_N with nonzero co-ordinates we use N multiplication schemes similar to (1.2) and perform $(I_1+1)(I_2+1) \ldots (I_N+1)-1$ multiplications; ε-condensation of P in the unit hypercube of R_N with respect to the variable x_k is possible if

$$\sum_{r=0}^{s_k} 4^r \sum_{m=r}^{s_k} \binom{2(I_k-r)}{m-r} \left| p_{I_k-r} \right| < 2^{2I_k-1} \varepsilon , \qquad k = 1(1)N . \qquad (2.8)$$

2.2. Smoothness indicators.

The numbers $s_k = s_k(x_k, \mathcal{E})$ in (2.8) may serve as indicators of the following attributes of $P(x_1, x_2, \ldots, x_N)$, obviously related to each other to a certain extent:

(i) Smoothness - the larger s_k the smoother P in the x_k-direction;

(ii) Weight of x_k - the larger s_k the smaller the importance of the variable x_k in the representation of P.

This leads naturally to the definition of principal variables as those for which the s's are least and, then, to the following problem.

2.3. Approximation of a multivariate polynomial by another polynomial of fewer variables.

Clearly, the x_k-dependence can only be removed from the representation (2.6) of $P(x_1, x_2, \ldots, x_N)$ if (2.8) holds for $s_k = I_k - 1$, yielding the condition

$$\sum_{r=0}^{I_k-1} \left[1 - 4^{-(I_k-r)} \binom{2(I_k-r)}{I_k-r} \right] \left| p_{I_k-r} \right| < \mathcal{E}, \qquad 1 \leqslant k \leqslant N . \qquad (2.9)$$

Assume (2.9) holds for k = 1, say, and the left hand side of (2.9) equals $\mathcal{E}_1 < \mathcal{E}$. Then $P \in \mathcal{P}_{I_1, I_2, \ldots, I_N}$ may be replaced by $P_1 \in \mathcal{P}_{I_2, I_3, \ldots, I_N}$, which is such that $\| P - P_1 \| < \mathcal{E}_1$. We may, of course, submit P_1 to the same reduction process. Let $\mathcal{E}_2 < \mathcal{E}$ be the tolerance parameter within which the variable x_2, say, may be eliminated from $P_1 \in \mathcal{P}_{I_2,I_3,\ldots,I_N}$, yielding $P_2 \in \mathcal{P}_{I_3,I_4,\ldots,I_N}$ such that $\|P_1 - P_2\| < \varepsilon_2$. Clearly, P_2 may replace P provided $\varepsilon_1 + \varepsilon_2 < \varepsilon$. If we can do this k times, then we end up with $P_k \in \mathcal{P}_{I_{k+1},I_{k+2},\ldots,I_N}$, an ε-condensed form of P with k fewer variables than P itself.

2.4. An algorithm for the condensation of multivariate polynomials.

In this Subsection we extend algorithm (1.10) to the multivariate case. The emphasis will actually be on bivariate polynomials as there is no essential difficulty in extending what follows to higher dimensions.

Taking the polynomial (2.1) and paralleling the analysis carried out in Subsections 1.1 and 1.2, we have, from (2.2),

$$P = \sum_{j=0}^{J-1} p_j\, y^j + p_J\, y^J$$

$$= P^1 + 2^{1-2J}\, p_J\, T_J^*(y) \tag{2.10}$$

where

$$P^1 = \sum_{j=0}^{J-1} p_j\, y^j + 2^{1-2J}\, p_J \sum_{j=0}^{J-1}{}' \binom{2J}{J-j} T_j^*(y) \tag{2.11}$$

$$= \sum_{j=0}^{J-1} p_j^1\, y^j ,$$

with

$$\begin{cases} p_0^1 = p_0 + (-1)^{J-1}\, 2^{1-2J}\, p_J \\ p_j^1 = p_j + 4^{j-J} p_J \sum_{k=j}^{J-1} (-1)^{k-j} \frac{k}{j} \binom{k+j-1}{k-j}\binom{2J}{J-k} , \qquad j = 1(1)J-1 , \end{cases} \tag{2.12}$$

is an ε-condensed form of P provided that

$$\varepsilon_0 = 2^{1-2J} \left| p_J \right| < \varepsilon .$$

It should be noted that the elimination of the second term on the right hand side of (2.10) amounts to deleting the J^{th} row of the coefficient matrix A of P and perturbing each of the other rows of A with the second term on the right hand side of (2.11).

Another condensation step will be possible if

$$\varepsilon_1 = 2^{1-2(J-1)} \left| p_{J-1}^1 \right| < \varepsilon - \varepsilon_0 .$$

This being so, then we change in (2.12) J into J-1, p_j into p_j^1 , $j = 0(1)J-2$, and start anew. The maximum number s_y of condensation steps is such that

$$\sum_{m=0}^{s_y} \varepsilon_m < \varepsilon \ , \quad \varepsilon_m = 2^{1-2(J-m)} \left| p_{J-m}^m \right| ,$$

(cf. (1.11)) and may be predicted by the use of (2.4). We end up with P^{s_y+1} in $\mathcal{P}_{I,J-s_y-1}$, $0 \leqslant s_y < J$:

$$P^{s_y+1} = \sum_{j=0}^{J-s_y-1} p_j^{s_y+1}\, y^j ,$$

with

$$\begin{cases} p_0^{m+1} = p_0^m + (-1)^{J-m-1}\, 2^{1-2(J-m)}\, p_{J-m}^m & \\ & m = 0(1)s_y \\ p_j^{m+1} = p_j^m + 4^{j-J+m}\, p_{J-m}^m \sum_{k=j}^{J-m-1} (-1)^{k-j} \frac{k}{j} \binom{k+j-1}{k-j}\binom{2(J-m)}{J-m-k}\,, & j = 1(1)J-m-1\,, \end{cases} \quad (2.13)$$

where $p_j^0 = p_j$, $j = 0(1)J$ (cf. (1.10)).

The condensation of $P(x,y)$ has been carried out with respect to the variable y (rows of A). Obviously, it could have been performed, in exactly the same way, with respect to the variable x (columns of A).

3. NUMERICAL CONDENSATION OF BERNSTEIN POLYNOMIALS

Bernstein approximants are applied in those numerical approximation problems referred to at the close of Chapter 1, where shape preservation is more important than closeness of fit. Being slowly convergent, fairly high degree approximants are required. However, a considerable reduction of degree may be achieved through condensation of the Bernstein approximants to a given f in C, under fairly weak smoothness conditions on f, keeping their shape approximating properties only slightly changed (see also Ortiz and M. da Silva (1978)).

3.1. Condensation of $B_n(f;x)$.

REMARK 3.1. The Bernstein operator B_n maps the whole class of functions f taking on the values f_k at the nodes k/n, $k = 0(1)n$, into one and the same polynomial $B_n f$. In particular, if $P \in \mathcal{P}_n$ is the polynomial which interpolates the table $\{k/n, f_k\}_{k=0}^{n}$, we have

$$B_n(P;x) \equiv B_n(f;x) = \sum_{k=0}^{n} \binom{n}{k} \Delta^k f_0 \, x^k , \tag{3.1}$$

and thus conditions under which $B_n f$ may be condensed can always be stated in terms of P.

REMARK 3.2. Recalling that

$$B_n(t^n;x) = \lambda_n x^n + \dots \quad = c_n T_n^*(x) + \dots$$

with

$$\lambda_n = \prod_{k=1}^{n-1} (1-\frac{k}{n}) = \prod_{k=1}^{[n/2]}{}' \left[\frac{k}{n}(1-\frac{k}{n})\right] \leqslant 2^{1-n} ,$$

the prime indicating that the product is to be doubled when n is even (see Subsection 4.1 of Chapter 1), and

$$c_n = 2^{1-2n} \lambda_n \leqslant 2^{2-3n} ,$$

we see that the Bernstein approximants are particularly suitable for numerical condensation.

Let us then assume that, in the process of approximating to the shape of a given curve, there is given a tolerance parameter ε related to the accuracy to which variations in shape cease to be detectable or relevant for the problem in hand. Let

$$P = P(x) = \sum_{j=0}^{n} b_j x^j$$

be such that $P(k/n) = f_k$, $k = 0(1)n$, then

$$B_n(P;x) = \sum_{j=0}^{n} b_j B_n(x^j;x)$$

$$= \sum_{i=0}^{n} \left(\sum_{j=i}^{n} a_{ij} b_j \right) x^i \qquad (3.2)$$

with

$$a_{ij} = \lambda_i \, n^{i-j} \, \sigma_{ij} \qquad (3.3)$$

given by (4.7) - (4.9) of Chapter 1. Equating coefficients of like powers of x in (3.1) and (3.2),

$$\binom{n}{i} \Delta^i f_0 = \sum_{j=i}^{n} a_{ij} b_j \, , \qquad i = 0(1)n \, , \qquad (3.4)$$

giving, for $i = n$,

$$\Delta^n f_0 = \lambda_n b_n = \frac{n!}{n^n} b_n \, ,$$

and leading to the following sufficient conditions on the smoothness of f for the ε-condensation of $B_n f$ from $\mathcal{P}_n$ to $\mathcal{P}_{n-1}$ (see (1.6)):

(i) In terms of the leading coefficient of $B_n f$ (see (3.1)):

$$\left| \Delta^n f_0 \right| < 2^{2n-1} \varepsilon \, .$$

(ii) In terms of the leading coefficient of P:

$$\left| b_n \right| < \frac{2^{2n-1} \, n^n}{n!} \varepsilon \, .$$

From (3.4),

$$\binom{n}{i} \left| \Delta^i f_0 \right| < \sum_{j=i}^{n} a_{ij} \left| b_j \right| \leqslant M \left\| A \right\|_\infty \, , \qquad (3.5)$$

where

$$M = \max_{0 \leqslant j \leqslant n} |b_j|$$

and

$$\|A\|_\infty = \max_{0 \leqslant i \leqslant n} r_i , \qquad r_i = \sum_{j=i}^{n} a_{ij} , \tag{3.6}$$

the matrix $A = (a_{ij})$ being, as seen in Subsection 4.1 of Chapter 1, upper triangular and column-stochastic. Therefore, an s-step condensation process to reduce $B_n f \in \mathcal{P}_n$ to $(B_n f)_\varepsilon \in \mathcal{P}_{n-s-1}$, $0 \leqslant s < n$, requires that (see (1.12))

$$\sum_{i=0}^{s} 4^i \sum_{j=0}^{s-i} \binom{2(n-i)}{j} < \frac{2^{2n-1}}{M \|A\|_\infty} < \frac{2^{2(n-1)}}{M} \varepsilon$$

be satisfied. Assuming this true, then algorithm (1.10) enables a step by step computation of $(B_n f)_\varepsilon$.

It should be noted that, in asserting that $\|A\|_\infty < 2$, use was made of the facts that $r_0 = a_{00} = 1$ and that the leading terms in the row sums r_i in (3.6) are $a_{ii} = \lambda_i$ and $a_{i,i+1} = \lambda_i \binom{i+1}{2}/n$, both >0 and <1, hence $r_i < 2$, $i = 0(1)n$.

3.1.1. Numerical example.

By way of illustration we consider the shape approximation problem of the polygonal function f with vertices (0,0), (.2,.6), (.6,.8), (.9,.7), and (1,0) by a single polynomial of a fairly low degree.

If we consider the approximant $B_{10}f$ and take $\varepsilon = 3.5 \times 10^{-3}$, $\varepsilon' = 7.0 \times 10^{-3}$, and $\varepsilon'' = 2.5 \times 10^{-2}$ as admissible condensation parameters, then it is possible to condense $B_{10}f$ to polynomial representations of degree 6, 5, and 4 respectively without exceeding those error bounds.

In each of the figures 2, 3, and 4 below we show the graph of f, $B_{10}f$, the condensed representation P_ε of $B_{10}f$ to degree r, and B_rf for r = 4, 5, and 6 respectively. We can appreciate in Figure 3 that for r = 5 the adjustment between $B_{10}f$ and its condensed form is fairly close.

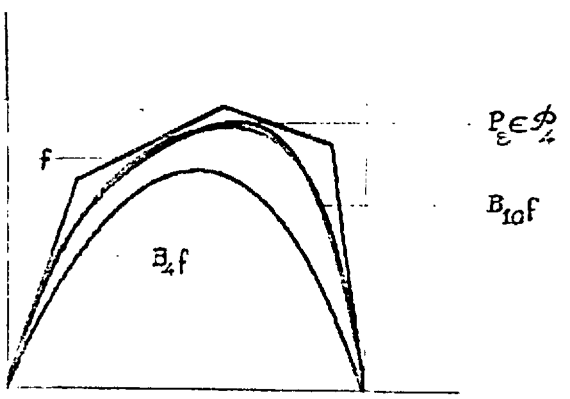


Figure 2

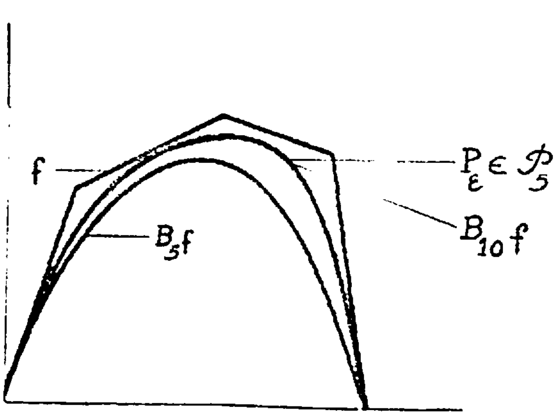


Figure 3

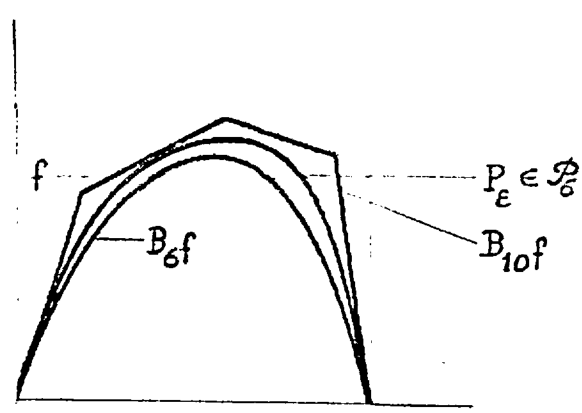


Figure 4

3.1.2. Numerical condensation of Bézier polynomials.

As seen in Section 5 of Chapter 1, the Bézier operator $\mathcal{B}_n$ associates with a given parametric function

$$F : [0,1] \longrightarrow \mathbb{R}_p , \qquad F(s) = (X_1(s), X_2(s), \ldots, X_p(s))^T , \qquad 0 \leqslant s \leqslant 1 , \qquad p \geqslant 1 ,$$

the parametric curve in $\mathbb{R}_p$

$$\mathcal{B}_n(F;s) = (B_n(X_1;s),\ B_n(X_2;s),\ \ldots\ ,\ B_n(X_p;s))^T .$$

The process of numerical condensation we have discussed for B_n extends trivially to $\mathcal{B}_n$. We say that

$$\mathcal{B}_{r_\varepsilon} F = (B_{r_\varepsilon} X_1,\ B_{r_\varepsilon} X_2,\ \ldots\ ,\ B_{r_\varepsilon} X_p)^T$$

is a condensed representation of $\mathcal{B}_n F$ provided that

$$\partial B_{r_\varepsilon} X_k < n , \qquad k = 1(1)p ,$$

and

$$\left\| \mathcal{B}_n F - \mathcal{B}_{r_\varepsilon} F \right\|_{\mathbb{R}_p} = \max_{1 \leqslant k \leqslant p} \left\| B_n X_k - B_{r_\varepsilon} X_k \right\| < \varepsilon .$$

3.2. Condensation of multivariate Bernstein-Bézier approximants.

Let $f = f(x,y) \in C[S_2]$, S_2 the unit (x,y)-square of $\mathbb{R}_2$. Referring to Section 3 of Chapter 1,

$$B_{n,m}(f;x,y) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_{ij}\, x^i\, y^j$$

with

$$a_{ij} = \binom{n}{i}\binom{m}{j} \Delta^{i}_{(x,1/n)} \Delta^{j}_{(y,1/m)} f(0,0) .$$

As an application of the material developed in Section 2, we find the following smoothness conditions on f for projecting $B_{n,m}f$ onto a proper subspace of $\mathcal{P}_{n,m}$ without introducing an error greater than ε in the numerical values assumed by $B_{n,m}f$ over S_2:

$$\text{(i)} \qquad \sum_{r=0}^{s_y} 4^r \sum_{k=0}^{s_y-r} \binom{2(m-r)}{k} \left| p_{m-r} \right| < 2^{2m-1}\varepsilon, \qquad 0 \leqslant s_y < m ,$$

where

$$p_{m-r} = p_{m-r}(x) = \sum_{i=0}^{n} a_{i,m-r}\, x^i ,$$

if the condensation of $B_{n,m}f$ is to be carried out with respect to the variable y; or

$$\text{(ii)} \qquad \sum_{r=0}^{s_x} 4^r \sum_{k=0}^{s_x-r} \binom{2(n-r)}{k} \left| p^*_{n-r} \right| < 2^{2n-1}\varepsilon, \qquad 0 \leqslant s_x < n ,$$

where

$$p^*_{n-r} = p^*_{n-r}(y) = \sum_{j=0}^{m} a_{n-r,j}\, y^j ,$$

if $B_{n,m}f$ is to be projected onto $\mathcal{P}_{n-s_x-1,m}$.

Use may be made of algorithm (2.13) for step by step computation of condensed forms $(B_{n,m}f)_\varepsilon$ of $B_{n,m}f$.

Figures 5, 7, 6, and 8 below represent $B_{10,10}f - (B_{10,10}f)_\varepsilon$, $B_{15,15}f - (B_{15,15}f)_\varepsilon$, and the corresponding contour maps respectively, with $\varepsilon = .01$ and $f = f(x,y) = \sin(x).\cos(y).\exp(-x^2-y^2)$.

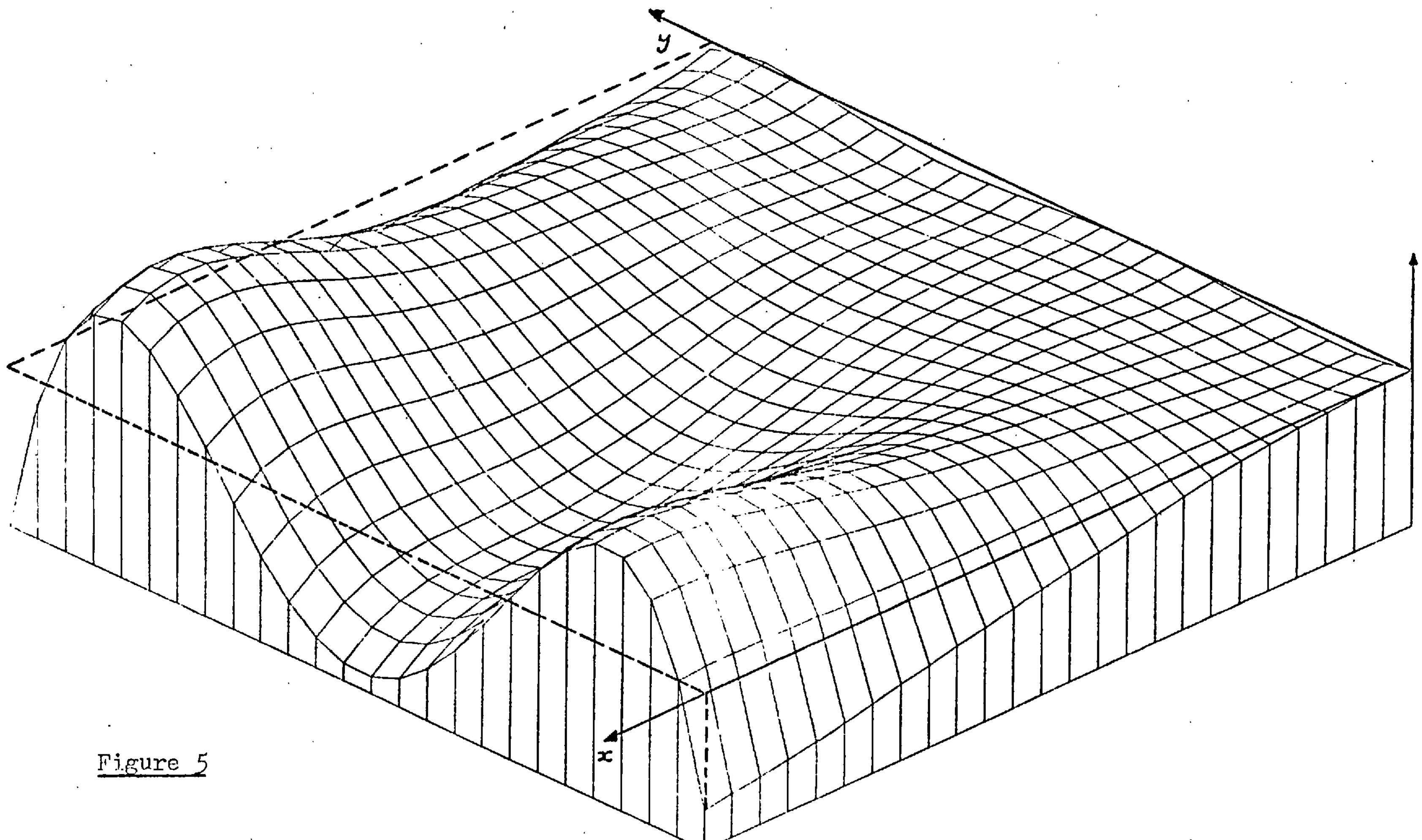


Figure 5

$z = B_{10,10}f - (B_{10,10}f)_{\varepsilon}$, $\varepsilon = .01$, $z_{max} = .0019$

$f = \sin(x).\cos(y).\exp(-x^2-y^2)$ $z_{min} = -.0023$

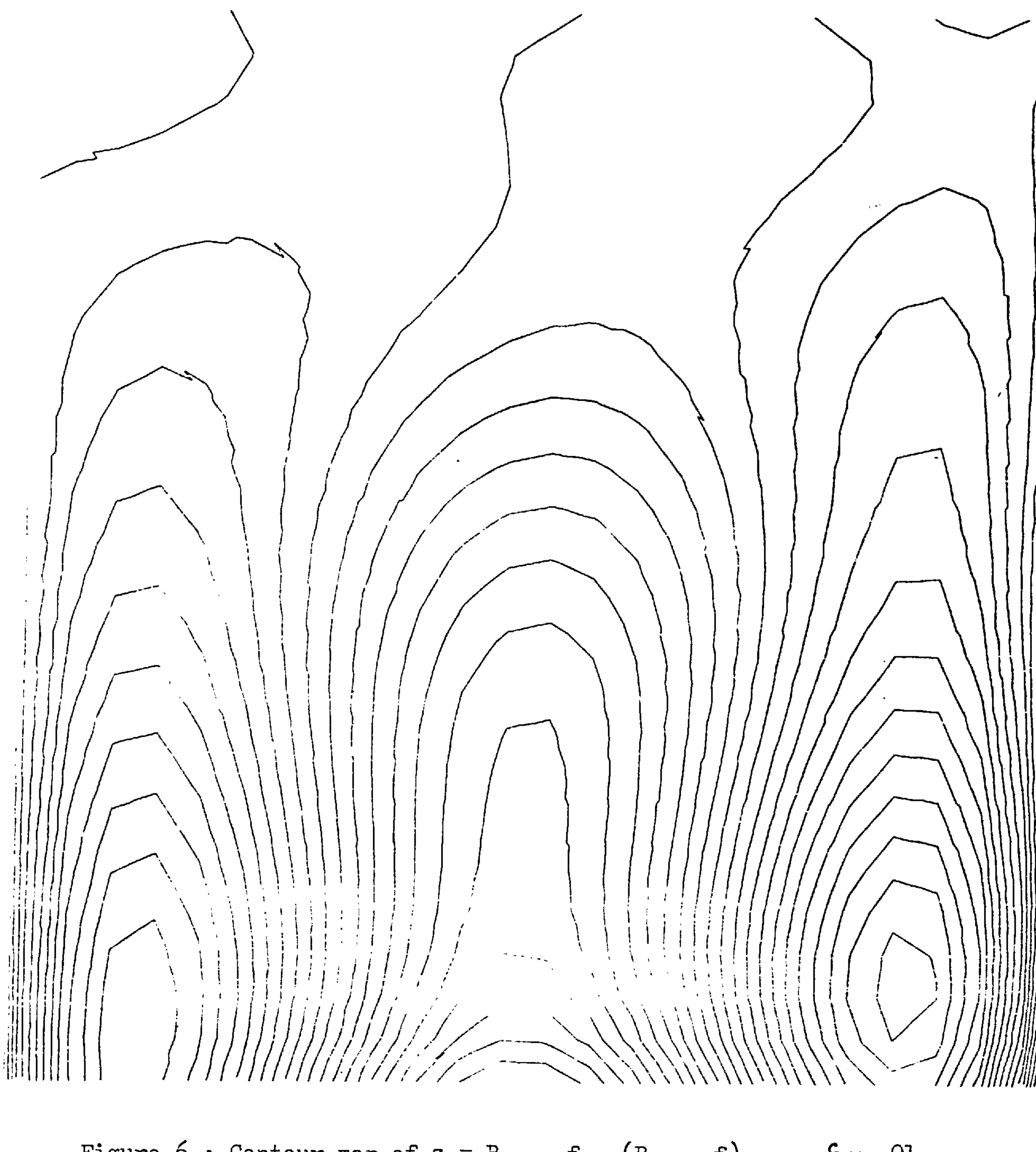


Figure 6 : Contour map of $z = B_{10,10}f - (B_{10,10}f)_{\varepsilon}$, $\varepsilon = .01$

$f = \sin(x).\cos(y).\exp(-x^2-y^2)$

$z_{max} = .0019$, $z_{min} = -.0023$

contour step = $(z_{max}-z_{min})/25$

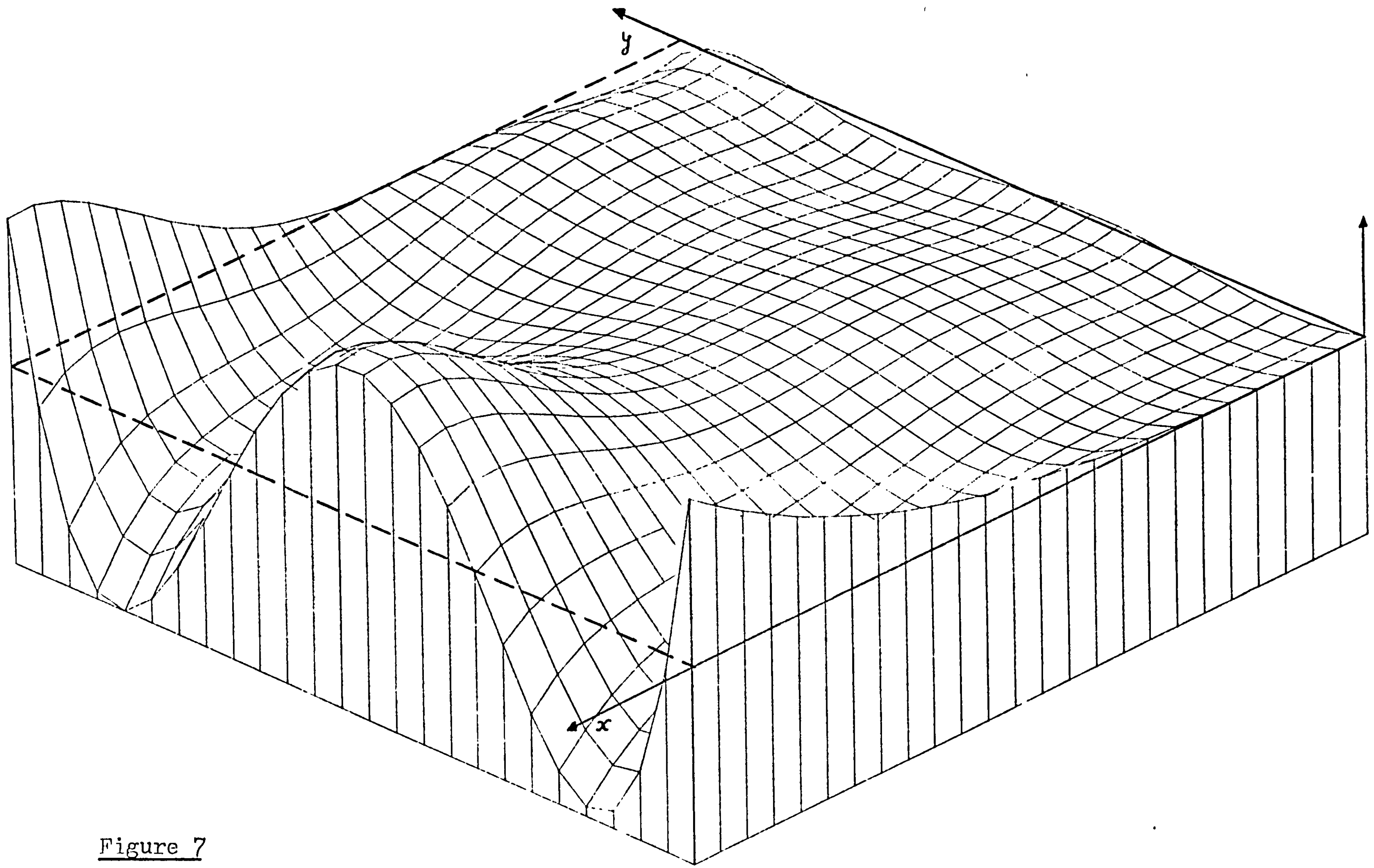


Figure 7

$z = B_{15,15}f - (B_{15,15}f)_{\varepsilon}$, $\varepsilon = .01$,

$f = \sin(x).\cos(y).\exp(-x^2-y^2)$

$z_{max} = .0029$

$z_{min} = -.0032$

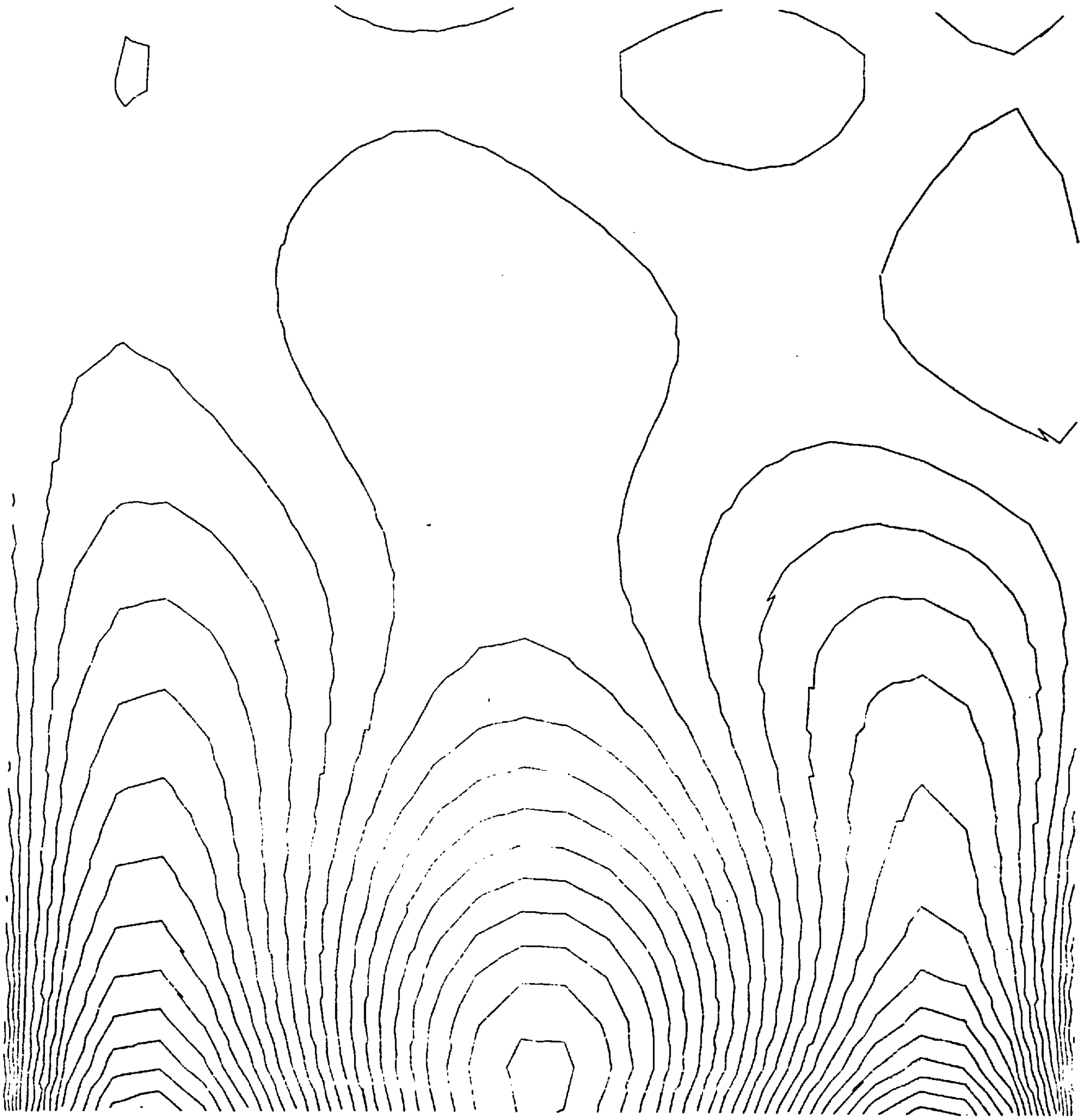


Figure 8 : Contour map of $z = B_{15,15}f - (B_{15,15}f)_{\varepsilon}$, $\varepsilon = .01$,

$f = \sin(x).\cos(y).\exp(-x^2-y^2)$

$z_{max} = .0029$, $z_{min} = -.0032$,

contour step = $(z_{max} - z_{min})/25$

Bearing in mind what has been said in Section 2 on the multivariate polynomial condensation problem, it is a simple matter to extend what has been done for $B_{n,m}$ to multivariate Bernstein polynomials $B_{\underline{n}}(f;\underline{x})$, $f \in C[S_N]$, S_N the unit hypercube of $\mathbb{R}_N$, and to Bézier hypersurfaces in $\mathbb{R}_p$:

$$\mathcal{B}_{\underline{n}}(F;\underline{x}) = (\ B_{\underline{n}}(X_1;\underline{x}),\ B_{\underline{n}}(X_2;\underline{x}),\ \ldots\ ,\ B_{\underline{n}}(X_p;\underline{x})\ )^T \ ,$$

where

$$F : S_N \longrightarrow \mathbb{R}_p \ , \qquad F(\underline{x}) = (\ X_1(\underline{x}),\ X_2(\underline{x}),\ \ldots\ ,\ X_p(\underline{x})\ )^T$$

(see Sections 3 and 5.4 of Chapter 1) .

CHAPTER 3

ARBITRARY ITERATES OF BERNSTEIN POLYNOMIALS

1. THE MATRIX FORM OF THE BERNSTEIN ITERATES

As seen before, the n^{th} degree Bernstein polynomial approximation to a given real $f(x)$ defined on $[0,1]$ is given by

$$B_n(f;x) = \sum_{k=0}^{n} f_k\, q_k(n,x) \ , \qquad f_k = f(k/n) \ , \tag{1.1}$$

where

$$q_k = q_k(n,x) = \binom{n}{k} x^k (1-x)^{n-k} = \sum_{i=k}^{n} c_{ik}\, x^i \ , \qquad k=0(1)n \ , \tag{1.2}$$

with

$$c_{ik} = \begin{cases} 0 & , \quad i<k \\ (-1)^{i-k} \binom{n}{i} \binom{i}{k} & , \quad 0 \leqslant k \leqslant i \leqslant n \ . \end{cases} \tag{1.3}$$

Bernstein iterates of natural order are defined recursively:

$$B_n^r(f;x) = B_n(B_n^{r-1}f;x) \ , \quad r>1 \ .$$

Owing to the facts that if $f \in \mathcal{P}_m$ then $B_n f \in \mathcal{P}_{\min\{m,n\}}$ and that B_n replaces $f \in C$ with a polynomial, it is no restriction, for $r>1$, to take f in $\mathcal{P}_N$, $N \leqslant n$. Actually, since $B_n 1 \equiv 1$, for all n in $\mathbb{N}$, we may take for domain (and range) of B_n the linear subspace $\mathcal{P}_N = \{x\ \mathcal{P}_{N-1}\}$ of polynomials of degree $\leqslant N$ vanishing at 0.

It follows from (1.1) - (1.3) that

$$B_n(x^j;x) = \sum_{k=1}^{n} \left(\frac{k}{n}\right)^j \sum_{i=k}^{n} (-1)^{i-k}\binom{n}{i}\binom{i}{k} x^i$$

$$= n^{-j} \sum_{i=1}^{n} \binom{n}{i}\left\{\sum_{k=1}^{i} (-1)^{i-k}\binom{i}{k} k^j\right\}x^i$$

$$= \sum_{i=1}^{j} a_{ij} x^i , \quad 1 \leqslant j \leqslant N , \tag{1.4}$$

where (cf. Subsection 4.1 of Chapter 1)

$$a_{ij} = \begin{cases} 0 , & i > j \\ \lambda_i n^{i-j} \sigma_{ij} , & 1 \leqslant i \leqslant j \leqslant N \end{cases} \tag{1.5}$$

with

$$\begin{aligned} \lambda_i = \lambda_i(n) &= \binom{n}{i} i!/n^i , \quad 1 \leqslant i \leqslant N, \\ &= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{i-1}{n}\right) , \end{aligned} \tag{1.6}$$

and

$$\sigma_{ij} = \begin{cases} 0 , & i > j \\ \frac{1}{i!} \sum_{k=1}^{i} (-1)^{i-k}\binom{i}{k}k^j , & 1 \leqslant i \leqslant j . \end{cases} \tag{1.7}$$

For a given

$$P(x) = \sum_{j=1}^{N} p_j x^j = X^T p$$

where $X^T = (x, x^2, \ldots, x^N)$ and $p = (p_1, p_2, \ldots, p_N)^T$, we have

$$B_n(P;x) = \sum_{j=1}^{N} p_j \, B_n(x^j;x)$$

$$= \sum_{j=1}^{N} p_j \sum_{i=1}^{j} a_{ij} \, x^i \quad , \text{ by } (1.4) \; ,$$

$$= \sum_{i=1}^{N} \Big(\sum_{j=i}^{N} a_{ij} \, p_j \Big) \, x^i$$

$$= X^T A_N \, p \; ,$$

with $A_N = (a_{ij})$ given by (1.5) - (1.7). Therefore,

$$B_n^r(P;x) = X^T A_N^r \, p \quad , \qquad r = 1,2,\ldots \; . \tag{1.8}$$

It is clear that (1.8) continues to hold for polynomials from $\dot{\mathcal{P}}_N$, namely, $p_0 + X^T p$. We have only to replace X by $(1,X)$, p by (p_0,p), and A_N by
$$A_{N+1} = \begin{bmatrix} 1 & 0 \\ 0 & A_N \end{bmatrix} .$$

Noting that all the eigenvalues of A_N lie in $(0,1]$, then equation (1.8) defines the <u>iterates of arbitrary order</u> $r \in \mathbb{R}$ of the Bernstein operator acting on $\mathcal{P}_N$.

2. THE TOTAL POSITIVITY OF B_n

Let $G = (g_{ij})$ be an n^{th} order real matrix. The k^{th} order minors of G formed from rows $i_1 < i_2 < \cdots < i_k$ and columns $j_1 < j_2 < \cdots < j_k$, $1 \leqslant k \leqslant N$, will be denoted by

$$G\begin{pmatrix} i_1, i_2, \ldots, i_k \\ j_1, j_2, \ldots, j_k \end{pmatrix} = \det\left(g_{i_m j_m}\right)_{m=1}^{k} = \begin{vmatrix} g_{i_1 j_1} & \cdots & g_{i_1 j_k} \\ \vdots & & \vdots \\ g_{i_k j_1} & \cdots & g_{i_k j_k} \end{vmatrix} . \qquad (2.1)$$

We say that $G \in TP$ (or is TP - totally positive) or that $G \in STP$ (or is STP - strictly totally positive) if all minors of G are nonnegative or strictly positive respectively.

If G is a lower (upper) triangular matrix, the minors (2.1) for which $i_m \geqslant j_m$ ($i_m \leqslant j_m$) for $1 \leqslant m \leqslant k$ are called the nontrivial minors of G. The remaining minors of G, the trivial minors, are obviously equal to 0. We say that $G \in \Delta TP$ or $G \in \Delta STP$ when the nontrivial minors of G are all nonnegative or strictly positive respectively.

The minors (2.1) for which $i_m = j_m$, $1 \leqslant m \leqslant k$, are termed the principal minors of G.

Being B_n a bijection in $\mathcal{P}_n$, for any given

$$P(x) = \sum_{j=0}^{n} p_j x^j = X^T p$$

$$= \sum_{k=0}^{n} \left(\sum_{j=0}^{k} \frac{\binom{k}{j}}{\binom{n}{j}} p_j \right) q_k(n,x) = Q^T C^{-1} p ,$$

where $Q^T = (q_0, q_1, \ldots, q_n)$ and $C = (c_{ik})$ is given by (1.3), we may write

$$P(x) = B_n(B_n^{-1}P;x) = \sum_{k=0}^{n} B_n^{-1}(P;k/n)\, q_k(n,x)$$

to obtain

$$(C^{-1}p)_k = \sum_{j=0}^{k} \frac{\binom{k}{j}}{\binom{n}{j}} p_j = B_n^{-1}(P;k/n) , \quad k=0(1)n .$$

REMARK 2.1. If $\underset{(a,b)}{Z}(f)$ denotes the number of real zeros of $f(x)$ in the indicated range, then, since:

$$v(P) \leqslant \underset{(0,1)}{Z}(P) = \underset{(0,1)}{Z}(Q^T C^{-1} p) = \underset{(0,\infty)}{Z}\Big(\sum_{k=0}^{n} \binom{n}{k} (C^{-1}p)_k z^k\Big), \; z = x/(1-x)$$

$$\leqslant v(C^{-1}p) = v(\{B_n^{-1}(P;k/n)\}_{k=0}^{n})$$

$$\leqslant v(B_n^{-1}P) \; ;$$

this being valid for all P in $\mathcal{P}_n$, we may replace in these inequalities P by B_nP (p by $A_{N+1}p$) to receive

$$v(B_nP) \leqslant \underset{(0,1)}{Z}(B_nP) = \underset{(0,1)}{Z}(Q^T C^{-1} A_{N+1} p)$$

$$\leqslant v(C^{-1}A_{N+1}p) = v(\{P(k/n)\}_{k=0}^{n})$$

$$\leqslant v(P) \; ;$$

and we get the well-known Schoenberg's result that the transformation $P \longrightarrow B_nP$ ($p \longrightarrow A_{N+1}p$) is variation diminishing while its inverse, $P \longrightarrow B_n^{-1}P$ ($p \longrightarrow A_{N+1}^{-1}p$) is variation increasing.

LEMMA 2.1.

$$A_N \in \Delta STP .$$

PROOF. The result follows immediately from Motzkin's theorem on variation diminishing matrix transformations (see, e.g., Schoenberg and Whitney (1951)) and the observation that the principal minors of A_N are all positive.

COROLLARY.

$$S = (\sigma_{ij}) \in \Delta STP .$$

PROOF. Using (1.5) and the homogeneity property of the determinant, we have

$$A_N\begin{pmatrix} i_1, i_2, \ldots, i_k \\ j_1, j_2, \ldots, j_k \end{pmatrix} = \lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_k} . n^{\sum_{m=1}^{k}(i_m - j_m)} . S\begin{pmatrix} i_1, i_2, \ldots, i_k \\ j_1, j_2, \ldots, j_k \end{pmatrix}$$

for $1 \leqslant k \leqslant N$, $1 \leqslant \begin{matrix} i_1 < i_2 < \cdots < i_k \\ j_1 < j_2 < \cdots < j_k \end{matrix} \leqslant N$, and the result is manifestly at hand.

3. THE POSITIVITY OF B_n^r

THEOREM 3.1. A_N^r is column-stochastic for each real $r \geq N-1$.

PROOF. Following Rosenbloom (1967)'s divided difference approach to defining matrix-valued functions, we have

$$(A_N^r)_{ik} = \lambda_i^r \delta_{ik} + \left[\lambda_i, \lambda_k\right]_{\lambda^r} \cdot a_{ik} + \sum_{j=1}^{k-i-1} \sum_{i<i_1<\cdots<i_j<k} \left[\lambda_i, \lambda_{i_1}, \ldots, \lambda_{i_j}, \lambda_k\right]_{\lambda^r} \cdot a_{ii_1} a_{i_1 i_2} \cdots a_{i_j k}, \quad 1 \leqslant i \leqslant k \leqslant N . \tag{3.1}$$

Observing that, for $1 \leqslant j < k-i$ and $r \geqslant k-i$,

$$\left[\lambda_i, \lambda_{i_1}, \ldots, \lambda_{i_j}, \lambda_k\right]_{\lambda^r} = \frac{1}{(j+1)!} \frac{d^{j+1}}{d\lambda^{j+1}} (\lambda^r)\Big|_{\lambda=\theta} = \binom{r}{j+1} \theta^{r-j-1} > 0$$

for some $\theta \in (\lambda_k, \lambda_i)$, it follows that

$$r \geqslant k-i \Longrightarrow (A_N^r)_{ik} > 0 , \quad 1 \leqslant i \leqslant k \leqslant N . \tag{3.2}$$

Let $V = V_N(n) = (v_{ij})$ denote the eigenmatrix of A_N normalized so that $v_{ii} = 1$, $1 \leqslant i \leqslant N$, and let the elements of V^{-1} be v_{ij}^*. The matrices V and V^{-1} are both upper triangular, and it is shown in Kelisky and Rivlin (1967) that the first row of V^{-1} consists of all 1's and that the column sums of V are all 0, except the first which is 1.

From the spectral representation

$$A_N^r = V \Lambda^r V^{-1}$$

where $\Lambda = \text{diag}(\lambda_i)$, and the properties of V and V^{-1} referred to above, we obtain

$$\sum_{i=1}^{k} (A_N^r)_{ik} = \sum_{i=1}^{k} \sum_{j=i}^{k} v_{ij} \, v^*_{jk} \lambda_j^r$$

$$= \sum_{j=1}^{k} \left(\sum_{i=1}^{j} v_{ij} \right) v^*_{jk} \lambda_j^r$$

$$= \sum_{j=1}^{k} \delta_{1j} \, v^*_{jk} \lambda_j^r$$

$$= v^*_{1k}$$

$$= 1 \, , \qquad 1 \leqslant k \leqslant N \, , \quad r \in \mathbb{R} \, . \tag{3.3}$$

That A_N^r is column-stochastic for each real $r \geqslant N-1$ follows now readily from properties (3.2) and (3.3).

<u>COROLLARY</u>. Like $B_n : \mathcal{P}_N \longrightarrow \mathcal{P}_N$, B_n^r $(r \geqslant N-1)$ is a linear positive operator of unit uniform norm.

We shall see in the next section how the condition $r \geqslant N-1$ can be relaxed to $r \geqslant 0$ and yet implying that B_n^r is either TP or nearly TP in the sense that the matrix representing it is either TP or replaceable, elementwise and arbitrarily closely, by a TP matrix.

4. THE LIMITING BEHAVIOUR OF B_n^r

4.1. The case of n fixed and $r \longrightarrow \infty$.

THEOREM 4.1.

a) $(A_N^r)_{1k} \uparrow 1 , \quad 1 \leqslant k \leqslant N ,$

b) $(A_N^r)_{ik} \downarrow 0 , \quad 2 \leqslant i \leqslant k \leqslant N .$

PROOF. Bearing in mind (3.1) and (3.2), recalling that $\lambda_1 = 1$, $0 < \lambda_i < 1$ for $2 \leqslant i \leqslant N$, and that

$$\left[\lambda_{m_1}, \lambda_{m_2}, \ldots, \lambda_{m_k}\right]_{\lambda^r} = \sum_{i=1}^{k} \frac{\lambda_{m_i}^r}{\Pi'(\lambda_{m_i})}, \qquad \Pi(\lambda) = \prod_{i=1}^{k} (\lambda - \lambda_{m_i}) , \quad (4.1)$$

then part b) follows immediately. As for part a),

$$(A_N^r)_{11} = 1$$

$$(A_N^r)_{1k} \uparrow \frac{a_{1k}}{1-\lambda_k} + \sum_{j=1}^{k-2} \sum_{1<i_1<\cdots<i_j<k} \frac{a_{1i_1} a_{i_1 i_2} \cdots a_{i_j k}}{(1-\lambda_{i_1})(1-\lambda_{i_2})\cdots(1-\lambda_k)}$$

$$= \prod_{m=2}^{k} \left(\frac{\sum_{i=1}^{m-1} a_{im}}{(1-\lambda_m)} \right) = 1 , \qquad 2 \leqslant k \leqslant N .$$

COROLLARY 1 (Kelisky - Rivlin). For $r = r_n \in \mathbb{N}$,

$$A_N^r \longrightarrow E = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & & O & \end{bmatrix} .$$

PROOF. This is obviously contained in a) and b), which not only hold for $r \in \mathbb{R}$ but enlighten the structural limiting properties of A_N^r .

COROLLARY 2. For each f in C

$$B_n^r(f;x) \longrightarrow B_n^{\infty}(f;x) \equiv B_1(f;x)$$

uniformly in $0 \leqslant x \leqslant 1$.

PROOF. We may replace f with $L_n(f;x) = X^T p$, the interpolating polynomial for f at the nodes k/n, k=0(1)n. Since $A_{N+1}^r = \begin{bmatrix} 1 & 0 \\ 0 & A_N^r \end{bmatrix}$, then, for n = N, Theorem 4.1 gives

$$A_{n+1}^r \longrightarrow A_{n+1}^{\infty} = \begin{bmatrix} 1 & 0 \\ 1 & E \end{bmatrix}$$

Therefore,

$$B_n^r(f;x) = X^T A_{n+1}^r p$$

$$\longrightarrow X^T A_{n+1}^{\infty} p = p_0 + \left(\sum_{i=1}^{n} p_i \right) x = f(0) + (f(1) - f(0))\, x \; .$$

Van der Steen, Sikkema (1966), and Kelisky and Rivlin (1967) have proved Corollary 2, only for $r = r_n \in \mathbb{N}$, by different methods.

COROLLARY 3. Let K be the (n+1)x(n+1) matrix representing $B_n : \mathcal{P}_n \longrightarrow \mathcal{P}_n$ when we take for $\mathcal{P}_n$ the Bernstein basis $\{q_k\}_{k=0}^n$, i.e., $B_n q_j = Q^T K e_j$, j=0(1)n, e_j being the (n+1)-component vector with 1 in the j^{th} position and 0 elsewhere. Then

$$K^r \longrightarrow K^\infty = \begin{bmatrix} 1 - 0/n & 0 & \dots & 0 & 0/n \\ 1 - 1/n & 0 & \dots & 0 & 1/n \\ 1 - 2/n & 0 & \dots & 0 & 2/n \\ \dots & \dots & \dots & \dots & \dots \\ 1 - n/n & 0 & \dots & 0 & n/n \end{bmatrix} .$$

PROOF. This follows from Corollary 2 and the observation that in K^∞ corresponding to B_n^∞, the k^{th} column of K^∞ consists of the coefficients of $B_n^\infty q_k$:

$$B_n^r q_0 \longrightarrow B_n^\infty q_0 = 1-x = \sum_{j=0}^{n} (1 - j/n)\, q_j$$

$$B_n^r q_k \longrightarrow B_n^\infty q_k = 0 , \quad k=1(1)n-1 ,$$

$$B_n^r q_n \longrightarrow B_n^\infty q_n = x = \sum_{j=0}^{n} (j/n)\, q_j .$$

Nielson, Riesenfeld, and Weiss (1976) have offered a proof of Corollary 3, only for $r = r_n \in \mathbb{N}$, using probabilistic arguments.

4.2. The case of $n \longrightarrow \infty$ and $r = r_m \longrightarrow t \in \mathbb{R}$, fixed, as $m \longrightarrow \infty$ independently of n.

The analysis of the limiting behaviour of the matrix $A_N^t(n)$, where t is a fixed real, N a fixed natural, and $n \longrightarrow \infty$, is considerably simplified if we observe that

$$a_{ii} = \lambda_i = 1 - \frac{\binom{i}{2}}{n} + O(n^{-2}) , \qquad 1 \leqslant i \leqslant N ,$$

$$a_{i,i+1} = \lambda_i \frac{\binom{i+1}{2}}{n} + O(n^{-2}) , \qquad 1 \leqslant i < N ,$$

$$a_{ik} = O(n^{i-k}) , \qquad 1 \leqslant i < k \leqslant N .$$

In view of this we have the following

LEMMA 4.1.

a) $$A_N = I + \frac{1}{n} C_N + \mathcal{O}(n^{-2})$$

b) $$= e^{\frac{1}{n} C_N} + \mathcal{O}(n^{-2})$$

c) $$= e^{\frac{1}{n}(C_N + \mathcal{O}(n^{-1}))} ,$$

where C_N is the bidiagonal matrix whose nonzero entries are given by

$$\begin{aligned} (C_N)_{ii} &= -\mu_i , \quad 1 \leqslant i \leqslant N \\ (C_N)_{i,i+1} &= \mu_{i+1} , \quad 1 \leqslant i < N \end{aligned} \quad , \quad \mu_j = \binom{j}{2} ,$$

with the usual convention that $\binom{i}{j} = 0$ if $i < j$, and $\mathcal{O}(n^{-k})$ denotes an $N \times N$ upper triangular matrix whose nonzero entries are $O(n^{-k})$.

REMARK 4.1. For the monomial x^j it follows that

$$n\left[B_n(x^j;x) - x^j\right] = X^T\left[n(A_N - I)\right]e_j$$

$$\longrightarrow X^T C_N e_j = \mu_j\, x^{j-1} - \mu_j\, x^j = \tfrac{1}{2}\, x(1-x)\, (x^j)'' .$$

Owing to the linearity of the operators B_n and d^2/dx^2 and to the facts that $\|B_n\| = 1$ and $\mathcal{P} = \bigcup_{N=0}^{\infty} \mathcal{P}_N$ is dense in C, we then get the Voronovskaya's result that

$$\lim_{n\to\infty} n\left[B_n(f;x) - f(x)\right] = \tfrac{1}{2}\, x(1-x)\, f''(x)$$

provided that f has a second derivative at $x \in [0,1]$, the convergence being uniform in $0 \leq x \leq 1$ whenever $f''(x)$ is continuous.

COROLLARY 1.

i) $\lim_{n\to\infty} A_N^n = e^{C_N}$

ii) $\lim_{n\to\infty} A_N^{[nt]} = \lim_{n\to\infty} A_N^{nt} = e^{tC_N}, \quad t \geq 0 .$

PROOF. Part i) is immediate. The first equality in ii) follows from the fact that $nt \sim [nt]$ and the second follows from i).

REMARK 4.2. Since each element of the sequence $\left\{A_N^{[nt]}\right\}$, $t>0$, is column-stochastic and $\triangle$STP, then so is the limit $\exp(tC_N)$. It is also clear that for $t>0$ and n sufficiently large, A_N^{nt} is either $\triangle$STP or replaceable, elementwise and arbitrarily closely, by a $\triangle$STP matrix, namely $A_N^{[nt]}$.

COROLLARY 2. Let $r = r_m \longrightarrow t \in \mathbb{R}$, fixed, as $m \longrightarrow \infty$ independently of n, then

$$\lim_{n,m \to \infty} A_N^{r_m} = I . \qquad (4.2)$$

PROOF.

$$\lim_{n,m \to \infty} A_N^{r_m} = (\lim_{n \to \infty} A_N)^t = I .$$

REMARK 4.3. For $t > 0$ and n sufficiently large A_N^t is either ΔSTP or can be approximated, elementwise and arbitrarily closely, by a ΔSTP matrix, namely, $\exp(\frac{t}{n} C_N)$. Indeed, from c) and b) of Lemma 4.1,

$$A_N^t = e^{\frac{t}{n} C_N} + \mathcal{O}(n^{-2}) .$$

REMARK 4.4. In correspondence with (4.2) we have, for each $P \in \mathcal{P}_N$ and $r_m \longrightarrow t \in \mathbb{R}$,

$$\lim_{n,m \to \infty} B_n^{r_m}(P;x) = P(x)$$

uniformly in $0 \leqslant x \leqslant 1$. Being $\mathcal{P}$ dense in C and, for $r_m > 0$, $\left\| B_n^{r_m} \right\| = 1$, then, for each f in C,

$$\lim_{n,m \to \infty} B_n^{r_m}(f;x) = f(x)$$

uniformly in $0 \leqslant x \leqslant 1$. For $r_m \longrightarrow 1$ we recover the Bernstein uniform approximation theorem.

4.3. The case of $r = r_n \longrightarrow \infty$.

Since the eigenvalues of A_N are all positive, then

$$\lim_{n\to\infty} A_N^{r_n} = \left(\lim_{n\to\infty} A_N^n \right)^{\lim_{n\to\infty} \frac{r_n}{n}} = \left(e^{C_N} \right)^{\lim_{n\to\infty} \frac{r_n}{n}}$$

exists iff $\lim_{n\to\infty} r_n/n$ exists. Let $r_n/n \longrightarrow t \in \mathbb{R}$, then $A_N^{r_n} \longrightarrow e^{tC_N}$, a column-stochastic ΔSTP matrix provided $t>0$.

REMARK 4.5. Referring to (3.1) and (4.1), the coefficients of $\lambda_i^{r_n}$, $\lambda_{i_m}^{r_n}$, $1 \leq m \leq j < k-i$, and $\lambda_k^{r_n}$ in

$$\left[\lambda_i, \lambda_{i_1}, \ldots, \lambda_{i_j}, \lambda_k\right]_{\lambda^{r_n}} \cdot a_{ii_1} a_{i_1 i_2} \cdots a_{i_j k} = \frac{a_{ii_1} a_{i_1 i_2} \cdots a_{i_j k}}{(\lambda_i - \lambda_k) \prod_{m=1}^{j} (\lambda_i - \lambda_{i_m})} \cdot \lambda_i^{r_n}$$

$$+ \sum_{m=1}^{j} \frac{a_{ii_1} a_{i_1 i_2} \cdots a_{i_j k}}{(\lambda_{i_m} - \lambda_i)(\lambda_{i_m} - \lambda_k) \prod_{\substack{p=1 \\ p \neq m}}^{j} (\lambda_{i_m} - \lambda_{i_p})} \cdot \lambda_{i_m}^{r_n}$$

$$+ \frac{a_{ii_1} a_{i_1 i_2} \cdots a_{i_j k}}{(\lambda_k - \lambda_i) \prod_{m=1}^{j} (\lambda_k - \lambda_{i_m})} \cdot \lambda_k^{r_n}, \qquad 1 \leq j < k-i,$$

are easily seen to have numerators $O(1/n^{k-i})$, denominators $O(1/n^{j+1})$, and thus all tend to 0 as $n \longrightarrow \infty$, except when $j = k-i-1$. Therefore,

$$b_{ik}(t) \equiv (e^{tC_N})_{ik} = e^{-\mu_i t} \delta_{ik} + \lim_{n\to\infty} \left[\lambda_i, \lambda_{i+1}, \ldots, \lambda_{k-1}, \lambda_k\right]_{\lambda^{r_n}} \cdot a_{i,i+1} \cdots a_{k-1,k}$$

and we recover, after some manipulation, the result by Kelisky and Rivlin (1967) that

$$b_{ik}(t) = \sum_{j=i}^{k} \beta_{i,j,k}\, e^{-\mu_j t}, \qquad 1 \leqslant i \leqslant k \leqslant N \ ; \quad t \geqslant 0 , \tag{4.3}$$

with

$$\beta_{i,j,k} = (-1)^{j-i} \frac{i}{k} \frac{\binom{k}{j}^2 \binom{j}{i}^2}{\binom{2j-2}{j-i}\binom{k+j-1}{2j-1}} , \qquad 1 \leqslant i \leqslant j \leqslant k . \tag{4.4}$$

In particular,

$$r_n/n \downarrow 0 \Longrightarrow A_N^{r_n} \longrightarrow I \quad \text{and} \quad r_n/n \uparrow \infty \Longrightarrow A_N^{r_n} \longrightarrow E .$$

The coefficients $\beta_{i,j,k}$ satisfy the following sets of seemingly nontrivial identities:

THEOREM 4.2.

a) $\displaystyle\sum_{i=1}^{j} \beta_{i,j,k} = 0 , \quad 2 \leqslant j \leqslant k ,$

b) $\displaystyle\sum_{j=i}^{k} \beta_{i,j,k}\, \mu_j^m = 0 , \quad 1 \leqslant i \leqslant k , \quad 0 \leqslant m < k-i .$

PROOF. Referring to the representation

$$e^{tC_N} = I + \sum_{m=1}^{\infty} \frac{t^m}{m!} C_N^m , \tag{4.5}$$

we see that the property that the column sums of $\exp(tC_N)$ are all equal to 1 is equivalent to the vanishing of the column sums of C_N^m , $m \geqslant 1$.

Direct multiplication reveals that $(C_N^m)_{ik} = 0$ for $k > i+m$, $m \geqslant 0$, and thus, from (4.5),

$$b_{ik}(t) = \sum_{m=k-i}^{\infty} \frac{(C_N^m)_{ik}}{m!} t^m , \qquad 1 \leqslant i < k , \tag{4.6}$$

showing that the functions $b_{ik}(t)$, $1 \leqslant i < k$, have $t = 0$ as a zero of multiplicity k-i. On the other hand, from (4.3),

$$b_{ik}(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \left(\sum_{j=i}^{k} \beta_{i,j,k} \mu_j^m \right) t^m , \tag{4.7}$$

and part b) follows.

Comparing like powers of t in (4.6) and (4.7) leads to

$$(C_N^m)_{ik} = (-1)^m \sum_{j=i}^{k} \beta_{i,j,k} \mu_j^m , \qquad 1 \leqslant i < k , \quad m \geqslant 0 .$$

The vanishing of the column sums of C_N^m , $m \geqslant 1$, implies that

$$\sum_{j=2}^{k} \left(\sum_{i=1}^{j} \beta_{i,j,k} \right) \mu_j^m = 0 , \qquad k \geqslant 2 , \quad m \geqslant 0 ,$$

and part a) follows from the arbitrariness of m.

<u>REMARK 4.6</u>. The set of identities a) and that corresponding to m = 0 in b) were first observed by Kelisky and Rivlin (1967).

With the change of variable $x = e^{-t}$ the range $[0,\infty]$ is transformed into $[0,1]$ and the functions $b_{ik}(t)$ may be written as polynomials in x:

$$b_{ik}(t) = \tilde{b}_{ik}(x) = x^{\mu_i} \sum_{j=i}^{k} \beta_{i,j,k} \, x^{\mu_j - \mu_i} ,$$

showing that $x = 0$ $(t = \infty)$ is a zero of $\tilde{b}_{ik}$ (b_{ik}) of multiplicity μ_i . Being $x = 1$ $(t = 0)$ a zero of $\tilde{b}_{ik}$ (b_{ik}) of multiplicity k-i, then we have the following

THEOREM 4.3.

$$\tilde{b}_{ik}(x) = x^{\mu_i}(1-x)^{k-i}\, p_{ik}(x)\ , \quad 1\leq i\leq k\ ,$$

with p_{ik} in $\mathcal{P}_s$, $s = (\mu_k-\mu_i)-(k-i) = \frac{1}{2}(k-i)(k+i-3)$, having every coefficient positive.

PROOF. The coefficients $c_j = c_j(i,k)$ of $p_{ik}(x) = \sum_{j=0}^{s} c_j\, x^j$ may be determined by equating coefficients of like powers of x in

$$\sum_{j=i}^{k} \beta_{i,j,k}\, x^{\mu_j-\mu_i} = (1-x)^{\nu}\, p_{ik}(x)\ , \quad \nu = k-i\ . \tag{4.8}$$

Assuming that $p_{ik}(x)$ has ω sign variations, then $(1-x)^{\nu}\, p_{ik}(x)$ will have at least $\omega+\nu$ (see Pólya and Szegö (1976, Probl. 30, p. 40)). On the other hand, by (4.4) and (4.8), the latter has precisely ν variations. Hence $\omega = 0$, i.e., all c_j's have the same sign. But $p_{ik}(x)>0$ for $x>0$ and therefore $c_j>0$, $0\leq j\leq s$.

5. CONVEXITY PRESERVING PROPERTIES OF B_n^r

There is a sharply contrasting behaviour between B_n^r and B_n^{-r} , $r>0$:

i) While B_n^r , having no eigenvalue >1, is contractive, variation diminishing, and norm not increasing, B_n^{-r} is a variation increasing dilatation which increases the norm unboundedly as $r \longrightarrow \infty$.

ii) For $r>0$ and $n>2$ the transformation B_n^r is convexity preserving or nearly so inasmuch as the matrix $A^r(n)$ representing it is TP or replaceable, elementwise and arbitrarily closely, by a TP matrix. In contrast, B_n^{-r} has no such property, as shown by the example

$$q_n(n,x) = x^n = B_n(B_n^{-1}x^n;x) = \sum_{k=0}^{n} B_n^{-1}(x^n;k/n)\, q_k(n,x) ,$$

which implies, by the linear independence of the q_k's, that

$$B_n^{-1}(x^n;k/n) = \begin{cases} 0 , & k=0(1)n-1 , \\ 1 , & k=n . \end{cases}$$

Therefore, while x^n is convex on $[0,1]$,

$$B_n^{-1}(x^n;k/n) = \prod_{k=0}^{n-1}\left(\frac{x - k/n}{1 - k/n}\right) = \frac{n^n}{n.}\, x(x-\frac{1}{n})\ldots(x-\frac{n-1}{n})$$

oscillates n times about zero.

REMARK 5.1. B_2^r is convexity preserving for each real r.

Indeed, let the function f be defined and convex on $[0,1]$. The ordinates $f_k = f(k/2)$, $k=0,1,2$, satisfy $\Delta^2 f_0 \geq 0$. Let

$$P(x) = B_2(f;x) = B_2(L_2 f;x) = f_0 + 2\Delta f_0\, x + \Delta^2 f_0\, x^2 .$$

Then

$$B_2^r(f;x) = B_2^{r-1}(P;x) = \begin{bmatrix}1, x, x^2\end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 1-2^{1-r} \\ & & 2^{1-r} \end{bmatrix} \begin{bmatrix} f_0 \\ 2\Delta f_0 \\ \Delta^2 f_0 \end{bmatrix} \qquad (5.1)$$

$$= f_0 + \left[2\Delta f_0 + (1 - 2^{1-r})\Delta^2 f_0\right] x + 2^{1-r}\Delta^2 f_0\, x^2$$

is manifestly convex for each real r.

It should be noticed that B_2^r is convexity preserving for every real r, whereas the 3x3 matrix in (5.1) is not positive for $r<1$, much less TP.

CHAPTER 4

THE LIMITING SEMIGROUP $\{\mathcal{B}_t;\ t \geq 0\}$ OF BERNSTEIN ITERATES

1. EXISTENCE, CHARACTERIZATION, AND REPRESENTATION OF $\mathcal{B}_t$

1.1. Existence and characterization of $\mathcal{B}_t$ acting on $\mathcal{P}_N$.

In Chapter 3 the main object of study was the matrix $A_{N+1}^{r_n}$ representing $B_n^{r_n}$ acting on $\mathcal{P}_N$. In the terminology of the operator semi-group theory $\{A_{N+1}^{r_n};\ r_n \geq 0\}$ is a semigroup of positive matrices. This semi-group was shown to converge to the semigroup of TP matrices $\{\exp(tC_{N+1};\ t \geq 0\}$ where

$$C_{N+1} = \lim_{n \to \infty} n(A_{N+1} - I)$$

iff $r_n/n \to t$ as $n \to \infty$. Under this condition, it was shown that

$$\exp(tC_{N+1}) = \lim_{n \to \infty} A_{N+1}^{r_n} = \lim_{n \to \infty} A_{N+1}^{[nt]} .$$

In view of this, if we call

$$\mathcal{B}_t p(x) = \lim_{n \to \infty} B_n^{r_n} p(x) = \lim_{n \to \infty} B_n^{[nt]} p(x)$$

$$= X^T \exp(tC_{N+1})\ p ,$$

where p denotes the coefficient vector of $P \in \mathcal{P}_N$ and $X^T = (1, x, x^2, \dots, x^N)$, that limit exists for any $t \geq 0$ and all $P \in \mathcal{P}_N$, i.e., there exists a totally positive semigroup $\{\mathcal{B}_t;\ t \geq 0\}$ on $\mathcal{P}_N$ to itself, with $\mathcal{B}_0 = I$ and such that, for each $t \geq 0$ and all $P \in \mathcal{P}_N$,

$$\lim_{n \to \infty} \left\| B_n^{[nt]} P - \mathcal{B}_t P \right\| = 0 .$$

Furthermore, the operators B_n (and likewise $B_n^{r_n}$, $r_n \geq 0$) have norm 1 since they are positive and preserve the unit function. Therefore $\{\mathcal{B}_t;\ t \geq 0\}$ is a totally positive, strongly continuous, contraction semigroup on $\mathcal{P}_N$, $\{\exp(tC_{N+1});\ t \geq 0\}$ being its matrix representation when we take for $\mathcal{P}_N$ the usual basis $\{x^k\}_{k=0}^N$.

The <u>infinitesimal generator</u> D of the semigroup $\{\mathcal{B}_t\}$ is defined as

$$D\, x^k = \lim_{t \downarrow 0} (\mathcal{B}_t(x^k; x) - x^k)/t \quad , \qquad k = 0(1)N , \qquad (1.1)$$

$$= \lim_{t \downarrow 0} X^T.(\exp(tC_{N+1} - I)/t).e_k$$

$$= X^T\, C_{N+1}\, e_k$$

$$= \mu_k\, x^{k-1} - \mu_k\, x^k \quad , \qquad \mu_k = \binom{k}{2} \quad ,$$

$$= \tfrac{1}{2}\, x(1-x)(x^k)'' ,$$

i.e.,

$$D = \tfrac{1}{2}\, x(1-x)\, d^2/dx^2 .$$

D is a linear differential operator acting on $\mathcal{D}(D) = C^2$, and C_{N+1} is the matrix representation of the restriction of D to $\mathcal{P}_N$.

Invoking the result of Voronovskaya (see Remark 4.1 in Chapter 3) we also have

$$D\,x^k = \lim_{n\to\infty} n(B_n(x^k;x) - x^k)\ , \qquad k = 0(1)N\ ,$$

hence

$$D = \lim_{n\to\infty} n(B_n - I)\ . \tag{1.2}$$

The foregoing results on the existence and characterization of the semigroup $\{\mathcal{B}_t;\ t\geqslant 0\}$ are contained in the theorem of Trotter(1958) on the convergence of the iterates of contractive mappings on Banach spaces. In what follows, however, there will be no need of the full strength of Trotter's result.

Introducing the notation

$$\begin{aligned} W_k(t,x) \equiv \mathcal{B}_t(x^k;x) &= \lim_{n\to\infty} B_n^{[nt]}(x^k;x)\ , \qquad k = 1(1)N\ , \\ &= X^T \exp(tC_N)\, e_k = \sum_{i=1}^{k} b_{ik}(t)\, x^i\ , \end{aligned} \tag{1.3}$$

and using the semigroup property of $\mathcal{B}_t$,

$$h^{-1}(W_k(t+h,x) - W_k(t,x)) = h^{-1}(\mathcal{B}_h - I)\, W_k(t,x)\ , \qquad h>0\ ,$$

and, similarly,

$$h^{-1}(W_k(t,x) - W_k(t-h,x)) = h^{-1}(\mathcal{B}_h - I)\mathcal{B}_{t-h}(x^k;x)\ , \qquad 0<h<t\ .$$

Letting $h \to 0$, recalling Remark 4.5 in Chapter 3 and using (1.1) we are led to the following initial value problem(Cauchy problem)

$$\begin{cases} \partial W_k/\partial t = D\,W_k\ , & W_k = W_k(t,x)\ , \\ W_k(0,x) = x^k\ , & W_k(\infty,x) = x\ , \\ W_k(t,0) = 0\ , & W_k(t,1) = 1 \end{cases} \tag{1.4}$$

of which (1.3) is the unique solution on

$$\Omega = \left\{ (t,x) \in \mathbb{R}_2\ :\ 0 \leq x \leq 1,\quad 0 \leq t \leq \infty \right\}. \tag{1.5}$$

<u>REMARK 1.1</u>. We recall, for clarity sake, that the end condition $W_k(\infty,x) = x$ corresponds to the fact

$$\mathcal{B}_\infty(x^k;x) = \lim_{t\to\infty} \mathcal{B}_t(x^k;x) = \lim_{n\to\infty} B_n^{r_n}(x^k;x) = B_1(x^k;x) = x$$

iff $\lim_{n\to\infty} r_n/n = \infty$, whereas the side conditions $W_k(t,0) = 0$ and $W_k(t,1) = 1$ correspond to the interpolating properties $B_n(x^k;0) = 0$ and $B_n(x^k;1) = 1$ respectively.

1.2. Spectral characteristics of $\mathcal{B}_t$.

From Chapter 3 (see Section 3 and Lemma 4.1 a))

$$\begin{aligned} C_N &= \lim_{n\to\infty} n(A_N(n) - I) \\ &= \lim_{n\to\infty} n(V_N(n)\,\Lambda_N(n)V_N^{-1}(n) - I) \\ &= \lim_{n\to\infty} V_N(n).\ \lim_{n\to\infty} n(\Lambda_N(n) - I).\ \lim_{n\to\infty} V_N^{-1}(n) \\ &= U_N.\ \mathrm{diag}(-\mu_j).\ U_N^{-1} . \end{aligned}$$

Instead of evaluating the eigenmatrix $U_N = (u_{ij})$ of C_N as $U_N = \lim_{n\to\infty} V_N(n)$ as in Kelisky and Rivlin (1967), we take advantage of the simple structure of C_N and use the equation

$$C_N U_N = U_N\ \mathrm{diag}(-\mu_j) ,$$

with U_N normalized so that its diagonal elements are equal to 1, to obtain

$$u_{ij} = -\frac{\mu_{i+1}}{\mu_j - \mu_i}\,u_{i+1,j} , \qquad i < j = 2(1)N$$

$$= (-1)^{j-i} \prod_{m=i}^{j-1} \left(\frac{\mu_{m+1}}{\mu_j - \mu_m} \right)$$

$$= \begin{cases} 0 & , \quad i > j \\ 1 & , \quad i = j \ ; \quad i,j = 1(1)N , \\ (-1)^{j-i} \dfrac{i}{j} \dfrac{\binom{j}{i}^2}{\binom{2j-2}{j-i}} & , \quad i < j . \end{cases}$$

REMARK 1.2. The matrix $U_N^{-1} = (u^*_{ij}) = \lim_{n\to\infty} V^{-1}(n)$ is obtained in a like manner. We use the equation

$$U_N^{-1} C_N = \mathrm{diag}(-\mu_j)\; U_N^{-1}$$

to get

$$u^*_{jk} = \frac{\mu_k}{\mu_k - \mu_j}\, u^*_{j,k-1} \quad , \qquad j<k = 2(1)N \; ,$$

$$= \prod_{m=j}^{k-1} \left(\frac{\mu_{m+1}}{\mu_{m+1} - \mu_j} \right)$$

$$= \begin{cases} 0 & , \quad j>k \\ 1 & , \quad j = k \;\; ; \quad j,k = 1(1)N \; , \\ \dfrac{j}{k}\, \dfrac{\binom{k}{j}^2}{\binom{k+j-1}{k-j}} & , \quad j<k \end{cases}$$

and we recover, once again, the Kelisky and Rivlin's result referred to in Remark 4.5 of Chapter 3 :

$$b_{ik}(t) \equiv (\exp(tC_N))_{ik} = \sum_{j=i}^{k} u_{ij}\, u^*_{jk}\, e^{-\mu_j t} = \sum_{j=i}^{k} \gamma_{i,j,k}\, e^{-\mu_j t} \; .$$

Iteration of

$$B_n\, v_j = \lambda_j\, v_j \quad , \qquad 1 \leq j \leq N \leq n \; , \tag{1.6}$$

where

$$v_j = v_j(n,x) = \sum_{i=1}^{j} v_{ij}(n)\, x^i \; ,$$

leads to

$$\mathcal{B}_t\, u_j = e^{-\mu_j t}\, u_j \; , \tag{1.7}$$

where

$$u_j = u_j(x) = \lim_{n\to\infty} v_j(n,x) = \sum_{i=1}^{j} (-1)^{j-i}\, \frac{i}{j}\, \frac{\binom{j}{i}^2}{\binom{2j-2}{j-i}}\, x^i \; .$$

Clearly, $u_j(0) = 0$ and, owing to the vanishing of the column sums of $V_N(n)$, $u_j(1) = 0$ as well. Thus, all the polynomials $u_j(x)$, $j \geq 2$, have the common factor $x(x-1)$. These polynomials are the only eigenfunctions of $\mathcal{B}_t$ with associated eigenvalues $\exp(-\mu_j t)$. We also note that 1 and x are eigenfunctions associated with the common eigenvalue 1.

It follows from (1.7) that

$$\frac{1}{t}(\mathcal{B}_t - I)\, u_j = \frac{1}{t}(e^{-\mu_j t} - 1)\, u_j$$

and, letting $t \downarrow 0$,

$$D\, u_j = -\mu_j\, u_j\ , \tag{1.8}$$

i.e.,

$$x(1-x)\, u_j''(x) + j(j-1)\, u_j(x) = 0\ .$$

But, for $j \geqslant 0$,

$$u_{j+2}(x) = x(x-1)\, \phi_j(x),$$

where $\phi_j(x)$ is a j^{th} degree polynomial with leading coefficient 1, due to the way we have normalized the matrix $V_N(n)$. Therefore

$$x(1-x)\, \phi_j''(x) - 2(2x-1)\, \phi_j'(x) + j(j+3)\, \phi_j(x) = 0,$$

giving

$$\phi_j(x) = P_j^{(1,1)}(2x-1) \Big/ \binom{2j+2}{j}\ ,$$

where $P_j^{(1,1)}(x)$ denotes the j^{th} degree Jacobi polynomial of parameters (1,1) normalized in the usual way, i.e., so that $P_j^{(1,1)}(1) = \binom{j+1}{j} = j+1$.

$\left\{\phi_j(x)\right\}_0^\infty$ is an orthogonal polynomial system on $[0,1]$ with respect to the weight function $x(1-x)$. Hence $\left\{u_j(x)\right\}_2^\infty$ is also an orthogonal polynomial system on that interval with respect to the weight function $(x(1-x))^{-1}$.

1.3. Integral representation of $\mathscr{B}_t$ acting on C.

Being $\mathscr{B}_t$ bounded and $\mathscr{D}(D)$ dense in C, then $\mathscr{B}_t$ can be extended to all of C; that is,

$$W(t,x) \equiv \mathscr{B}_t f(x) = \lim_{n\to\infty} B_n^{[nt]} f(x)$$

exists for each f in C, and we set out to work out an explicit representation for $\mathscr{B}_t f$. To begin with, we assume that $f \in \mathscr{D}(D)$. Then so is $\mathscr{B}_t f$ because $\mathscr{B}_t Df = D\mathscr{B}_t f = \frac{d}{dt}(\mathscr{B}_t f)$ (see, e.g., Butzer and Berens (1967,p.9)) and we set up, as in Subsection 1.1, the following Cauchy problem on Ω given by (1.5) :

$$\begin{cases} \partial W/\partial t = D\,W \quad , & W = W(t,x) \quad , \\ W(0,x) = f(x) \quad , & W(\infty,x) = B_1 f(x) \quad , \\ W(t,0) = f(0) \quad , & W(t\ ,1) = f(1) \end{cases} \tag{1.9}$$

(see Remark 1.1 on the end and side conditions).

With problem (1.4) as a guide, we expand W(t,x) into the eigenfunction system $\{u_j(x)\}_0^\infty$, with coefficients which are functions of t :

$$W(t,x) = \sum_{j=0}^{\infty} c_j(t)\, u_j(x) \; .$$

The differential equation in (1.9) now separates into the ordinary differential equation

$$c_j'(t) + \mu_j c_j(t) = 0$$

which is solved by

$$c_j(t) = \bar{c}_j \, e^{-\mu_j t} .$$

Our problem is now reduced to the determination of the coefficients $\bar{c}_j$. Making use of the initial condition in (1.9),

$$f(x) = \sum_{j=0}^{\infty} \bar{c}_j \, u_j(x)$$

$$= \bar{c}_0 + \bar{c}_1 x + \sum_{j=2}^{\infty} \bar{c}_j \, u_j(x) . \tag{1.10}$$

Recalling that, for $j \geq 2$, $u_j(0) = u_j(1) = 0$, we obtain

$$\bar{c}_0 = f(0) \quad \text{and} \quad \bar{c}_1 = f(1) - f(0).$$

Defining

$$\tilde{f}(y) = f(y) - f(0) - (f(1) - f(0))\, y$$

$$= f(y) - B_1 \, f(y) , \tag{1.11}$$

we have

$$\tilde{f}(y) = \sum_{j=2}^{\infty} \bar{c}_j \, u_j(y) . \tag{1.12}$$

Multiplying both sides of (1.12) by $u_k(y)/(y(1-y))$, $k \geq 2$, and integrating from $y = 0$ to $y = 1$, gives

$$\bar{c}_k = \frac{1}{h_k} \int_0^1 \frac{\tilde{f}(y)\, u_k(y)}{y(1-y)} \, dy \quad , \quad k \geq 2 ,$$

where

$$h_k = \int_0^1 \frac{u_k^2(y)}{y(1-y)} \, dy = \frac{k-1}{k(2k-1) \binom{2k-2}{k}^2} .$$

Finally, we obtain

$$W(t,x) = f(0) + (f(1) - f(0))\, x + \sum_{j=2}^{\infty} \frac{e^{-\mu_j t}}{h_j}\, u_j(x) \int_0^1 \frac{\tilde{f}(y)\, u_j(y)}{y(1-y)}\, dy ,$$

or, which is the same,

$$\mathcal{B}_t\, f(x) = B_1\, f(x) + \int_0^1 G(t;x,y)\, (f(y) - B_1\, f(y))\, dy , \tag{1.13}$$

where

$$G(t;x,y) = \frac{1}{y(1-y)} \sum_{j=2}^{\infty} \frac{e^{-\mu_j t}}{h_j}\, u_j(x)\, u_j(y)$$

$$= \frac{1}{y(1-y)} \sum_{j=0}^{\infty} \frac{e^{-\mu_{j+2} t}}{h_{j+2}} \cdot \frac{x(x-1)}{\binom{2j+2}{j}} P_j^{(1,1)}(2x-1) \cdot \frac{y(y-1)}{\binom{2j+2}{j}} P_j^{(1,1)}(2y-1)$$

$$= x(1-x) \sum_{j=0}^{\infty} H_j\, e^{-\alpha_j t}\, P_j^*(x)\, P_j^*(y) , \tag{1.14}$$

with

$$H_j = \frac{(j+2)(2j+3)}{j+1}$$

$$\alpha_j = \mu_{j+2} = \frac{(j+1)(j+2)}{2} , \quad j = 0,1,2,\ldots ,$$

$$P_j^*(x) = P_j^{(1,1)}(2x-1) .$$

The restriction that f should be in $\mathcal{D}(D)$ can now be removed as the representation (1.13) is clearly valid for each f in C.

Since $\mathcal{B}_t\, f(0) = f(0)$ and $\mathcal{B}_t\, f(1) = f(1)$, there is no loss in generality in assuming that $f(0) = f(1) = 0$. In this case

$$\mathcal{B}_t\, f(x) = \int_0^1 G(t;x,y)\, f(y)\, dy\ .$$

REMARK 1.3. For any $t > 0$ and all f in C, the function $\mathcal{B}_t\, f(x)$ is analytic on $0 \leqslant x \leqslant 1$. Indeed, consulting (1.14), the factor $\exp(-\alpha_j t)$ makes convergent not only the infinite sum representing $G(t;x,y)$ but all its derivatives of arbitrary order, with respect to t or x, for all $t > 0$ and $0 \leqslant x \leqslant 1$.

2. SMOOTHING EFFECTS OF $\mathcal{B}_t$

LEMMA 2.1. For each fixed nonnegative t, the kernel $G(t;x,y)$ of the transformation $\mathcal{B}_t$ is strictly positive in the interior of the unit square $S_2 = \{(x,y) \in \mathbb{R}_2 : 0 \leq x,y \leq 1\}$.

PROOF. Using the orthogonality of the shifted Jacobi polynomials appearing in (1.14) we see that

$$\int_0^1 G(s;x,y_1)\, G(t;y_1,y)\, dy_1 = G(s+t;x,y) \; . \qquad (2.1)$$

It is also easy to show that , for $k \geq 1$,

$$G(t;x,y) = \int_0^1 \int_0^1 \cdots \int_0^1 G\left(\frac{t}{k+1};x,y_1\right).G\left(\frac{t}{k+1};y_1,y_2\right). \cdots .G\left(\frac{t}{k+1};y_k,y\right)\, dy_1 dy_2 \ldots dy_k \; . \qquad (2.2)$$

The kernel $G(t;x,y)$ is clearly continuous on S_2 and, along the diagonal x=y, it is everywhere positive except at the endpoints. Hence, there exists a neighbourhood of the diagonal, say $\vartheta_\epsilon = \{(x,y) : |x - y| < \epsilon\}$, in which $G(t;x,y) > 0$. Now, for any point (x,y) in S_2, there is a finite set $y_1, y_2, \ldots, y_k$ such that all points (x,y_1), $(y_1,y_2), \ldots, (y_k,y)$ lie in ϑ_ϵ . The strict positivity of $G(t;x,y)$ for any $t \geq 0$ and all (x,y) in S_2 except at the corners x=0 and x=1 follows now readily from (2.2).

We have borrowed this elegant idea from Karlin and McGregor (1960).

LEMMA 2.2. For each $t \geq 0$, $G(t;x,y)$ is strictly totally positive in the sense that if m is any positive integer, $0 < x_1 < x_2 < \cdots < x_m < 1$, and $0 \leq y_1 < y_2 < \cdots < y_m \leq 1$, then

$$G\left(t;\begin{matrix}x_1,\ldots,x_m\\y_1,\ldots,y_m\end{matrix}\right) = \begin{vmatrix} G(t;x_1,y_1) & \ldots & G(t;x_1,y_m)\\ \vdots & & \vdots\\ G(t;x_m,y_1) & \ldots & G(t;x_m,y_m)\end{vmatrix} > 0 .$$

PROOF. In terms of the determinantal polynomials

$$P^*\begin{pmatrix}n_1,\ldots,n_m\\x_1,\ldots,x_m\end{pmatrix} = \begin{vmatrix} P^*_{n_1}(x_1) & \ldots & P^*_{n_1}(x_m)\\ \vdots & & \vdots\\ P^*_{n_m}(x_1) & \ldots & P^*_{n_m}(x_m)\end{vmatrix} , \quad 0\leqslant n_1<\ldots<n_m,$$

there is a representation

$$G\left(t;\begin{matrix}x_1,\ldots,x_m\\y_1,\ldots,y_m\end{matrix}\right) = x_1(1-x_1)\ \ldots\ x_m(1-x_m). \sum_{0\leqslant n_1<\ldots<n_m} H_{n_1}\ldots H_{n_m}\, .$$

$$e^{-(\alpha_{n_1}+\ \ldots\ +\alpha_{n_m})t}\, P^*\begin{pmatrix}n_1,\ldots,n_m\\x_1,\ldots,x_m\end{pmatrix} P^*\begin{pmatrix}n_1,\ldots,n_m\\y_1,\ldots,y_m\end{pmatrix}$$

analogous to (1.14). In correspondence with (2.1) we also have

$$G\left(s+t;\begin{matrix}x_1,\ldots,x_m\\y_1,\ldots,y_m\end{matrix}\right) = \int_0^1\ldots\int_0^1 G\left(s;\begin{matrix}x_1,\ldots,x_m\\z_1,\ldots,z_m\end{matrix}\right) G\left(t;\begin{matrix}z_1,\ldots,z_m\\y_1,\ldots,y_m\end{matrix}\right)$$

$$.\, dz_1\ldots dz_m .$$

Using this and repeating the argument we used when dealing with the case m=1 it follows that G(t;x,y) is strictly totally positive on the unit hypercube except at the corners where $x_i=0$, i=1(1)m, at which G(t;x,y) vanishes.

<u>THEOREM 2.1</u>. The semigroup $\{\mathcal{B}_t;\ t \geqslant 0\}$ is <u>variation diminishing</u>.

<u>PROOF</u>. For singular integrals, variation diminishing and total positivity are equivalent properties (see Butzer and Nessel (1971, p. 150).

<u>REMARK 2.1</u>. As an immediate consequence of this result, all the shape preserving properties of the Bernstein operators carry over to the semigroup $\{\mathcal{B}_t\}$. In other words, the graphs of f and $\mathcal{B}_t f$ have the same shape. In particular, if f is monotone or convex, so is $\mathcal{B}_t f$.

<u>REMARK 2.2</u>. An immediate consequence of (1.13) is the invariance of $\mathcal{P}_1$ under $\mathcal{B}_t$. From this and the positivity of $\{\mathcal{B}_t;\ t \geqslant 0\}$ it follows that $\|\mathcal{B}_t\| = \|\mathcal{B}_t 1\| = 1$, just like B_n and its iterates of nonnegative order.

3. APPLICATIONS OF $\mathcal{B}_t$

3.1. Saturation theory for the Bernstein approximation in C.

The Bernstein saturation problem is to determine a positive, non-increasing function ϕ_n (the saturation order) with the property that $\phi_n \downarrow 0$ as $n \to \infty$ and to characterize two classes S (the saturation class) and T (the trivial class) of functions f in C such that

$$B_n(f;x) - f(x) = O(\phi_n) \quad \text{iff} \quad f \in S$$

and

$$B_n(f;x) - f(x) = o(\phi_n) \quad \text{iff} \quad f \in T .$$

The class S consists of all functions f in C optimally approximated by $B_n f$, i.e., no higher order of approximation than ϕ_n can occur except for T, which B_n leaves intact.

K. de Leeuw (1959) was the first to solve this problem following the Voronovskaya's result that the boundedness of f on $[0,1]$ and the existence of f'' at a point $x \in [0,1]$ implies

$$B_n(f;x) - f(x) = \frac{x(1-x)}{2n} f''(x) + o(1/n) \tag{3.1}$$

and the Lorentz' (1953, p. 22) conjecture that the relation

$$B_n(f;x) - f(x) = o(1/n)$$

cannot be true for all $x \in [a,b] \subseteq [0,1]$ unless f is a linear polynomial on $[a,b]$.

An improved solution (in the sense that the behaviour of the saturation order near the endpoints of $[0,1]$ is taken into account) was given by Lorentz (1966, p. 102) through an involved, functional-analytic technique.

There are two alternatives to Lorentz' approach to the theory of saturation of linear positive algebraic polynomial approximation operators, these are :

a) The parabola technique of Bajsanski and Bojanic (1964) where asymptotic relations of Voronovskaya's type (3.1) play a major role. See e.g., DeVore (1972), Lorentz and Schumaker (1972), and Berens (1972) for further developments and applications.

b) The operator semigroup method, first applied by Karlin and Ziegler (1970) and Micchelli (1973). Here, the idea is to derive from a given sequence $\{L_n\}$ of linear approximation operators a continuous semigroup $\{T_t;\ t>0\}$ by taking limits of appropriate iterates of L_n, namely, $L_n^{r_n}$, where $r_n \phi_n \longrightarrow t>0$ as $n \longrightarrow \infty$, ϕ_n being the saturation order. The saturation properties of $\{T_t\}$ are shown to be the same as those of $\{L_n\}$ and saturation for a continuous semigroup is well established in Butzer and Berens (1967).

THEOREM 3.1 (Lorentz-Micchelli). For f in C the following statements are equivalent :

(i) $|f'(x) - f'(y)| \leqslant M|x - y|$, $0 \leqslant x,y \leqslant 1$;

(ii) $|B_n(f;x) - f(x)| \leqslant \frac{M}{2n} x(1-x)$, $n \geqslant 1$, $0 \leqslant x \leqslant 1$;

(iii) $|\mathcal{B}_t(f;x) - f(x)| \leqslant \frac{Mt}{2} x(1-x)$, $t \geqslant 0$, $0 \leqslant x \leqslant 1$.

Moreover,

$$B_n(f;x) - f(x) = o(x(1-x)/n) \quad \text{iff} \quad f \in \mathcal{P}_1 .$$

PROOF. See G. Lorentz (1966, p. 102) for the equivalence of (i) and (ii) of which the last assertion is an immediate consequence, and C. Micchelli (1973) for the equivalence of (ii) and (iii) and (iii) and (i).

To sum up, the Bernstein approximation procedure is saturated with order $x(1-x)/n$, trivial class $\mathcal{P}_1$, and saturation class S consisting of all functions f in C for which f' exists and belongs to the classical Lipschitz class Lip 1.

3.2. Characterizations of convexity.

THEOREM 3.2 (Karlin-Ziegler-Micchelli). The following are necessary and sufficient conditions for f to be convex on $[0,1]$:

(i) $B_n(f;k/n) \geqslant f(k/n)$, $k = 0(1)n$; $n \geqslant 1$;

(ii) $B_n(f;x) \geqslant f(x)$, $n \geqslant 1$; $0 \leqslant x \leqslant 1$;

(iii) $\mathcal{B}_t(f;x) \geqslant f(x)$, $t \geqslant 0$; $0 \leqslant x \leqslant 1$.

PROOF. See S. Karlin and Z. Ziegler (1970) and C. Micchelli (1973).

The next result involves convexity and monotonicity.

THEOREM 3.3. Let $f \in C$ then f is convex on $[0,1]$ iff, for all $0 \leqslant x \leqslant 1$ and $0 \leqslant s \leqslant t$,

$$\mathcal{B}_t f(x) \geqslant \mathcal{B}_s f(x) . \tag{3.2}$$

PROOF. If f is convex and $t > s$, then, by part (iii) of Theorem 3.2

$$\mathcal{B}_{t-s}f(x) \geqslant f(x) , \quad 0 \leqslant x \leqslant 1 ,$$

and the necessity part follows upon application of $\mathcal{B}_s$ to both sides of this inequality.

The sufficiency part follows at once from (3.2) on letting $s \downarrow 0$ and using part (iii) of Theorem 3.2 once again.

3.3. Linear operators commuting with B_n.

Let T be a linear operator mapping C into itself and commuting with B_n :

$$TB_n = B_n T .$$

The characterization of such a transformation was first given by Konheim and Rivlin (1968). It was given later by Karlin and Ziegler (1970) as an application of the iteration method.

Defining $W = T(I - B_1)$, then Karlin and Ziegler's result is that

$$Wf(x) = a + bx + cf(x) + df(1-x) ,$$

where a and b are linear functionals on f and c and d constants depending on B_n.

See Subsection 3.6.3 for a detailed extension of this result to functions of two independent variables.

3.4. Saturation theory for de Leeuw-like operators.

We call de Leeuw-like operators the following polynomial approximation operators defined for f in C by

$$K_n(f;x) = \sum_{k=0}^{n} l^*_{nk}(f)\, q_k(n,x)$$

with

$$l^*_{nk}(f) = \begin{cases} f(0) & , \quad k = 0 \\ n \displaystyle\int_{-1/(2n)}^{+1/(2n)} f(k/n + t)dt, & \quad k = 1(1)n-1 , \\ f(1) & , \quad k = n . \end{cases}$$

$K_n f$ generalizes $B_n f$, which corresponds to the point evaluation functionals $l^*_{nk}(f) = f(k/n)$, and has been introduced by de Leeuw (1959) in his treatment of the Bernstein saturation problem. Actually, de Leeuw's definition is slightly different, viz

$$\sum_{k=1}^{n-1} l^*_{nk}(f)\, q_k .$$

He has shown, through a number of lemmas, that these operators possess the same saturation properties as those of B_n and we show here, as another application of the iteration method, that the same result holds for the operators K_n.

<u>LEMMA 3.1</u>. For each f in C

$$K_n(f;x) \longrightarrow f(x) \quad \text{as } n \longrightarrow \infty$$

uniformly in $0 \leqslant x \leqslant 1$.

<u>PROOF</u>. Let us compute $K_n(f_i;x)$, $f_i = x^i$, $i = 0,1,2$.

$$l^*_{nk}(f_0) = 1 , \quad k = 0(1)n ;$$

$$l^*_{nk}(f_1) = n \int_{-1/(2n)}^{+1/(2n)} (k/n + t)\, dt = k/n , \quad k = 0(1)n ;$$

$$l^*_{nk}(f_2) = \begin{cases} 0 , & k = 0 \\ (\frac{k}{n})^2 + \frac{1}{12n^2} , & k = 1(1)n-1 \\ 1 , & k = n . \end{cases}$$

Therefore,

$$K_n(f_0;x) = \sum_{k=0}^{n} q_k(n,x) = 1 ;$$

$$K_n(f_1;x) = \sum_{k=0}^{n} \frac{k}{n} q_k(n,x) = x ;$$

$$K_n(f_2;x) = \sum_{k=1}^{n-1} ((\frac{k}{n})^2 + \frac{1}{12n^2})\, q_k(n,x) + x^n$$

$$= B_n(f_2;x) + \frac{1}{12n^2} (1 - x^n - (1 - x)^n)$$

$$= x^2 + \frac{x(1-x)}{n} + \frac{1}{12n^2} (1 - x^n - (1 - x)^n)$$

$$\longrightarrow x^2 \text{ , uniformly, as } n \longrightarrow \infty ,$$

and the result follows on appealing to the theorem of Korovkin (1960) on the convergence of sequences of linear positive operators on C.

We turn next to the limiting semigroup of the iterates of the operators K_n.

LEMMA 3.2. For $t \geqslant 0$ and f in C,

$$\lim_{n \to \infty} \left\| K_n^{[nt]} f - \mathcal{B}_t f \right\| = 0 .$$

PROOF. For g in C^2,

$$K_n(g;x) - B_n(g;x) = \sum_{k=1}^{n-1} \left(n\int_{-1/(2n)}^{+1/(2n)} (g(k/n + t) - g(k/n))\, dt \right) q_k(n,x)$$

$$= \sum_{k=1}^{n-1} \left(n\int_{-1/(2n)}^{+1/(2n)} (g(k/n + t) - g(k/n) - g'(k/n)t)\, dt \right) q_k(n,x)$$

$$\left|K_n(g;x) - B_n(g;x)\right| \leqslant \frac{\|g''\|}{2} \left(n\int_{-1/(2n)}^{+1/(2n)} t^2\, dt \right) \sum_{k=1}^{n-1} q_k(n,x)$$

$$\leqslant \frac{\|g''\|}{24n^2} ,$$

whence

$$\left\| K_n g - B_n g \right\| = O(1/n^2) .$$

Making use of the identity

$$u^k - v^k = (u - v)\,(u^{k-1} + u^{k-2}v + \ldots + uv^{k-2} + v^{k-1})$$

and of the fact that both K_n and B_n have unit norm, we conclude that

$$\left\| K_n^k g - B_n^k g \right\| \leqslant k \left\| K_n g - B_n g \right\|$$

$$= O(k/n^2) .$$

For $k = [nt]$, $t \geqslant 0$, $n \longrightarrow \infty$, this gives

$$\lim_{n\to\infty} \left\| K_n^{[nt]} g - \mathcal{B}_t g \right\| = 0$$

for each g in C^2. But C^2 is dense in C and the result follows.

The next result shows that K_n and B_n have exactly the same saturation properties.

THEOREM 3.4. For f in C the following statements are equivalent:

(i) $f' \in \mathrm{Lip}_M 1 \,[0,1]$

(ii) $n\left|K_n(f;x) - f(x)\right| \leqslant \frac{M}{2} x(1-x) + o(1)$

where $o(1) \downarrow 0$ uniformly in $0 \leqslant x \leqslant 1$ as $n \longrightarrow \infty$.

Moreover,

$$K_n(f;x) - f(x) = o(1/n) \quad \text{iff} \quad f \in \mathcal{P}_1 .$$

PROOF. (i) $\Longrightarrow$ (ii). We follow the analysis in Lorentz(1966, p. 102) to obtain

$$\left| f(x) - f(y) - f'(x)(x-y)\right| \leqslant \frac{M}{2}(x-y)^2.$$

Let y be fixed but arbitrary. Being K_n a linear positive operator which preserves 1 and x, then

$$\left| f(x) - K_n(f;x)\right| \leqslant \frac{M}{2} K_n((x-y)^2;x)$$

$$\leqslant \frac{M}{2n}\left(x(1-x) + \frac{1}{12n} g_n(x)\right)$$

with

$$g_n(x) = 1 - x^n - (1-x)^n$$

and (ii) follows since $0 \leqslant g_n(x) \leqslant 1$ for all $n \geqslant 1$ and $0 \leqslant x \leqslant 1$.

(ii) $\Longrightarrow$ (i). Since

$$K_n(y(1-y);x) = (1 - 1/n)\, x(1-x) - \frac{1}{12n^2} g_n(x)$$

$$\leqslant \lambda_2\, x(1-x)\ , \quad \lambda_2 = 1 - 1/n\ ,$$

and, for r in $\mathbb{N}$,

$$K_n^r(f;x) - f(x) = \sum_{j=0}^{r-1} K_n^j(\,K_n f - f;x) \;,$$

then

$$\left| K_n^r(f;x) - f(x) \right| \leqslant \frac{M}{2n} \sum_{j=0}^{r-1} K_n^j(y(1-y) + o(1);x)$$

$$\leqslant \frac{M}{2}(1-\lambda_2^r)\; x(1-x) + \frac{Mr}{2n}\, o(1) \;.$$

Taking $r = [nt]$, $t \geqslant 0$, $n \longrightarrow \infty$, and using Lemma 3.2, gives

$$\left| \mathcal{B}_t(f;x) - f(x) \right| \leqslant \frac{Mt}{2}\, x(1-x) \;,$$

and (i) follows by Theorem 3.1.

The last assertion is equivalent to

$$\left| K_n(f;x) - f(x) \right| \leqslant \epsilon_n / n$$

with $\epsilon_n \downarrow 0$ as $n \longrightarrow \infty$, then $f \in \mathrm{Lip}\, \frac{1}{\epsilon_n}\, [0,1]$, and since $\epsilon_n > 0$ is arbitrary, f' is constant.

3.5. Approximation of smooth functions by polynomial operators of Micchelli's type.

LEMMA 3.3. Let f be defined and nonnegative on [0,1] then

$$B_n(\sqrt{f}\ ;\ x) \leqslant \sqrt{B_n(f;x)}\ .$$

PROOF.

$$B_n(\sqrt{f}\ ;\ x) = \sum_{k=0}^{n} \sqrt{f_k}\, q_k(n,x)\ , \qquad f_k = f(k/n)\ ,$$

$$= \sum_{k=0}^{n} (f_k\, q_k(n,x))^{\frac{1}{2}}\, (q_k(n,x))^{\frac{1}{2}}\ .$$

Employing the inequality of Cauchy-Schwarz we may write

$$B_n(\sqrt{f}\ ;\ x) \leqslant \Big(\sum_{k=0}^{n} f_k\, q_k(n,x)\Big)^{\frac{1}{2}} \Big(\sum_{k=0}^{n} q_k(n,x)\Big)^{\frac{1}{2}}$$

$$\leqslant (B_n(f;x))^{\frac{1}{2}}\ .$$

Let $\omega(f;\delta)$ be the modulus of continuity of $f \in C$, i.e.,

$$\omega(f;\delta) = \sup_{\substack{0 \leqslant x,y \leqslant 1 \\ |x-y| \leqslant \delta}} \Big| f(y) - f(x) \Big|\ , \qquad \delta > 0.$$

The subadditivity of $\omega(f;\delta)$ as a function of δ implies that

$$\omega(f;\lambda\delta) \leq (1+\lambda)\,\omega(f;\delta) \tag{3.3}$$

for all $\lambda, \delta > 0$ (see Lorentz (1966, p.44)).

Micchelli (1973) states without proof the following

LEMMA 3.4. For f in C and N in $\mathbb{N}$,

$$\left|(B_n - I)^N f(x)\right| \le \frac{3}{2}(2^N - 1)\,\omega(f;n^{-\frac{1}{2}}) .$$

PROOF. Since

$$(B_n - I)^N f(x) = f(x) + \sum_{m=1}^{N} (-1)^m \binom{N}{m} B_n^m(f;x) - (1-1)^N f(x),$$

then

$$\left|(B_n - I)^N f(x)\right| \le \sum_{m=1}^{N} \binom{N}{m} \left|(B_n^m - I) f(x)\right| .$$

For $x,y \in [0,1]$ and $\delta > 0$ we have

$$\left|f(y) - f(x)\right| \le \omega(f;|y-x|) = \omega(f;\frac{|y-x|}{\delta}\delta)$$

$$\le (1 + \frac{1}{\delta}|y-x|)\,\omega(f;\delta) , \qquad \text{by (3.3).}$$

In this inequality we assume that x is fixed but arbitrary. Since B_n^m is a positive operator which preserves constants, we obtain

$$\left|B_n^m(f(y);x) - f(x)\right| \le (1 + \frac{1}{\delta} B_n^m(|y-x|;x))\,\omega(f;\delta)$$

$$\le (1 + \frac{1}{2\delta n^{\frac{1}{2}}})\,\omega(f;\delta) ,$$

after observing that $|y-x| = +\sqrt{(y-x)^2}$ and using Lemma 3.3. Therefore,

$$\left|(B_n - I)^N f(x)\right| \le (\sum_{m=1}^{N} \binom{N}{m})\,(1 + \frac{1}{2\delta n^{\frac{1}{2}}})\,\omega(f;\delta)$$

and the result follows upon taking $\delta = n^{-\frac{1}{2}}$.

REMARK 3.1. For $N = 1$, Lemma 3.4 contains the Popoviciu's result that, for each f in C and n in $\mathbb{N}$,

$$\|B_n f - f\| \leqslant \frac{3}{2}\omega(f;n^{-\frac{1}{2}}) .$$

This inequality has been sharpened by Sikkema (1961) and Schurer and Steutel (1976,1977), who have determined the best possible constant for f in C and f in C^1 respectively.

Micchelli (1973) has introduced and studied the approximation properties of the operators

$$T_{n,N} = I - (I - B_n)^N \quad , \quad n,N\in \mathbb{N} .$$

Along the same lines, we introduce and study the operators

$$U_{n,N+1} = I + \Delta(B_n) ,$$

where

$$\Delta(B_n) = \prod_{k=0}^{N} (B_n - \lambda_k I) ,$$

$$\lambda_0 = \lambda_1 = 1, \; \lambda_k = 1 - \mu_k/n + o(n^{-1}) \; , \quad \mu_k = \binom{k}{2} ,$$

the λ_k's being the eigenvalues of the matrix A_{N+1} representing B_n acting on $\mathcal{P}_N$.

Unlike Micchelli's, our operator preserves $\mathcal{P}_N$. Indeed, for each $P(x) = X^T p$ in $\mathcal{P}_N$,

$$\Delta(B_n)\; P(x) = X^T \Delta(A_{N+1})\; p = 0$$

by Cayley-Hamilton theorem.

Having the fact that $\Delta(B_n) = (B_n - I)^{N+1} + o(1)$ and Lemma 3.4 in mind, it appears that $U_{n,N+1}f$ provides no better an approximation to any f in C than $B_n f$ itself. However, this is not the case for sufficiently smooth functions. Indeed, for f in C^{2N+2}, it follows from (1.2) that

$$n^{N+1}\ \Delta(B_n) \longrightarrow \prod_{k=0}^{N} (D + \mu_k I) = \Delta(D)$$

as $n \longrightarrow \infty$ and thus

$$\lim_{n \to \infty} n^{N+1} (U_{n,N+1} - I)\ f(x) = \Delta(D)f(x)\ .$$

Therefore, the order of approximation of $f \in C^{2N+2}$ by $U_{n,N+1}f$ is $O(1/n^{N+1})$, whereas that of f by $B_n f$ cannot be improved beyond $O(1/n)$, no matter how smooth f may be.

The above condition that $f \in C^{2N+2}$ may be slightly relaxed. To this end, let K denote the subset of C consisting of functions f such that $f, f', \ldots, f^{(2N+1)} \in C$ and $f^{(2N+1)} \in \text{Lip}\ 1$ on $[0,1]$.

THEOREM 3.5. For f in K,

$$n^{N+1}\ (U_{n,N+1}(f;x) - f(x)) = O(1)$$

uniformly in $0 \leqslant x \leqslant 1$.

PROOF. This follows from the observation that

$$n^{N+1} \Delta(B_n)\ f(x) = (n(B_n - I) + O(1))^{N+1}\ f(x)$$

and the fact that

$$n^k (B_n - I)^k\ f(x) = O(1)$$

for every k in $\mathbb{N}$, uniformly in $0 \leqslant x \leqslant 1$ (see Theorem 4.4 in Micchelli (1973)).

THEOREM 3.6. Let $f \in C$ and N be a nonnegative integer. If

$$n^{N+1} \left| U_{n,N+1}(f;x) - f(x) \right| \leqslant \frac{1}{2} Mx(1-x) + o(1)$$

uniformly in $0 \leqslant x \leqslant 1$, then $f, f', \dots, f^{(2N+1)} \in C$ and $(\prod_{k=1}^{N}(D + \mu_k I))f$ has a continuous extension to $[0,1]$ whose derivative is in Lip $\frac{1}{M}$.

<u>PROOF</u>. Since, for r in $\mathbb{N}$,

$$B_n^r - \lambda_k^r I = S_k(B_n - \lambda_k I) , \quad k = 0(1)N ,$$

with

$$S_k = S_k(n,r) = \sum_{j=0}^{r-1} \lambda_k^j B_n^{r-j-1} ,$$

then

$$\Delta(B_n^r) = S\,\Delta(B_n)$$

with

$$\Delta(B_n^r) = \prod_{k=0}^{N}(B_n^r - \lambda_k^r I) \quad \text{and} \quad S = \prod_{k=0}^{N} S_k .$$

Setting $v_2 = v_2(x) = x(x-1)$, then

$$S_2 v_2 = r \lambda_2^{r-1} v_2$$

$$S_k v_2 = \frac{\lambda_2^r - \lambda_k^r}{\lambda_2 - \lambda_k} v_2 , \quad k \neq 2 ,$$

$$S v_2 = r \lambda_2^{r-1} \prod_{\substack{k=0 \\ k\neq 2}}^{N} \left(\frac{\lambda_2^r - \lambda_k^r}{\lambda_2 - \lambda_k}\right) v_2$$

and

$$\left|\Delta(B_n^r) f(x)\right| \leq \left|S\,\Delta(B_n) f(x)\right|$$

$$\leq \frac{Mr \lambda_2^{r-1}}{2n} \prod_{\substack{k=0 \\ k\neq 2}}^{N} \left(\frac{\lambda_2^r - \lambda_k^r}{n(\lambda_2 - \lambda_k)}\right) x(1-x) + o(1/n^{N+1}) .$$

For $r = r_n = [nt]$, $t \geqslant 0$, $n \longrightarrow \infty$,

$$\frac{r\,\lambda_2^{r-1}}{n} \longrightarrow t\,e^{-t} = t - o(t) \quad \text{as} \quad t \downarrow 0$$

$$\frac{\lambda_2^r - \lambda_k^r}{n(\lambda_2 - \lambda_k)} \longrightarrow \frac{e^{-\mu_2 t} - e^{-\mu_k t}}{\mu_2 - \mu_k} = t + o(t) , \quad k \neq 2 ,$$

and

$$\left| \Delta(\mathcal{B}_t)\, f(x) \right| \leqslant \frac{Mt^{N+1}}{2} (x(1-x) + o(1))$$

with

$$\Delta(\mathcal{B}_t) = \prod_{k=0}^{N} (\mathcal{B}_t - e^{-\mu_k t} I) .$$

That f enjoys the differentiability properties stated in the theorem follows now from this inequality if we let $0 < a \leqslant x \leqslant b < 1$, define

$$g_t(x) = \frac{1}{t^{N+1}} \int_a^x \frac{\Delta(\mathcal{B}_t)\, f(s)}{\frac{1}{2}s(1-s)}\, ds ,$$

and follow the lines of the argument used in the proof of Theorem 4.5 in Micchelli (1973).

<u>COROLLARY.</u> Let f be a real-valued function defined on $[0,1]$ and N a nonnegative integer. If

$$n^{N+1} \left| U_{n,N+1}(f;x) - f(x) \right| = o(1)$$

uniformly in $0 \leqslant x \leqslant 1$, then f is a linear polynomial on $[0,1]$.

<u>PROOF.</u> For any $r \in \mathbb{N}$,

$$\left\| \Delta(B_n^r)\, f \right\| \leqslant \left\| S \right\| \left\| \Delta(B_n)\, f \right\| \leqslant r^{N+1} \epsilon_n \,/\, n^{N+1} ,$$

with $\varepsilon_n \downarrow 0$ as $n \longrightarrow \infty$. Choosing $r = r_n$ so that $r_n/n \longrightarrow \infty$ and $(r_n/n)\varepsilon_n \downarrow 0$,

$$\lim_{n \longrightarrow \infty} \Delta(B_n^r) f = (B_1 - I)^2 B_1^{N-1} f = 0 ,$$

from which it follows, by the idempotency of B_1, that $f = B_1 f$, i.e., f is a linear polynomial on $[0,1]$.

3.6. Iterates of multivariate Bernstein polynomials: Properties and applications.

3.6.1. The bivariate Bernstein operator B_n^* acting on $\mathcal{P}_{N,N}$.

The generation and approximation properties of N-dimensional Bernstein polynomials $B_{\underline{n}}(f;S_N;\underline{x})$ for f in $C[S_N]$ have been considered in Section 3 of Chapter 1. In this Subsection emphasis will be on the bivariate Bernstein operator $B_n^*P(x,y) = B_{n,n}(P;S_2;x,y)$ acting on $\mathcal{P}_{N,N}$ as most of the 2-dimensional results extend to any finite number of dimensions without essential difficulty.

The bivariate polynomial

$$P = P(x,y) = \sum_{i,j=0}^{N} p_{ij}\, x^i y^j \in \mathcal{P}_{N,N}$$

can be written in matrix notation simply as

$$P = X^T p\, Y ,$$

where $X^T = (1,x,x^2,\dots,x^N)$, $Y^T = (1,y,y^2,\dots,y^N)$, and $p = (p_{ij})$, $i,j=0(1)N$, is the $(N+1)x(N+1)$ coefficient matrix associated with P. In this compact notation we have the following

LEMMA 3.5. For P as above

$$B_n^* P = X^T A_{N+1}(n)\, p\, A_{N+1}^T(n)\, Y .$$

PROOF.

$$B_n^*(x^i y^j;x,y) = B_n(x^i;x).B_n(y^j;y)$$

$$= X^T A_{N+1}(n)\, e_i . Y^T A_{N+1}(n)\, e_j$$

$$B_n^*(x^i y^j;x,y) = X^T A_{N+1}(n)\, M_{ij}\, A_{N+1}^T(n)\, Y,$$

where $M_{ij} = e_i e_j^T$ is, of course, the $(N+1)x(N+1)$ zero matrix with a 1 in the (i,j) position. Owing to the linearity of B_n^*, the result follows on multiplying both sides of this equation by p_{ij} and summing over i and j:

$$B_n^* P = X^T A_{N+1}(n) \left(\sum_{i,j=0}^{N} p_{ij} M_{ij} \right) A_{N+1}^T(n)\, Y$$

$$= X^T A_{N+1}(n)\, p\, A_{N+1}^T(n)\, Y .$$

LEMMA 3.6. For P in $\mathscr{P}_{N,N}$

$$\lim_{n\to\infty} n(B_n^* P - P) = D^* P$$

with

$$D^* = \tfrac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2} + \tfrac{1}{2}y(1-y)\frac{\partial^2}{\partial y^2} .$$

PROOF. A simple computation reveals that

$$A_{N+1}(n)\, M_{ij}\, A_{N+1}^T(n) - M_{ij} = (A_{N+1}(n) - I)\, M_{ij}\, (A_{N+1}(n) - I)^T +$$

$$(A_{N+1}(n) - I)\, M_{ij} + M_{ij}(A_{N+1}(n) - I)^T$$

and that

$$\lim_{n\to\infty} n(A_{N+1}(n)\, M_{ij}\, A_{N+1}^T(n) - M_{ij}) = C_{N+1}\, M_{ij} + M_{ij}\, C_{N+1}^T . \tag{3.4}$$

Now

$$X^T C_{N+1}\, M_{ij}\, Y = y^j\, X^T C_{N+1}\, e_i$$

$$= y^j(\mu_i x^{i-1} - \mu_i x^i)$$

$$= \tfrac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}\,(x^i y^j) .$$

Multiplying both sides by p_{ij} and summing over i and j yiels

$$X^T C_{N+1}\ p\ Y = \tfrac{1}{2}x(1-x)\frac{\partial^2 p}{\partial x^2}.$$

In like manner,

$$\begin{aligned} X^T M_{ij} C^T_{N+1}\ Y &= Y^T C_{N+1} M_{ji}\ X \\ &= x^i\ Y^T C_{N+1}\ e_j \\ &= x^i(\mu_j y^{j-1} - \mu_j y^j) \\ &= \tfrac{1}{2}y(1-y)\frac{\partial^2}{\partial y^2}(x^i y^j)\ , \end{aligned}$$

and

$$X^T\ p\ C^T_{N+1}\ Y = \tfrac{1}{2}y(1-y)\frac{\partial^2 p}{\partial y^2}\ .$$

The result follows now readily on performing the foregoing operations on the left side of (3.4) and appealing to Lemma 3.5.

REMARK 3.2. Owing to the linearity of the operators B_n^* and D^* and to the facts that $\|B_n^*\| = 1$ and $\bigcup_{N=0}^{\infty} \mathcal{P}_{N,N}$ is dense in $C^2[S_2]$,

$$\lim_{n\to\infty} n(B_n^*\ f(x,y) - f(x,y)) = D^* f(x,y) \tag{3.5}$$

uniformly in S_2 .

This should be confronted with the result of Stancu (1963 b),1964) that

$$B_n^*\ f(x,y) - f(x,y) = \frac{x(1-x)}{2n} f''_{x^2}(\xi,y) + \frac{y(1-y)}{2n} f''_{y^2}(x,\eta) + \frac{x(1-x)\ y(1-y)}{4n^2} f^{(iv)}_{x^2y^2}(\xi,\eta)\ , \qquad \xi,\eta\in(0,1)\ .$$

3.6.2. Iteration of B_n^* and the limiting semigroup $\{\mathcal{B}_t^* \; ; \; t \geq 0\}$.

Using Lemma 3.5 and the results obtained in Section 4.3 of Chapter 3 on the matrix $A_{N+1}^{r_n}(n)$, it follows that

$$\mathcal{B}_t^* \equiv \lim_{n \to \infty} (B_n^*)^{r_n}$$

exists as a linear positive contraction operator on $\mathcal{P}_{N,N}$ to itself iff $r_n/n \to t \geq 0$ as $n \to \infty$.

For any P in $\mathcal{P}_{N,N}$ let p be its coefficient matrix. Then

$$\mathcal{B}_t^* P = X^T . \lim_{n \to \infty} \left[A_{N+1}^{r_n}(n) \; p \; (A_{N+1}^{r_n}(n))^T \right] . \; Y$$

$$= X^T \left[\exp(tC_{N+1}) \; p \; (\exp(tC_{N+1}))^T \right] Y \; . \qquad (3.6)$$

If t=0 then, clearly,

$$\mathcal{B}_0^* P = P \; , \quad \text{i.e.,} \; \mathcal{B}_0^* = I \; . \qquad (3.7)$$

If $t \to \infty$ then

$$\lim_{n \to \infty} A_{N+1}^{r_n}(n) \; p \; (A_{N+1}^{r_n}(n))^T = A_{N+1}^{\infty} \; p \; (A_{N+1}^{\infty})^T$$

$$= \left[\begin{array}{cc:c} p_{00} & \sum_{j=1}^{N} p_{0j} & \\ \sum_{i=1}^{N} p_{i0} & \sum_{i,j=1}^{N} p_{ij} & \mathbf{O} \\ \hdashline & \mathbf{O} & \end{array} \right]$$

$$= \left[\begin{array}{cc:c} P(0,0) & P(0,1)-P(0,0) & \\ P(1,0)-P(0,0) & P(1,1)+P(0,0)-P(0,1)-P(1,0) & \mathbf{O} \\ \hdashline & \mathbf{O} & \end{array} \right]$$

and

$$\mathcal{B}^*_{\infty} P(x,y) = \lim_{n\to\infty} (B^*_n)^{r_n} P(x,y)$$

$$= P(0,0) + \Big[P(1,0)-P(0,0)\Big]x + \Big\{P(0,1)-P(0,0) +$$

$$\Big[P(1,1)+P(0,0)-P(0,1)-P(1,0)\Big]x\Big\}y$$

$$= P(0,0)(1-x)(1-y) + P(0,1)(1-x)y + P(1,0)x(1-y) + P(1,1)xy$$

$$= B^*_1 P(x,y) \ ,$$

the bilinear polynomial interpolating P at the four corners of the unit square.

It follows from (3.6) that

$$\mathcal{B}^*_{s+t} P = X^T\Big\{e^{(s+t)C_{N+1}} \ p \Big[e^{(s+t)C_{N+1}}\Big]^T\Big\} Y$$

$$= X^T\Big\{e^{sC_{N+1}}\Big[e^{tC_{N+1}} \ p \ (e^{tC_{N+1}})^T\Big](e^{sC_{N+1}})^T\Big\} Y \ .$$

We have, in the right brackets, the coefficient matrix of $\mathcal{B}^*_t P$ and, in braces, that of $\mathcal{B}^*_s(\mathcal{B}^*_t P)$. Therefore,

$$\mathcal{B}^*_{s+t} P = \mathcal{B}^*_s \mathcal{B}^*_t P \ . \tag{3.8}$$

It is also easy to see that

$$\lim_{t\downarrow 0} \frac{1}{t}\Big[e^{tC_{N+1}} \ p \ (e^{tC_{N+1}})^T - p\Big] = C_{N+1} \ p + p \ C^T_{N+1}$$

and

$$\lim_{t\downarrow 0} \frac{1}{t}(\mathcal{B}^*_t - I) \ P = X^T(C_{N+1} \ p + p \ C^T_{N+1})Y$$

$$= D^* P \ . \tag{3.9}$$

With the results (3.5) - (3.9) and Remark 3.2 in mind, we may assert the existence of a totally positive, strongly continuous, contraction semigroup $\{\mathcal{B}_t^* \ ; \ t \geqslant 0\}$ on $C[S_2]$ with infinitesimal generator D^* acting on $C^2[S_2]$ and such that, for any f in $C[S_2]$ and all $t \geqslant 0$,

$$\lim_{n \to \infty} (B_n^*)^{[nt]} f = \mathcal{B}_t^* f .$$

We now proceed to determine an explicit representation for $\mathcal{B}_t^* f$, $f \in C[S_2]$, by working on the lines set out in Subsection 1.4 and omitting the details.

Introducing the notation

$$W_{km}(t,x,y) = \mathcal{B}_t^* (x^k y^m;x,y)$$

we are led to the Cauchy problem

$$\begin{cases} \dfrac{\partial}{\partial t} W_{km}(t,x,y) = D^* W_{km}(t,x,y) \\ W_{km}(0,x,y) = x^k y^m \end{cases} \qquad k,m=0(1)N ,$$

whose solution is

$$W_{km}(t,x,y) = X^T e^{tC_{N+1}} M_{km} (e^{tC_{N+1}})^T Y$$

$$= (\sum_{i=0}^{k} b_{ik}(t)\, x^i) (\sum_{j=0}^{m} b_{jm}(t)\, y^j) .$$

Now let

$$g_{km} = g_{km}(x,y) = u_k(x).u_m(y) , \quad k,m=0(1)N . \tag{3.10}$$

Then

$$\mathcal{B}_t^*(g_{km};x,y) = \mathcal{B}_t(u_k;x).\mathcal{B}_t(u_m;y)$$

$$= e^{-\mu_k t} u_k(x) . e^{-\mu_m t} u_m(y)$$

$$= e^{-(\mu_k+\mu_m)t} g_{km}(x,y) \ , \qquad (3.11)$$

and

$$D^* g_{km} = -(\mu_k+\mu_m) g_{km} \ .$$

Setting

$$W(t,x,y) = \mathscr{B}_t^*(f;x,y) \ , \quad f \in \mathscr{D}(D^*) \ ,$$

we find the Cauchy problem

$$\begin{cases} \frac{\partial}{\partial t} W(t,x,y) = D^* W(t,x,y) \\ W(0,x,y) = f(x,y) \ , \end{cases}$$

which can be solved in much the same way as in the 1-dimensional case. We have, successively:

$$W(t,x,y) = \sum_{k,m=0}^{\infty} c_{km}(t) \, g_{km}(x,y)$$

$$= \sum_{k,m=0}^{\infty} \bar{c}_{km} \, e^{-(\mu_k+\mu_m)t} g_{km}(x,y) \ ;$$

$$f(x,y) = \sum_{k,m=0}^{\infty} \bar{c}_{km} \, g_{km}(x,y)$$

$$= \bar{c}_{00} + \bar{c}_{10}x + \bar{c}_{01}y + \bar{c}_{11}xy + \sum_{k,m=2}^{\infty} \bar{c}_{km} \, g_{km}(x,y)$$

$$= f(0,0)(1-x)(1-y) + f(1,0)x(1-y) + f(0,1)y(1-x) + f(1,1)xy$$

$$+ \sum_{k,m=2}^{\infty} \bar{c}_{km} \, g_{km}(x,y)$$

$$= B_1^* f(x,y) + \sum_{k,m=2}^{\infty} \bar{c}_{km} \, g_{km}(x,y) \ ;$$

$$\tilde{f}(\xi,\eta) = f(\xi,\eta) - B_1^*(f;\xi,\eta)$$

$$= \sum_{k,m=2}^{\infty} \bar{c}_{km}\, g_{km}(\xi,\eta) \;;$$

$$\bar{c}_{km} = \frac{1}{h_k h_m} \int_{S_2} \tilde{f}(\xi,\eta) \frac{u_k(\xi)}{\xi(1-\xi)} \frac{u_m(\eta)}{\eta(1-\eta)}\, d\xi\, d\eta \,, \qquad k,m \geqslant 2 \;;$$

$$\mathcal{B}_t^*(f;x,y) = B_1^*(f;x,y) + \int_{S_2} G(t;x,\xi)\, G(t;y,\eta)\, \tilde{f}(\xi,\eta)\, d\xi\, d\eta \,.$$

The analysis carried out above can be easily extended to the N-dimensional case yielding the following

<u>THEOREM 3.7</u>. There exists a semigroup $\{\mathcal{B}_t^* \;;\; t \geqslant 0\}$ of class (C_0) on $C[S_N]$ such that

$$\lim_{n\to\infty} (B_n^*)^{r_n} f = \mathcal{B}_t^* f$$

iff $\lim_{n\to\infty} r_n/n \to t \geqslant 0$. The semigroup $\{\mathcal{B}_t^*\}$ is totally positive, contractive, generated by the linear differential operator

$$D^* \equiv \sum_{i=1}^{N} \tfrac{1}{2} x(1-x) \frac{\partial^2}{\partial x_i^2} \,,$$

which is such that

$$\lim_{n\to\infty} n(B_n^* f(\underline{x}) - f(\underline{x})) = D^* f(\underline{x}) \tag{3.12}$$

uniformly in $\underline{x}$ in S_N provided that $f \in C^2[S_N]$, and given by

$$\mathcal{B}_t^* (f;\underline{x}) = B_1^* (f;\underline{x}) + \int_{S_N} K(t;\underline{x},\underline{y})\, (f(\underline{y}) - B_1^*(f;\underline{y}))\, d\underline{y}$$

with

$$K(t;\underline{x},\underline{y}) = \prod_{i=1}^{N} G(t;x_i,y_i) ,$$

and

$$G(t;x_i,y_i) = \frac{1}{y_i(1-y_i)} \sum_{k=2}^{\infty} \frac{e^{-\mu_k t}}{h_k} u_k(x_i)\, u_k(y_i) , \qquad i=1(1)N,$$

with the u's and h's as in Subsection 1.4. $B_1^*(f;\underline{x})$ is the multilinear polynomial interpolating $f(\underline{x})$ at the vertices of S_N and is such that

$$\lim_{n\to\infty} (B_n^*)^{r_n} f = B_1^* f \quad \text{iff} \quad \lim_{n\to\infty} r_n/n = \infty .$$

CORROLLARY 1. For f in $C[S_N]$

$$B_n^* f(\underline{x}) - f(\underline{x}) = o(1/n) \Longrightarrow f = B_1^* f .$$

PROOF. The left side of the implication means that $B_n^* f - f = \epsilon_n/n$ with $\epsilon_n \downarrow 0$ as $n \to \infty$. Let $\{r_n\}$ be a sequence of nonnegative integers such that $r_n/n \to \infty$ and $r_n \epsilon_n/n \downarrow 0$ as $n \to \infty$. Then

$$\| B_1^* f - f \| = \lim_{n\to\infty} \| (B_n^*)^{r_n} f - f \| \leqslant \lim_{n\to\infty} r_n \| B_n^* f - f \| = 0.$$

In other words, the trivial class of B_n^* is $\mathcal{P}_{1,1,\ldots,1}$, the subspace of N-dimensional polynomials linear in each variable.

As an immediate consequence of the fact that B_n^* is a positive operator which leaves $\mathcal{P}_{1,1,\ldots,1}$ invariant we have the following

COROLLARY 2. If f in $C[S_N]$ is convex on S_N then

(i) $B_n^* f(\underline{x}) \geqslant f(\underline{x}), \quad n \geqslant 1, \quad \underline{x} \in S_N$;

(ii) $\mathcal{B}_t^* f(\underline{x}) \geqslant f(\underline{x}), \quad t \geqslant 0, \quad \underline{x} \in S_N$.

PROOF. Let $\underline{y}$ in S_N be fixed but arbitrary then there are real constants $c_1,\ldots,c_N$ such that

$$f(\underline{x}) \geqslant f(\underline{y}) + \sum_{i=1}^{N} c_i(x_i - y_i) , \quad \underline{x},\underline{y} \in S_N .$$

Applying B_n^* to both sides of this inequality gives

$$B_n^*(f(\underline{x});\underline{y}) \geqslant f(\underline{y}), \quad \text{all } \underline{y} \text{ in } S_N,$$

and this, under iteration, yields (ii).

REMARK 3.3. If $D^*f = 0$ on some subset F of S_N then, clearly,

$$B_n^*(f;\underline{x}) - f(\underline{x}) = o(1/n)$$

and it is interesting to note the following consequences of this fact:

(i) The only solution of $D^*f = 0$ with continuous second derivatives on S_N has the form $B_1^* f$.

(ii) The local saturation class theorem that, for each g(x) in C and $0 < a < x < b < 1$,

$$B_n(g;x) - g(x) = o(1/n) \Longrightarrow g = B_1 g \quad \text{on } [a,b]$$

(see de Leeuw (1959) and Bajsanski and Bojanic (1964)) is not true for B_n^*; that is,

$$B_n^*(f;\underline{x}) - f(\underline{x}) = o(1/n) \not\Longrightarrow f = B_1^* f \quad \text{on } F ,$$

whenever F is a closed subset of S_N. Indeed, if $g_1 = g_1(x)$ and $g_2 = g_2(y)$ are some nonlinear twice continuously differentiable functions satisfying

$$D(g_1;x) = g_1(x) \quad \text{on} \quad 0 < a \leqslant x \leqslant b < 1$$

and

$$D(g_2;y) = -g_2(y) \quad \text{on} \quad 0 < c \leqslant y \leqslant d < 1 \ ,$$

then

$$f = f(x,y) = g_1(x).g_2(y) \neq B_1^* f$$

and yet

$$D^* f = 0 \quad \text{on} \quad F = [a,b] \times [c,d] \ .$$

<u>COROLLARY 3</u>. For f in $C^2[S_2]$ the following statements are equivalent:

$$\text{(i)} \quad \left| D^* f(\underline{x}) \right| \leqslant M \ , \quad \underline{x} \in S_N \ ;$$

$$\text{(ii)} \quad n \left| B_n^* f(\underline{x}) - f(\underline{x}) \right| \leqslant M + o(1) \ , \quad \underline{x} \in S_N \ ;$$

$$\text{(iii)} \quad \left| \mathcal{B}_t^* f(\underline{x}) - f(\underline{x}) \right| \leqslant Mt \ , \quad \underline{x} \in S_N \ , \quad t \geqslant 0 \ .$$

<u>PROOF</u>. (i)$\Longrightarrow$(ii). Immediate from (3.12).

(ii)$\Longrightarrow$(iii). For $r \in \mathbb{N}$,

$$\left| (B_n^*)^r f(\underline{x}) - f(\underline{x}) \right| \leqslant r \left| B_n^* f(\underline{x}) - f(\underline{x}) \right| \leqslant Mr/n + o(1/n) \ ,$$

and (iii) follows upon taking $r = r_n = [nt]$, $t \geqslant 0$, $n \longrightarrow \infty$.

(iii)$\Longrightarrow$(i). Immediate from the fact that $D^* f = \lim_{t \downarrow 0} (\mathcal{B}_t^* f - f)/t$.

3.6.3. Linear operators commuting with bivariate Bernstein polynomials.

Let T be a linear operator mapping $C[S_2]$ into itself and commuting with B_n^* :

$$TB_n^* = B_n^*T .$$

Owing to the density of the space of bivariate polynomials in $C[S_2]$, it suffices to require this to hold for polynomials.

Let $f \in C[S_2]$ and (θ,ζ) be any of the points (0,0),(0,1),(1,0),(1,1). We show first that

$$B_1^* B_n^* = B_n^* B_1^* .$$

Indeed,

$$B_n^* B_1^* f(x,y) = B_1^* f(x,y)$$

as $\mathcal{P}_{1,1}$ is left invariant under B_n^*, and

$$B_1^* B_n^* f(x,y) = B_n^* f(0,0)(1-x)(1-y) + B_n^* f(1,0)x(1-y) + B_n^* f(0,1)(1-x)y + B_n^* f(1,1)xy$$

$$= B_1^* f(x,y) ,$$

since $B_n^* f(\theta,\zeta) = f(\theta,\zeta)$.

It is now easily seen that the linear operator

$$W = T(I - B_1^*)$$

has the following properties:

$$WB_n^* = B_n^*W \tag{3.13}$$

and

$$Wf(\theta,\zeta) = 0 \ , \tag{3.14}$$

which imply that W anihilates $\mathcal{P}_{1,1}$. Indeed, if $g \in \mathcal{P}_{1,1}$, then $Wg = WB_n^* g = B_n^* Wg$, i.e., Wg is left invariant under B_n^* and thus $Wg \in \mathcal{P}_{1,1}$. Property (3.14) now implies that $Wg = 0$ and,since g is arbitrary in $\mathcal{P}_{1,1}$, $W\mathcal{P}_{1,1} = 0$.

Iteration of B_n^* in (3.13) leads to

$$W\mathcal{B}_t^* = \mathcal{B}_t^* W \ , \qquad t \geqslant 0 \ ,$$

and application of W to both sides of

$$\mathcal{B}_t^* g_{ij} = e^{-(\mu_i+\mu_j)t} g_{ij} \ ,$$

where $g_{ij} = g_{ij}(x,y) = u_i(x).u_j(y)$ are, for $i,j \geqslant 2$, the only common eigenfunctions of $\mathcal{B}_t^*$ corresponding to the eigenvalues $\exp(-(\mu_i+\mu_j)t)$ (see (3.10) and 3.11)), gives

$$\mathcal{B}_t^*(Wg_{ij}) = e^{-(\mu_i+\mu_j)t}(Wg_{ij}) \ ,$$

whence

$$Wg_{ij} = c_{ij} g_{ij} \ , \qquad i,j \geqslant 2 \ ,$$

for some constants c_{ij}. Recalling from Subsection 1.3 that

$$u_i(1-x) = (-1)^i u_i(x) \ , \qquad i \geqslant 2 \ ,$$

we find

$$B_3(u_i;x) = 3u_i(1/3)x(1-x)^2 + 3u_i(2/3)x^2(1-x)$$

$$= \begin{cases} -3u_i(1/3)u_2(x) \ , & \text{if } i \text{ even} \ , \\ 6u_i(1/3)u_3(x) \ , & \text{if } i \text{ odd} \ , \end{cases}$$

and

$$B_3^* W g_{ij}(x,y) = c_{ij}\, B_3(u_i;x).B_3(u_j;y)$$

$$= \begin{cases} 9g_{ij}(1/3,1/3)c_{ij}g_{22}(x,y)\ , & \text{if } i \text{ even },\ j \text{ even}, \\ -18g_{ij}(1/3,1/3)c_{ij}g_{23}(x,y)\ , & \text{if } i \text{ even },\ j \text{ odd }, \\ -18g_{ij}(1/3,1/3)c_{ij}g_{32}(x,y)\ , & \text{if } i \text{ odd },\ j \text{ even}, \\ 36g_{ij}(1/3,1/3)c_{ij}g_{33}(x,y)\ , & \text{if } i \text{ odd },\ j \text{ odd }, \end{cases}$$

On the other hand

$$W\, B_3^* g_{ij}(x,y) = \begin{cases} 9g_{ij}(1/3,1/3)c_{22}g_{22}(x,y)\ , & \text{if } i \text{ even },\ j \text{ even}, \\ -18g_{ij}(1/3,1/3)c_{23}g_{23}(x,y)\ , & \text{if } i \text{ even },\ j \text{ odd }, \\ -18g_{ij}(1/3,1/3)c_{32}g_{32}(x,y)\ , & \text{if } i \text{ odd },\ j \text{ even}, \\ 36g_{ij}(1/3,1/3)c_{33}g_{33}(x,y)\ , & \text{if } i \text{ odd },\ j \text{ odd }. \end{cases}$$

Use of classical properties of the shifted Jacobi polynomials of parameters (1,1), $P_n^*(x)$, shows that $P_n^*(1/3) \neq 0$ for all n. From this we infer that $g_{ij}(1/3,1/3) \neq 0$ for all i and j, and therefore

$$c_{ij} = \begin{cases} c_{22}\ , & \text{if } i \text{ even, } j \text{ even}, \\ c_{23}\ , & \text{if } i \text{ even, } j \text{ odd }, \\ c_{32}\ , & \text{if } i \text{ odd },\ j \text{ even}, \\ c_{33}\ , & \text{if } i \text{ odd },\ j \text{ odd }. \end{cases}$$

Let f(x,y) be any polynomial. We may express it in the form

$$f(x,y) = B_1^*\, f(x,y) + \sum_{i,j \geqslant 2} a_{ij}\, g_{ij}(x,y)\ .$$

Then we have

$$\tilde{f}(x,y) \equiv f(x,y) - B_1^* f(x,y) = \sum_{i,j\geq 2} a_{ij}\, g_{ij}(x,y)$$

and

$$Wf(x,y) = W\tilde{f}(x,y) = \sum_{i,j\geq 2} \alpha_{ij} W g_{ij}(x,y)$$

$$= c_{22}\, \Sigma_{00} + c_{23}\, \Sigma_{01} + c_{32}\, \Sigma_{10} + c_{33}\, \Sigma_{11} \;,$$

Σ_{00}, Σ_{01}, Σ_{10}, and Σ_{11} standing for the summations over i and j even, i even and j odd, i odd and j even, and i and j odd respectively.

Owing to the symmetry properties of the basic functions g_{ij}, these summations are given by

$$\Sigma_{00} = \frac{1}{4}\left\{\tilde{f}(x,y) + \tilde{f}(1-x,y) + \tilde{f}(x,1-y) + \tilde{f}(1-x,1-y)\right\}$$

$$\Sigma_{01} = \frac{1}{4}\left\{\tilde{f}(x,y) + \tilde{f}(1-x,y) - \tilde{f}(x,1-y) - \tilde{f}(1-x,1-y)\right\}$$

$$\Sigma_{10} = \frac{1}{4}\left\{\tilde{f}(x,y) - \tilde{f}(1-x,y) + \tilde{f}(x,1-y) - \tilde{f}(1-x,1-y)\right\}$$

$$\Sigma_{11} = \frac{1}{4}\left\{\tilde{f}(x,y) - \tilde{f}(1-x,y) - \tilde{f}(x,1-y) + \tilde{f}(1-x,1-y)\right\} .$$

Therefore,

$$Wf(x,y) = C_{00}(1-x)(1-y) + C_{01}(1-x)y + C_{10}x(1-y) + C_{11}xy +$$
$$\overline{C}_{00}f(x,y) + \overline{C}_{01}f(1-x,y) + \overline{C}_{10}f(x,1-y) + \overline{C}_{11}f(1-x,1-y) \;,$$

where C_{00}, C_{01}, C_{10}, C_{11} are linear functionals on f and $\overline{C}_{00}$, $\overline{C}_{01}$, $\overline{C}_{10}$, $\overline{C}_{11}$ constants depending on B_n^* .

CHAPTER 5

ADDENDUM

BERNSTEIN TYPE APPROXIMATION ON $C[0,\infty]$

Let $C[0,\infty]$ denote the subspace of $C[0,\infty)$ consisting of continuous real-valued functions f on $[0,\infty)$ for which $\lim_{t\to\infty} f(t)$ exists.

It is well known that $C[0,\infty]$ is a separable Banach space normed by

$$\|f\| = \sup_{0\leqslant t\leqslant\infty} |f(t)|$$

and spanned by $\{e^{-nt};\ n=0,1,2,\ldots\}$. We also note that the transformation $x = e^{-t}$ maps $C[0,\infty]$ on $C[0,1]$.

An approximation process on $C[0,\infty]$ closely related to the Bernstein construction is the following.

In correspondence with a given $f\in C[0,\infty]$, exhibiting at most polynomial growth at ∞, let us consider the following sequence of operators

$$S_n(f;x) = e^{-nx}\sum_{k=0}^{\infty}\frac{(nx)^k}{k!}f(k/n)\ , \tag{1}$$

commonly referred to as the Szász operators. An analysis paralleling the one carried out in Section 1.2 of Chapter 1 to express $B_n(f;x)$ in terms of finite differences shows that

$$S_n(f;x) = \exp(nx\Delta_{1/n})\, f(0) \tag{2}$$

$$= \sum_{k=0}^{\infty} \frac{(nx)^k}{k!} \Delta^k_{1/n}\, f(0)\ , \tag{3}$$

a replica of Taylor's expansion except that derivatives are replaced by differences.

The result of Szász that

$$S_n(f;x) \longrightarrow f(x) \quad \text{as} \quad n \longrightarrow \infty \tag{4}$$

uniformly in every finite interval $0 \leqslant x \leqslant a$ (see, e.g., Cheney and Sharma (1964)) is an immediate consequence of the so-called first exponential formula of semigroup theory (see, e.g., Hille and Phillips (1974, p.302),

$$\lim_{h \downarrow 0} \left\| T_t f - \exp(tD_h)f \right\| = 0, \quad D_h \equiv (T_h - I)/h\ ,$$

applied to the semigroup $\{T_t;\ t \geqslant 0\}$ of translations in $C[0,\infty]$, i.e.,

$$T_t f(x) \equiv f(x+t) = \lim_{h \downarrow 0} \exp(tD_h) f(x)\ , \tag{5}$$

the limit existing uniformly with respect to x in $[0,\infty]$ and with respect to t in every finite interval $[0,\beta]$. In fact, taking $h = 1/n$ and $x = 0$ it follows from (5) that

$$f(t) = \lim_{n \to \infty} \exp(tn\Delta_{1/n})\, f(0)$$

and (4) follows from this and (2).

As we shall see, most of the Bernstein approximation properties pass on to the Szász operators. Namely,

a) S_n is again an interpolation operator in the sense that the values of the argument function at a certain finite number of points determine the result of operating on that function.

b) The operator S_n is linear and positive as follows at once from the definition (1).

c) That S_n maps $\mathcal{P}_N$, $n \geq N$, onto itself and leaves $\mathcal{P}_1$ invariant follows immediately from the representation (3). Also,

$$S_n(t^2;x) = x^2 + \frac{x}{n} \quad . \tag{6}$$

Regarded as a linear operator in $\mathcal{P}_N$, the matrix $\mathcal{A}_{N+1}(n)$ representing S_n when we take for $\mathcal{P}_N$ the basis $\{x^k\}_{k=0}^N$ may be obtained as follows.

$$\begin{aligned} S_n(t^j;x) &= \sum_{i=0}^{\infty} \frac{n^i}{i!} \Delta^i_{1/n} t^j \Big|_{t=0} x^i \\ &= \sum_{i=0}^{j} \sigma_{ij}\, n^{i-j}\, x^i \\ &= \sum_{i=0}^{j} \frac{a_{ij}}{\lambda_i} x^i , \end{aligned}$$

showing that

$$\mathcal{A}_{N+1}(n) = \Lambda^{-1}_{N+1}(n)\, A_{N+1}(n) , \tag{7}$$

where $A_{N+1}(n) = (a_{ij})$, $0 \leq i \leq j \leq N \leq n$, is the $(N+1)x(N+1)$ matrix representation of B_n acting on $\mathcal{P}_N$ and $\Lambda_{N+1}(n) = \text{diag}(\lambda_i)$, $i=0(1)N$, with $\lambda_i = a_{ii}$.

Being the product of two TP matrices,

d) $\mathcal{A}_{N+1}(n)$ is TP and therefore S_n is variation diminishing. As a result, all those shape preserving properties we have studied for the

Bernstein polynomials carry over to the Szász operators.

We now take up the task of iterating S_n proceeding in much the same way as when dealing with B_n. Since, by Lemma 4.1 a) of Chapter 3,

$$A_{N+1}(n) = I + \frac{1}{n} C_{N+1} + O(1/n^2) ,$$

(7) gives

$$\mathcal{A}_{N+1}(n) = \Lambda_{N+1}^{-1}(n) + \frac{1}{n} \Lambda_{N+1}^{-1}(n)\, C_{N+1} + O(1/n^2)$$

and

$$\lim_{n\to\infty} n(\mathcal{A}_{N+1}(n) - I) = \lim_{n\to\infty} n(\Lambda_{N+1}^{-1}(n) - I) + \lim_{n\to\infty} \Lambda_{N+1}^{-1}(n)\, C_{N+1} \qquad (8)$$

$$= \operatorname{diag}(\mu_j) + C_{N+1} , \qquad j=0(1)N , \qquad (9)$$

$$= \begin{bmatrix} 0 & \mu_1 & & & \mathbf{0} \\ & 0 & \mu_2 & & \\ & & \ddots & \ddots & \\ & & & & \mu_N \\ & \mathbf{0} & & & 0 \end{bmatrix} .$$

Let $\{r_n\}$ be a sequence of nonnegative reals such that $r_n/n \to t$ as $n \to \infty$. Let $(s_{ik}(t))$, $0 \leqslant i \leqslant k \leqslant N$, be the matrix representing the limiting operator

$$\mathcal{S}_t = \lim_{n\to\infty} S_n^{r_n}$$

acting on $\mathcal{P}_N$. Clearly,

$$\lim_{n\to\infty} \mathcal{A}_{N+1}^{r_n}(n) = \lim_{n\to\infty} \Lambda_{N+1}^{-r_n}(n) . \lim_{n\to\infty} A_{N+1}^{r_n}(n)$$

$$= \operatorname{diag}(e^{\mu_i t}) . e^{tC_{N+1}} , \qquad i=0(1)N ,$$

giving the following explicit representation for the entries of the limiting matrix

$$s_{ik}(t) = \sum_{j=i}^{k} \beta_{i,j,k}\, e^{-(\mu_j-\mu_i)t}$$

with $\beta_{i,j,k}$ given by (4.4) of Chapter 3.

An analysis paralleling the one carried out for the semigroup $\{\mathcal{B}_t; t \geqslant 0\}$ shows the existence of a totally positive semigroup $\{\mathcal{S}_t;\ t \geqslant 0\}$ of class (C_0) on $C[0,\infty]$ given by

$$\mathcal{S}_t f = \lim_{n \to \infty} S_n^{[nt]} f$$

and generated by the linear differential operator $\frac{1}{2} x\, d^2/dx^2$ with domain $C^2[0,\infty]$.

As a last application of the iteration method we give the saturation theory for the Szász operators.

<u>THEOREM.</u> Let $f \in C[0,\infty]$ exhibit at most polynomial growth at ∞ then the following statements are equivalent:

(i) $f \in \mathrm{Lip}\,\frac{1}{M}[0,\infty)$;

(ii) $\left| S_n(f;x) - f(x) \right| \leqslant \frac{Mx}{2n}$, $\quad n \geqslant 1$, $\quad x \geqslant 0$;

(iii) $\left| \mathcal{S}_t(f;x) - f(x) \right| \leqslant \frac{Mx}{2}\, t$, $t \geqslant 0$, $\quad x \geqslant 0$.

Moreover,

$$S_n(f;x) - f(x) = o(x/n) \iff f \in \mathcal{P}_1 .$$

PROOF. That (i) $\Longrightarrow$ (ii) follows immediately from the inequality

$$\left| f(x) - f(y) - f'(x)(x-y) \right| \leqslant \frac{M}{2}(x-y)^2 ,$$

the positivity of S_n, and (6).

That (ii) $\Longrightarrow$ (iii) is an immediate consequence of the positivity of S_n and of the invariance of x under S_n. Indeed, since

$$S_n^k(f;x) - f(x) = \sum_{j=0}^{k-1} S_n^j(S_n f - f;x) ,$$

then

$$\left| S_n^k(f;x) - f(x) \right| \leqslant \sum_{j=0}^{k-1} S_n^j(\left| S_n f - f \right|;x) \leqslant \frac{Mk}{2n} x$$

and (iii) follows upon taking $k = [nt]$, $t \geqslant 0$, $n \longrightarrow \infty$.

Finally, we show that (iii) $\Longrightarrow$ (i) by showing that it is true on every closed interval $[a,b] \subset [0,\infty)$. This follows upon letting $0 < a \leqslant x \leqslant b < \infty$, defining

$$g_t(x) = \frac{1}{t} \int_a^x \frac{(\mathcal{S}_t - I) f(s)}{\frac{s}{2}} \, ds ,$$

and following the lines of the argument used in the proof that (iii) $\Longrightarrow$ (i) in Theorem 3.2 in Micchelli (1973).

The last assertion follows at once from the equivalence of (i) and (ii).

REFERENCES

1. B. BAJSANSKY and R. BOJANIC (1964), A note on approximation by Bernstein polynomials, Bulletin Amer. Math. Soc. 70, 675-677.

2. R. E. BARNHILL and R. F. RIESENFELD (1974), Computer aided geometric design, Academic Press, N. Y.

3. H. BERENS (1972), Pointwise saturation of positive operators, J. Approx. Theory 6, 135-146.

4. P. BÉZIER (1972), Numerical control-mathematics and applications, transl. by A. R. Forrest and A. F. Pankhurst, Wiley, N. Y.

5. R. G.-BONNE and P. SABLONNIÈRE (1976), Comparaison des formes de courbes paramétrées et de leurs approximants-spline, Publication 76 de l'Université de Lille.

6. R. G.-BONNE and P. SABLONNIÈRE (1977), Propriétés de diminution de la variation et de la forme pour approximants-spline, Seminaire 263 d'Analyse Numerique de Lille.

7. P. L. BUTZER and H. BERENS (1967), Semigroups of operators and approximation, Springer-Verlag, N. Y.

8. P. L. BUTZER and R. J. NESSEL (1971), Fourier analysis and approximation Vol. 1, Birkhauser Verlag, Basel.

9. E. W. CHENEY and A. SHARMA (1964), Bernstein power series, Canadian J. of Math. 16, 241-252.

10. R. A. DEVORE (1972), The approximation of continuous functions by positive linear operators, Lecture Notes in Mathematics Vol. 293, Springer-Verlag, N. Y.

11. A. R. FORREST (1971), Coons' surfaces and multivariable functional interpolation, Cambridge CAD Group Doc. 38.

12. A. R. FORREST (1972), Interactive interpolation and approximation by Bézier polynomials, The Comp. J. 15, 71-79.

13. W. GORDON and R. F. RIESENFELD (1974 a)), Bernstein-Bézier methods for the computer aided design of free-form curves and surfaces, J. of the A. C. M. 21, 293-310.

14. W. GORDON and R. F. RIESENFELD (1974 b)), B-spline curves and surfaces, Proceedings 2., 95-126.

15. E. HILLE and R. PHILLIPS (1974), Functional Analysis and Semigroups, Amer. Math. Soc. Colloquium Publ.XXXI, Providence.

16. S. KARLIN (1968), Total Positivity, Vol. 1, Stanford University Press, Stanford.

17. S. KARLIN and J. McGREGOR (1960), Classical diffusion processes and total positivity, J. of Math. Anal. and Appl. 1, 163-183.

18. S. KARLIN and J. McGREGOR (1962), On a genetics model of Moran, Proceedings of the Camb. Phil. Soc. 58, 299-311.

19. S. KARLIN and Z. ZIEGLER (1970), Iteration of positive approximation operators, J. of Approx. Theory 3, 310-339.

20. R. P. KELISKY and T. J. RIVLIN (1967), Iterates of Bernstein polynomials, Pacific J. of Math. 21, 511-520.

21. D. G. KENDALL (1954), Bernstein polynomials and semigroups of operators, Math. Scand. 2, 185-186.

22. A. G. KONHEIM and T. J. RIVLIN (1968), The bounded linear operators that commute with the Bernstein polynomials, Bulletin Amer. Math. Soc. 74, 111-114.

23. P. P. KOROVKIN (1960), Linear operators and approximation theory, Hindustan Publishing Co., Delhi.

24. C. LANCZOS (1938), Trigonometric interpolation of empirical and analytical functions, J. Math. Phys. 17, 129-199.

25. C. LANCZOS (1952), Introduction, in Tables of Chebyshev Polynomials $S_n(x)$ and $C_n(x)$, National Bureau of Standards, Appl. Maths. Series, pp. V-XXIX.

26. C. LANCZOS (1956), Applied Analysis, Prentice Hall, New Jersey.

27. K. de LEEUW (1959), On the degree of approximation by Bernstein polynomials, J. d'Anal. Math. 7, 89-104.

28. G. G. LORENTZ (1953), Bernstein polynomials, University of Toronto Press, Toronto.

29. G. G. LORENTZ (1966), Approximation of functions, Holt, Rinehart and Wiston, N. Y.

30. G. G. LORENTZ and L. L. SCHUMAKER (1972), Saturation of positive operators, J. Approx. Theory 5, 413-424.

31. M. MARSDEN and I. J. SCHOENBERG (1966), On variation diminishing spline approximation methods, Math. (Cluj) 8, 61-82.

32. C. A. MICCHELLI (1973), The saturation class and iterates of the Bernstein polynomials, J. Approx. Theory 8, 1-18.

33. G. M. NIELSON, R. F. RIESENFELD, and N. A. WEISS (1976), Iterates of Markov operators, J. Approx. Theory 17, 321-331.

34. E. L. ORTIZ (1977), Polynomial condensation in one and several variables with applications, in Topics in Numerical Analysis III, ed. by J. J. H. Miller.

35. E. L. ORTIZ and M. R. da SILVA (1978), Polynomial approximation of shapes based on condensed representations of Bernstein polynomials in one and several variables, to be published by the London Math. Soc.

36. G. PÓLYA and I. J. SCHOENBERG (1958), Remarks on de la Vallée Poussin means and convex conformal maps of the circle, Pacific

J. of Math. 8, 295-334.

37. G. PÓLYA and G. SZEGÖ (1976), Problems and Theorems in Analysis, Vol. II transl. by C. E. Billigheimer, Springer-Verlag, N. Y.

38. P. C. ROSENBLOOM (1967), Bounds of functions of matrices, The Amer. Math. Monthly 74, 920-926.

39. I. J. SCHOENBERG (1959), On variation diminishing approximation methods, in On Numerical Approximation, ed. by R. E. Langer, 249-274.

40. I. J. SCHOENBERG (1967), On spline functions, in Inequalities, ed. by O. Shisha, 255-291, Academic Press, N. Y.

41. I. J. SCHOENBERG and A. WHITNEY (1951), A theorem on polygons in n dimensions with applications to variation-diminishing and cyclic variation-diminishing linear transformations, Comp. Math. 9, 141-160.

42. F. SCHURER (1962), On the order of approximation with generalized Bernstein polynomials, Indag. Math. 24, 484-488.

43. F. SCHURER (1963), On the approximation of functions of many variables with linear positive operators, Indag. Math. 25, 313-327.

44. F. SCHURER and F. W. STEUTEL (1976), On an inequality of Lorentz in the theory of Bernstein polynomials, Lecture Notes in Mathematics 501, pp. 332-338, Springer-Verlag, N. Y.

45. F. SCHURER and F. W. STEUTEL (1977), The degree of local approximation of functions in $C_1[0,1]$ by Bernstein polynomials, J. Approx. Theory 19, 69-82.

46. P. C. SIKKEMA (1961), Der wert einiger konstanten in der theorie der approximation mit Bernstein-polynomen, Numer. Math. 3, 107-116.

47. P. C. SIKKEMA (1966), Über potenzen von verallgemeinerten Bernstein-operatoren, Math. (Cluj) 8, 173-180.

48. D. D. STANCU (1960 a)), Some Bernstein polynomials in two variables and their applications, Sov. Math. 1, 1025-1027.

49. D. D. STANCU (1960 b)), On some polynomials of Bernstein type (Rumanian with Russian and French summary), Acad. Rep. Pop. Rom. Fil. Yasi, Stud. Cerc. St. Mat. 11, 221-233.

50. D. D. STANCU (1963 a)), A method for obtaining polynomials of Bernstein type of two variables, Amer. Math. Monthly 70, 260-264.

51. D. D. STANCU (1963 b)), Evaluation of the remainder term in approximation formulas by Bernstein polynomials, Math. Comp. 17, 270-278.

52. D. D. STANCU (1964) The remainder of certain linear approximation formulas in two variables, SIAM J. Numer. Anal. 1, 139-163.

53. Y. SUZUKI (1967), Saturation of local approximation by linear positive operators of Bernstein type, Tôhoku Math. J. 19, 429-453.

54. H. F. TROTTER (1958), Approximation of semi-groups of operators, Pacific J. Math. 8, 887-919.