JOURNAL OF APPROXIMATION THEORY 50, 111–126 (1987)

Variation Diminishing Properties of Bernstein Polynomials on Triangles

T. N. T. GOODMAN

Department of Mathematical Sciences, The University, Dundee DD1 4HN, Scotland

Communicated by Philip J. Davis

Received October 2, 1984; revised March 21, 1985

DEDICATED TO THE MEMORY OF GÉZA FREUD

We show that if f is any function on a triangle T, then the variation of the gradient of the nth Bernstein polynomial $B_n(f)$ on T cannot exceed the variation of the gradient of the nth Bézier net $\hat{f}_n$ on T. We deduce that if f is in $C^2(T)$, then the variation of the gradient of $B_n(f)$ on T is bounded by a given constant (depending on T) times the variation of the gradient of f. It is further shown that the variation of $B_n(f)$ on T is bounded by $2n/(n+1)$ times the variation of $\hat{f}_n$ on T. © 1987 Academic Press, Inc.

1. INTRODUCTION

In order to motivate our discussion we first recall some properties of univariate Bernstein polynomials [3]. Let p be any polynomial of degree $n \geqslant 1$. Then we can write p in the form

$$p(x) = \sum_{i=0}^{n} a_i \binom{n}{i} (1-x)^i x^{n-i}. \tag{1}$$

Then

$$\int_0^1 |p(x)|\, dx \leqslant \sum_{i=0}^{n} |a_i| \binom{n}{i} \int_0^1 (1-x)^i x^{n-i} dx$$

which gives

$$\int_0^1 |p(x)|\, dx \leqslant \frac{1}{n+1} \sum_{i=0}^{n} |a_i|. \tag{2}$$

0021-9045/87 $3.00

Applying (2) to p' and p'' gives

$$\int_0^1 |p'(x)|\,dx \leqslant \sum_{i=0}^{n-1} |a_{i+1} - a_i|, \tag{3}$$

$$\int_0^1 |p''(x)|\,dx \leqslant n \sum_{i=0}^{n-2} |a_{i+2} - 2a_{i+1} + a_i|. \tag{4}$$

Inequalities (3) and (4) can be expressed neatly by introducing the function $\hat{a}$ on $[0, 1]$ which is linear on $[i/n, (i+1)/n]$ for $i = 0, 1, ..., n-1$ and satisfies $\hat{a}(i/n) = a_i$ $(i = 0, 1, ..., n)$. We denote by $V(f, [0, 1])$ the total variation of a function f on $[0, 1]$ and write $V_1(f, [0, 1]) = V(f', [0, 1])$. Then (3) and (4) can be written as

$$V(p, [0, 1]) \leqslant V(\hat{a}, [0, 1]), \tag{5}$$

$$V_1(p, [0, 1]) \leqslant V_1(\hat{a}, [0, 1]). \tag{6}$$

Another related result, due to Pólya and Schoenberg [5], is that if $S(f, [0, 1])$ denotes the number of times a function f changes sign on $[0, 1]$, then

$$S(p, [0, 1]) \leqslant S(\hat{a}, [0, 1]). \tag{7}$$

We now see how inequalities (5), (6), and (7) are equivalent to "variation diminishing" properties of Bernstein polynomials. For any function f on $[0, 1]$ we define the Bernstein polynomial $B_n(f)$ by

$$B_n(f)(x) = \sum_{i=0}^{n} f\left(\frac{i}{n}\right)\binom{n}{i} (1-x)^i x^{n-i}. \tag{8}$$

Let $\hat{f}_n$ denote the function on $[0, 1]$ which is linear on $[i/n, (i+1)/n]$ for $i = 0, 1, ..., n-1$ and interpolates f at i/n $(i = 0, 1, ..., n)$. Then from (5), (6), and (7) we have

$$V(B_n(f), [0, 1]) \leqslant V(\hat{f}_n, [0, 1]), \tag{9}$$

$$V_1(B_n(f), [0, 1]) \leqslant V_1(\hat{f}_n, [0, 1]), \tag{10}$$

$$S(B_n(f), [0, 1]) \leqslant S(\hat{f}_n, [0, 1]). \tag{11}$$

It is easily seen that

$$V(\hat{f}_n, [0, 1]) \leqslant V(f, [0, 1])$$

$$V_1(\hat{f}_n, [0, 1] \leqslant V_1(f, [0, 1]) \qquad (f \in C^1([0, 1])),$$

$$S(\hat{f}_n, [0, 1]) \leqslant S(f, [0, 1])$$

and so

$$V(B_n(f), [0, 1]) \leqslant V(f, [0, 1]), \tag{12}$$

$$V_1(B_n(f), [0, 1]) \leqslant V_1(f, [0, 1]) \qquad (f \in C^1([0, 1])), \tag{13}$$

$$S(B_n(f), [0, 1]) \leqslant S(f, [0, 1]). \tag{14}$$

We now turn to a consideration of analogous results for Bernstein polynomials on triangles. Let T_1, T_2, T_3 be the vertices of the triangle T. We shall identify any point P in T with its barycentric coodinates (u, v, w), where

$$P = uT_1 + vT_2 + wT_3, \qquad u + v + w = 1, \qquad u \geqslant 0, v \geqslant 0, w \geqslant 0.$$

For any function f on T we define the Bernstein polynomial $B_n(f, T)$, or simply $B_n(f)$, by

$$B_n(f)(P) := \sum_{i+j+k=n} f\left(\frac{i}{n}, \frac{j}{n}, \frac{k}{n}\right) \frac{n!}{i!j!k!} u^i v^j w^k. \tag{15}$$

We denote by $\hat{f}_n$ the function on T which interpolates f at $\{(i/n, j/n, k/n)$: $i+j+k=n\}$ and which is linear on each element of Ω_n, the regular triangulation of T at the same points $\{(i/n, j/n, k/n)\}$. This function $\hat{f}_n$ is called the *nth Bézier net of* [2]. For future reference we now give an explicit description of the triangulation Ω_n. For $i, j, k \geqslant 0$, $i+j+k=n-1$, we let U_{ijk} denote the triangle with vertices

$$\left(\frac{i+1}{n}, \frac{j}{n}, \frac{k}{n}\right), \qquad \left(\frac{i}{n}, \frac{j+1}{n}, \frac{k}{n}\right), \qquad \left(\frac{i}{n}, \frac{j}{n}, \frac{k+1}{n}\right).$$

For $i, j, k \geqslant 0$, $i+j+k=n-2$, we let W_{ijk} denote the triangle with vertices

$$\left(\frac{i}{n}, \frac{j+1}{n}, \frac{k+1}{n}\right), \qquad \left(\frac{i+1}{n}, \frac{j}{n}, \frac{k+1}{n}\right), \qquad \left(\frac{i+1}{n}, \frac{j+1}{n}, \frac{k}{n}\right).$$

Then $\Omega_n = \{U_{ijk}: i+j+k=n-1\} \cup \{W_{ijk}: i+j+k=n-2\}$. This is illustrated for $n=4$ in Fig. 1, where the triangles U_{ijk} and W_{ijk} are shown in white and black, respectively.

We require the analogs of V and V_1 to vanish only for constant functions and for linear functions, respectively. Since we also require them to be invariant under rotations, we define

$$V(f, T) := \int_T (f_x^2 + f_y^2)^{1/2}, \tag{16}$$

$$V_1(f, T) := \int_T (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2)^{1/2}. \tag{17}$$

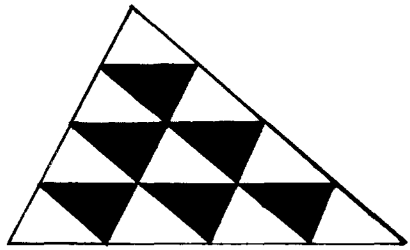

FIGURE 1

Definition (17) does not make sense for the function $\hat{f}_n$, except in the sense of taking a limit as now described. For $i, j, k \geqslant 0$, $i+j+k=n-2$, the change in the gradient of $\hat{f}_n$ across the line segment l_{ijk} joining $((i+1)/n, (j+1)/n, k/n)$ and $((i+1)/n, j/n, (k+1)/n)$ is easily seen to have magnitude

$$\left| f\left(\frac{i+2}{n}, \frac{j}{n}, \frac{k}{n}\right) + f\left(\frac{i}{n}, \frac{j+1}{n}, \frac{k+1}{n}\right) - f\left(\frac{i+1}{n}, \frac{j+1}{n}, \frac{k}{n}\right) - f\left(\frac{i+1}{n}, \frac{j}{n}, \frac{k+1}{n}\right) \right| n |T_2 T_3| / 2\Delta \tag{18}$$

where $T_2 T_3$ denotes the vector from T_2 to T_3 and $\Delta = \text{Area } T$.

So $V_1(\hat{f}_n, N)$ for some small neighbourhood N around l^1_{ijk} can be regarded as (18) multiplied by the length of l^1_{ijk}. Denoting this by $V^1_{ijk}(\hat{f}_n)$ we have

$$V^1_{ijk}(\hat{f}_n) = \left| f\left(\frac{i+2}{n}, \frac{j}{n}, \frac{k}{n}\right) + f\left(\frac{i}{n}, \frac{j+1}{n}, \frac{k+1}{n}\right) - f\left(\frac{i+1}{n}, \frac{j+1}{n}, \frac{k}{n}\right) - f\left(\frac{i+1}{n}, \frac{j}{n}, \frac{k+1}{n}\right) \right| |T_2 T_3|^2 / 2\Delta \tag{20}$$

Corresponding formulae hold for the variation $V^2_{ijk}(\hat{f}_n)$ and $V^3_{ijk}(\hat{f}_n)$ around the line joining $((i+1)/n, (j+1)/n, k/n)$ to $(i/n, (j+1)/n, (k+1)/n)$ and around the line joining $((i+1)/n, j/n, (k+1)/n)$ to $(i/n, (j+1)/n, (k+1)/n)$, respectively. We therefore define

$$V_1(\hat{f}_n, T) = \sum_{s=1}^{3} \sum_{i+j+k=n-2} V^s_{ijk}(\hat{f}_n). \tag{21}$$

We have now formulated analogs of (9) and (10), namely

$$V(B_n(f), T) \leqslant V(\hat{f}_n, T), \tag{22}$$

$$V_1(B_n(T) \leqslant V_1(\hat{f}_n, T). \tag{23}$$

It is not clear to the author how to formulate an analog of inequality (11). We prove in Section 2 that inequality (23) does indeed hold and is sharp. This says, roughly speaking, that $B_n(f)$ cannot be further from being linear than is $\hat{f}_n$. In Section 4 we show that (22) does not hold in general but instead we have

$$V(B_n(f), T) \leqslant \frac{2n}{n+1} V(\hat{f}_n, T). \tag{24}$$

It remains to consider analogs of (12) and (13). Now for any given values $\{f_{ijk}: i, j, k \geqslant 0,\ i+j+k=n\}$, we can find a smooth function f satisfying $f(i/n, j/n, k/n) = f_{ijk}$ $(i+j+k=n)$ for which $V(f, T)$ is arbitrarily small. Thus there is no constant C such that $V(B_n(f, T) \leqslant CV(f, T)$ for all smooth functions f. In contrast we show in Section 3 that there is a constant C (depending on the angles of T) such that

$$V_1(B_n(f), T) \leqslant CV_1(f, T)$$

for all f in $C^2(T)$. Such a constant C can be given explicitly in terms of the angles of T; in particular C can be taken <2 if T is close enough to equilateral.

Finally we mention a recent result of Chang and Davis [1] which has some similarities with our results and was partly responsible for the writing of this paper. The authors show that if $\hat{f}_n$ is convex on T, then $B_n(f)$ is convex; but also give an example of a convex function f on T for which $B_n(f)$ is not convex.

2. Theorem 1

THEOREM 1. *For any $n \geqslant 1$ and any function f on the triangle T,*

$$V_1(B_n(f), T) \leqslant V_1(\hat{f}_n, T). \tag{23}$$

Our method of proof is to derive (23) first for the special case when $B_n(f)$ is of degree 2 and then subdivide T and approximate $B_n(f)$ on each sub-triangle by a quadratic polynomial. We shall denote by Π_r the space of polynomials of degree r and write $f_{ijk} = f(i/n, j/n, k/n)$, $\Delta = \text{Area } T$.

LEMMA 1. *Inequality* (23) *holds if $B_n(f)$ is in Π_2.*

Proof. Choose h in Π_1 so that $(B_n(f) - h)(u, v, w) = au^2 + bv^2 + cw^2$ for some a, b, c. Putting $g = f - h$, we have $B_n(g) = B_n(f) - B_n(h) = B_n(f) - h$, since B_n reproduces Π_1. Then $V_1(B_n(f), T) = V_1(B_n(g) + h, T) =$

$V_1(B_n(g), T)$ and $V_1(\hat{f}_n, T) = V_1(\hat{g}_n + h, T) = V_1(\hat{g}_n, T)$. So it is sufficient to prove

$$V_1(B_n(g), T) \leqslant V_1(\hat{g}_n, T). \tag{25}$$

A straightforward calculation shows that

$$\begin{aligned} V_1(B_n(g), T) = \frac{1}{2\Delta} \{ & a^2|T_2T_3|^4 + b^2|T_3T_1|^4 + c^2|T_1T_2|^4 \\ & + 2ab(T_2T_3 \cdot T_3T_1)^2 + 2bc(T_3T_1 \cdot T_1T_2)^2 \\ & + 2ca(T_1T_2 \cdot T_2T_3)^2\}^{1/2}. \end{aligned} \tag{26}$$

We now calculate $V_1(\hat{g}_n, T)$. Since $B_n(g) = (au^2 + bv^2 + cw^2) \times (u+v+w)^{n-2}$, we have for $i+j+k=n$,

$$g\left(\frac{i}{n}, \frac{j}{n}, \frac{k}{n}\right) = \frac{1}{n(n-1)} \{ai(i-1) + bj(j-1) + ck(k-1)\}.$$

Applying (20) now gives for $i+j+k=n-2$,

$$V_{ijk}^1(\hat{g}_n) = \frac{|a| \cdot |T_2T_3|^2}{n(n-1)\,\Delta}. \tag{27}$$

Deriving corresponding formulae for $V_{ijk}^2(\hat{g}_n)$ and $V_{ijk}^3(\hat{g}_k)$ and substituting into (21) gives

$$V_1(\hat{g}_n, T) = \frac{1}{2\Delta} \{|a| \cdot |T_2T_3|^2 + |b| \cdot |T_3T_1|^2 + |c| \cdot |T_1T_2|^2\}. \tag{28}$$

From (26) and (28) we can easily deduce (25). ∎

Now let

$$T_{ij} := \tfrac{1}{2}(T_i + T_j) \qquad (1 \leqslant i, j \leqslant 3) \tag{29}$$

and let $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ denote respectively the triangles $T_1T_{12}T_{31}$, $T_{12}T_2T_{23}$, $T_{31}T_{23}T_3$, $T_{23}T_{31}T_{12}$. Take any function f on T and for $r = 1, \ldots, 4$ write

$$B_n(f)|\Delta_r = \sum_{i+j+k=n} a_{ijk}^r \frac{n!}{i!j!k!} u^i v^j w^k, \tag{30}$$

where u, v, w now denote barycentric coordinates for Δ_r. (E. g., if $r=1$, $(u, v, w) = uT_1 + vT_{12} + wT_{31}$.)

LEMMA 2. *We have the equations*

$$a^1_{ijk}=\frac{1}{2^{j+k}}\sum_{\beta=0}^{j}\sum_{\gamma=0}^{k}\binom{j}{\beta}\binom{k}{\gamma}f_{n-\beta-\gamma,\beta,\gamma}, \tag{31}$$

$$a^2_{ijk}=\frac{1}{2^{k+i}}\sum_{\gamma=0}^{k}\sum_{\alpha=0}^{i}\binom{k}{\gamma}\binom{i}{\alpha}f_{\alpha,n-\gamma-\alpha,\gamma}, \tag{32}$$

$$a^3_{ijk}=\frac{1}{2^{i+j}}\sum_{\alpha=0}^{i}\sum_{\beta=0}^{j}\binom{i}{\alpha}\binom{j}{\beta}f_{\alpha,\beta,n-\alpha-\beta}, \tag{33}$$

$$a^4_{ijk}=\frac{1}{2^{n}}\sum_{\alpha=0}^{i}\sum_{\beta=0}^{j}\sum_{\gamma=0}^{k}\binom{i}{\alpha}\binom{j}{\beta}\binom{k}{\gamma}f_{k+\beta-\gamma,i+\gamma-\alpha,j+\alpha-\beta}. \tag{34}$$

Furthermore

$$\sum_{r=1}^{4}\sum_{i+j+k=n}a^r_{ijk}=4\sum_{i+j+k=n}f_{ijk}. \tag{35}$$

Proof. Equations (32)–(34) follow immediately from Theorem 8 of [1]. To prove (35) we note that the function F_{ijk} on T given by $F_{ijk}(u,v,w)=(n!/i!j!k!)\,u^iv^jw^k$ satisfies

$$\int F_{ijk}=\frac{2\Delta}{(n+1)(n+2)}. \tag{36}$$

This can be seen either by direct calculation or by noting that $F_{ijk}=2\Delta/(n+1)(n+2)$ times a multivariate B-spline (Corollary 2 of [4]). From (36) and (15) we see that

$$\int B_n(f)=\frac{2\Delta}{(n+1)(n+2)}\sum_{i+j+k=n}f_{ijk}. \tag{37}$$

Then (35) follows on applying (37) to the triangles $\Delta_1,\ldots,\Delta_4$ and noting that $\sum_{r=1}^{4}\int_{\Delta_r}B_n(f)=\int_T B_n(f)$. ∎

Formulae (31)–(34) can be expressed more neatly by introducing the following averaging operators. For any array $\{b_{ijk}\}$ we define

$$(A_1b)_{ijk}=\tfrac{1}{2}(b_{i,j-1,k+1}+b_{i,j+1,k-1}), \tag{38}$$

$$(A_2b)_{ijk}=\tfrac{1}{2}(b_{i+1,j,k-1}+b_{i-1,j,k+1}), \tag{39}$$

$$(A_3b)_{ijk}=\tfrac{1}{2}(b_{i-1,j+1,k}+b_{i+1,j-1,k}). \tag{40}$$

LEMMA 3. *Let* $\{b_{ijk} : i+j+k=2n\}$ *be any array satisfying* $b_{2i,2j,2k} = f_{ijk}$. *Then for* $i+j+k=n$,

$$a^1_{ijk} = (A_2^k A_3^j b)_{i+n,j,k}, \tag{41}$$

$$a^2_{ijk} = (A_1^k A_3^i b)_{i,j+n,k}, \tag{42}$$

$$a^3_{ijk} = (A_1^j A_2^i b)_{i,j,k+n}, \tag{43}$$

$$a^4_{ijk} = (A_1^i A_2^j A_3^k b)_{n-i,n-j,n-k}. \tag{44}$$

Proof. To derive (41) we note that from (40),

$$(A_3^j b)_{i+n,j,k} = \frac{1}{2^j} \sum_{\beta=0}^{j} \binom{j}{\beta} b_{i+n+j-2\beta,2\beta,k}$$

and from (39)

$$\begin{aligned}(A_2^k A_3^j b)_{i+n,j,k} &= \frac{1}{2^{j+k}} \sum_{\beta=0}^{k} \binom{k}{\gamma} \sum_{\beta=0}^{j} \binom{j}{\beta} b_{i+n+j-2\beta+k-2\gamma,2\beta,2\gamma} \\ &= \frac{1}{2^{j+k}} \sum_{\beta=0}^{j} \sum_{\gamma=0}^{k} \binom{j}{\beta}\binom{k}{\gamma} f_{n-\beta-\gamma,\beta,\gamma} \\ &= a^1_{ijk}\end{aligned}$$

by (31). Formulae (42)–(44) follow similarly. ∎

From Eq. (35) and Lemma 3 we can immediately deduce

LEMMA 4. *If* $\{b_{ijk} : i+j+k=2n\}$ *is any array, then*

$$\begin{aligned}\sum_{i+j+k=n} &\{(A_2^k A_3^j b)_{i+n,j,k} + (A_1^k A_3^i b)_{i,j+n,k} \\ &+ (A_1^j A_2^i b)_{i,j,k+n} + (A_1^i A_2^j A_3^k b)_{n-i,n-j,n-k}\} \\ = 4 &\sum_{i+j+k=n} b_{2i,2j,2k}.\end{aligned}$$

LEMMA 5. *For* $r=1,\dots,4$, *let* g^r *denote any function on* Δ_r *for which* $B_n(f)|\Delta_r = B_n(g^r, \Delta_r)$. *Then*

$$\sum_{r=1}^{4} V_1(\hat{g}^r_n, \Delta_r) \leqslant V_1(\hat{f}_n, T). \tag{45}$$

Proof. Note that using barycentric coordinates on Δ_r, $g^r(i/n,\ j/n,$

$k/n) = a^r_{ijk}$ $(i+j+k=n)$. Now consider triangle Δ_1 and take i, j, k with $i+j+k=n-2$. Then from (41), after some simplification,

$$\begin{aligned} &a^1_{i+2,j,k} + a^1_{i,j+1,k+1} - a^1_{i+1,j+1,k} - a^1_{i+1,j,k+1} \\ &\quad = \tfrac{1}{4} A_2^k A_3^j (b_{i+2+n,j,k} + b_{i-2+n,j+2,k+2} - b_{i+n,j+2,k} - b_{i+n,j,k+2}). \end{aligned}$$

So from (20),

$$V^1_{ijk}(\hat{g}_n) \leqslant \tfrac{1}{4}(A_2^k A_3^j v)_{i-2+n,j,k} \, |T_2 T_3|^2/2\Delta, \tag{46}$$

where for $i+j+k=2n-4$,

$$v_{ijk} = |b_{i+4,j,k} + b_{i,j+2,k+2} - b_{i+2,j+2,k} - b_{i+2,j,k+2}|. \tag{47}$$

Similarly we have for $i+j+k=n-2$,

$$\begin{aligned} &V'_{ijk}(\hat{g}^2_n) \leqslant \tfrac{1}{4}\,(A_1^k A_3^j v)_{i,j-2+n,k} \, |T_2 T_3|^2/2\Delta, \\ &V'_{ijk}(\hat{g}^3_n) \leqslant \tfrac{1}{4}(A_1^j A_2^i v)_{i,j,k-2+n} \, |T_2 T_3|^2/2\Delta, \\ &V'_{ijk}(\hat{g}^4_n) \leqslant \tfrac{1}{4}(A_1^i A_2^j A_3^k v)_{n-2-i,n-2-j,n-2-k} \, |T_2 T_3|^2/2\Delta. \end{aligned} \tag{48}$$

From (46), (48), and Lemma 4 (with b replaced by v and n by $n-2$) we have

$$\begin{aligned} \sum_{r=1}^{4} \sum_{i+j+k=n-2} V'_{ijk}(\hat{g}^r_n) &\leqslant \sum_{i+j+k=n-2} V_{2i,2j,2k} |T_2 T_3|^2/2\Delta \\ &= \sum_{i+j+k=n-2} |f_{i+2,j,k} + f_{i,j+1,k+1} - f_{i+1,j+1,k} \\ &\qquad - f_{i+1,j,k+1}| \, |T_2 T_3|^2/2\Delta \\ &= \sum_{i+j+k=n-2} V'_{ijk}(\hat{f}_n) \end{aligned}$$

by (47), the fact that $b_{2i,2j,2k} = f_{ijk}$, and (20).

Deriving corresponding inequalities for $\sum_{r=1}^4 \sum_{i+j+k=n-2} V^s(\hat{g}^r_n)$ $(s=2,3)$ and applying (21) gives

$$\begin{aligned} \sum_{r=1}^{4} V_1(\hat{g}^r_n, \Delta_r) &= \sum_{r=1}^{4} \sum_{s=1}^{3} \sum_{i+j+k=n-2} V^s_{ijk}(\hat{g}^r_n) \\ &\leqslant \sum_{s=1}^{3} \sum_{i+j+k=n-2} V^s_{ijk}(\hat{f}_n) \\ &= V_1(\hat{f}_n, T). \quad \blacksquare \end{aligned}$$

Proof of Theorem 1. For any p in Π_n, $p = B_n(g)$ for some function g on T and $\hat{g}_n$ depends only on p and not the choice of g. Then $V_1(\hat{g}_n, T)$ and $V_1(p, T)$ are both norms on Π_n/Π_1 and so there is a constant C such that

$$V_1(\hat{g}_n, T) \leqslant CV_1(B_n(g), T) \tag{49}$$

for all functions g on T. Moreover since (49) is invariant under translation, rotation or change of scale, it also holds for any triangle similar to T.

Now take $m \geqslant 1$ and recall that Ω_m is the triangulation of T defined after Eq. (15). Take any triangle Δ in Ω_m and write $B_n(f)|\Delta = q(\Delta) + r(\Delta)$, where $q(\Delta)$ is in Π_2 and all second-order derivatives of $r(\Delta)$ vanish at some point in Δ. Then the second-order derivatives of $r(\Delta)$ tend to zero as $m \to \infty$ uniformly over all Δ in Ω_m and so

$$V_1(B_n(f), \Delta) = V_1(q(\Delta), \Delta) + o\left(\frac{1}{m^2}\right). \tag{50}$$

Now choose functions g, h on Δ such that $B_n(f)|\Delta = B_n(g, \Delta)$ and $q(\Delta) = B_n(h, V)$. Then from Lemma 1, with T replaced by Δ,

$$\begin{aligned} V_1(q(\Delta), \Delta) &\leqslant V_1(\hat{h}_n, \Delta) \\ &\leqslant V_1(\hat{g}_n, \Delta) + V_1((h-g)\hat{}_n, \Delta) \\ &\leqslant V_1(\hat{g}_n, \Delta) + CV_1(B_n(h-g), \Delta) \end{aligned} \tag{51}$$

by (49). Now $B_n(h-g, \Delta) = B_n(h, \Delta) - B_n(g, \Delta) = q(\Delta) - B_n(f) = -r(\Delta)$ and so

$$V_1(B_n(h-g), \Delta) = V_1(r(\Delta), \Delta) = o(1/m^2). \tag{52}$$

Combining (50), (51), and (52) gives

$$V_1(B_n(f), \Delta) \leqslant V_1(\hat{g}_n(\Delta), \Delta) + o(1/m^2)$$

where we have replaced $\hat{g}_n$ by $\hat{g}_n(\Delta)$ to make clear it depends on Δ. Thus

$$\begin{aligned} V_1(B_n(f), T) &= \sum_{\Delta \in \Omega_m} V_1(B_n(f), \Delta) \\ &\leqslant \sum_{\Delta \in \Omega_m} V_1(\hat{g}_n(\Delta), \Delta) + o(1). \end{aligned} \tag{53}$$

If $m = 2^s$ we can apply Lemma 5 successively to give

$$\sum_{\Delta \in \Omega_m} V_1(\hat{g}_n, \Delta) \leqslant V_1(\hat{f}_n, T). \tag{54}$$

Combining (53) and (54) and letting $s \to \infty$ then completes the proof of Theorem 1. ∎

We see from the proof of Lemma 1 that equality holds in (23) when $B_n(f)(u, v, w) = u^2$.

3. Theorem 2

THEOREM 2. *For any f in $C^2(T)$,*

$$V_1(B_n(f), T) \leqslant CV_1(f, T), \tag{55}$$

where if T has angles α, β, γ, then

$$C^{-2} = \min\{a^2 + b^2 + c^2 + 2bc \cos^2 \alpha + 2ca \cos^2 \beta + 2ab \cos^2 \gamma \\ : |a| + |b| + |c| = 1\}. \tag{56}$$

It is easily seen that if T is equilateral, then $C = \sqrt{34}/3$. Now from Theorem 1 it is sufficient to prove that for f in $C^2(T)$,

$$V_1(\hat{f}_n, T) \leqslant CV_1(f, T). \tag{57}$$

In a similar manner to the proof of Theorem 1 we shall first prove this for f in Π_2 and then subdivide T and approximate f on each subtriangle by a quadratic polynomial.

LEMMA 6. *Inequality* (57) *holds if f is in Π_2.*

Proof. Choose h in Π_1 so that $(f-h)(u, v, w) = au^2 + bv^2 + cw^2$ for some a, b, c. Putting $g = f - h$, we have $V_1(\hat{f}_n, T) = V_1(\hat{g}_n + h, T) = V_1(\hat{g}_n, T)$ and $V_1(f, T) = V_1(g, T)$. Thus it is sufficient to prove that

$$V_1(\hat{g}_n, T) \leqslant CV_1(g, T). \tag{58}$$

From (26) we have

$$V_1(g, T) = \frac{1}{2\Delta}\{a^2|T_2T_3|^4 + b^2|T_3T_1|^4 + c^2|T_1T_2|^4 \\ + 2ab|T_2T_3|^2|T_3T_1|^2 \cos^2 \gamma + 2bc|T_3T_1|^2|T_1T_2|^2 \cos^2 \alpha \\ + 2ca|T_1T_2|^2|T_2T_3|^2 \cos^2 \beta\}^{1/2}. \tag{59}$$

We now calculate $V_1(\hat{g}_n, T)$. For $i+j+k=n$, $g(i/n, j/n, k/n) = (1/n^2)(ai^2 + bj^2 + ck^2)$. So from (20), for $i+j+k=n-2$,

$$V_{ijk}^1(\hat{g}_n) = |a| \cdot |T_2T_3|^2/n^2\Delta$$

with corresponding formulae for $V_{ijk}^2(\hat{g}_n)$ and $V_{ijk}^3(\hat{g}_n)$. Substituting into (21) gives

$$V_1(\hat{g}_n, T) = \frac{n-1}{2n\Delta}\{|a||T_2T_3|^2 + |b||T_3T_1|^2 + |c||T_1T_2|^2\}. \tag{60}$$

Then (58) follows from (56), (59), and (60). ∎

LEMMA 7. *For any function f on T,*

$$V_1(\hat{f}_n, T) \leqslant V_1(\hat{f}_{2n}, T). \tag{61}$$

Proof. We note that $f_{2i,2j,2k} = f_{ijk}$. So for $i+j+k = n-2$,

$$\begin{aligned}
&|f_{i+2,j,k} + f_{i,j,k+1} - f_{i+1,j+1,k} - f_{i+1,j,k+1}|\\
&\quad\leqslant |f_{2i+4,2j,2k} + f_{2i+2,2j+1,2k+1} - f_{2i+3,2j+1,2k} - f_{2i+3,2j,2k+1}|\\
&\qquad + |f_{2i+3,2j+1,2k} + f_{2i+1,2j+2,2k+1} - f_{2i+2,2j+2,2k} - f_{2i+2,2j+1,2k+1}|\\
&\qquad + |f_{2i+3,2j,2k+1} + f_{2i+1,2j+1,2k+2} - f_{2i+2,2j+1,2k+1} - f_{2i+2,2j,2k+2}|\\
&\qquad + |f_{2i+2,2j+1,2k+1} + f_{2i,2j+2,2k+2} - f_{2i+1,2j+2,2k+1} - f_{2i+1,2j+1,2k+2}|
\end{aligned}$$

and so from (20),

$$\begin{aligned}
V_{ijk}^1(\hat{f}_n) \leqslant{}& V_{2i+2,2j,2k}^1(\hat{f}_{2n}) + V_{2i+1,2j+1,2k}^1(\hat{f}_{2n}) + V_{2i+1,2j,2k+1}^1(\hat{f}_{2n})\\
&+ V_{2,2j+1,2k+1}^1(\hat{f}_{2n}).
\end{aligned}$$

Deriving corresponding inequalities for $V_{ijk}^2(\hat{f}_n)$ and $V_{ijk}^3(\hat{f}_n)$ and applying (21) gives (61). ∎

Proof of Theorem 2. Take $m \geqslant 1$ and any triangle Δ in Ω_m. Write $f|\Delta = q + r$ (depending on Δ), where q is in Π_2 and $r_{xx}, r_{xy}, r_{yy} \to 0$ as $m \to \infty$ uniformly over Δ in Ω_m. Then for $s \geqslant 1$, $\hat{f}_{ms}|\Delta = \hat{q}_{ms} + \hat{r}_{ms}$ and by Lemma 6,

$$\begin{aligned}
V_1(\hat{f}_{ms}, \Delta) &= V_1(\hat{q}_{ms}, \Delta) + o(1/m^2)\\
&\leqslant CV_1(q, \Delta) + o(1/m^2)\\
&= CV_1(f, \Delta) + o(1/m^2).
\end{aligned}$$

Now

$$\begin{aligned}
V_1(\hat{f}_{ms}, T) &= \sum_{\Delta \in \Omega_m} V_1(\hat{f}_{ms}, \Delta) + O(1/s)\\
&\leqslant C \sum_{\Delta \in \Omega_m} V_1(f, \Delta) + m^2 o(1/m^2) + O(1/s)\\
&= CV_1(f, T) + m^2 o(1/m^2) + O(1/s).
\end{aligned} \tag{62}$$

If $ms = 2^r n$ we can apply Lemma 7 successively to give

$$V_1(\hat{f}_n, T) \leqslant V_1(\hat{f}_{ms}, T). \tag{63}$$

Combining (62) and (63) and letting $m, s \to \infty$ gives (57). ∎

4. Theorem 3

We recall the definition of U_{ijk} given after Eq. (15) and define $U = U_n(T) = \bigcup \{U_{ijk} : i+j+k = n-1\}$. (In Fig. 1, $U = U_4(T)$ is the region of T coloured white.) We can now state the following result which clearly implies (24).

THEOREM 3. *For any $n \geqslant 1$ and any function f on T,*

$$V(B_n(f), T) \leqslant \frac{2n}{n+1} V(\hat{f}_n, U). \tag{64}$$

If f is linear we have equality in (64) for then $V(B_n(f), T) = V(f, T) = |\nabla f| \Delta = |\nabla f| (2n/(n+1))$ Area $U = (2n/(n+1))\, V(f, U) = (2n/(n+1))\, V(\hat{f}_n, U)$.

We shall prove Theorem 3 by subdividing T and approximating $B_n(f)$ on each triangle by a linear function.

LEMMA 8. *If S denotes the triangle with vertices A, B, C, and g is the linear function on S satisfying $g(A) = a$, $g(B) = b$, $g(C) = c$, then*

$$V(g, S) = \tfrac{1}{2} |(b-c)\, A + (c-a)\, B + (a-b)\, C|. \tag{65}$$

Proof. Straightforward calculation.

LEMMA 9. *If g^r and Δ_r are as in Lemma 5 ($r = 1,..., 4$), then*

$$\sum_{r=1}^{4} V(\hat{g}^r_n, U_n(\Delta_r)) \leqslant V(\hat{f}_n, U_n(T)). \tag{66}$$

Proof. For $i, j, k \geqslant 0$ with $i+j+k = 2n-2$ we let Δ_{ijk} denote the subtriangle of T with vertices

$$\left(\frac{i+2}{2n}, \frac{j}{2n}, \frac{k}{2n}\right), \qquad \left(\frac{i}{2n}, \frac{i+2}{2n}, \frac{k}{2n}\right), \qquad \left(\frac{i}{2n}, \frac{j}{2n}, \frac{k+2}{2n}\right).$$

We note that for $i+j+k = n-1$, $\Delta_{2i,2j,2k} = U_{ijk}$. For $i+j+k = n$, $r = 1,..., 4$

we define a^r_{ijk} as in (30), while for $i+j+k=2n$ we define b_{ijk} as in Lemma 3. If L denotes the linear function on Δ_{ijk} satisfying

$$L\left(\frac{i+2}{2n}, \frac{j}{2n}, \frac{k}{2n}\right)=b_{i+2,j,k}, \qquad L\left(\frac{i}{2n}, \frac{i+2}{2n}, \frac{k}{n}\right)=b_{i,j+2,k},$$

$$L\left(\frac{i}{2n}, \frac{j}{2n}, \frac{k+2}{2n}\right)=b_{i,j,k+2}$$

we write

$$X_{ijk}=V(L, \Delta_{ijk}). \tag{67}$$

Then

$$V(\hat{f}_n, U_n(T))=\sum_{i+j+k=n-1} V(\hat{f}_n, U_{ijk})=\sum_{i+j+k=n-1} X_{2i,2j,2k}. \tag{68}$$

Now for $i+j+k=n-1$, consider the triangle $U_{i+n,j,k}$. From (41) we see, after some simplification, that

$$\begin{aligned}(a^1_{i,j+1,k}-a^1_{i,j,k+1}, a^1_{i,j,k+1}-a^1_{i+1,j,k}, a^1_{i+1,j,k}-a^1_{i,j+1,k})\\ =\tfrac{1}{2}A_2^k A_3^j(b_{i+n-1,j+2,k}-b_{i+n-1,j,k+2}, b_{i+n-1,j,k+2}-b_{i+n+1,j,k},\\ b_{i+n+1,j,k}-b_{i+n-1,j+2,k}).\end{aligned}$$

From Lemma 8 it follows that

$$V(\hat{g}'_n, U_{i+n,j,k})\leqslant \tfrac{1}{4}(A_2^k A_3^j X)_{i+n-1,j,k} \tag{69}$$

and so

$$\begin{aligned}V(\hat{g}'_n, U_n(\Delta_1))&=\sum_{i+j+k=n-1} V(\hat{g}'_n, U_{i+n,j,k})\\ &\leqslant \tfrac{1}{4}\sum_{i+j+k=n-1}(A_2^k A_3^j X)_{i+n-1,j,k}.\end{aligned} \tag{70}$$

Similarly we have

$$V(\hat{g}^2_n, U_n(\Delta_2))\leqslant \tfrac{1}{4}\sum_{i+j+k=n-1}(A_1^k A_3^i X)_{i,j+n-1,k}, \tag{71}$$

$$V(\hat{g}^3_n, U_n(\Delta_3))\leqslant \tfrac{1}{4}\sum_{i+j+k=n-1}(A_1^j A_2^i X)_{i,j,k+n-1}, \tag{72}$$

$$V(\hat{g}^4_n, U_n(\Delta_4))\leqslant \tfrac{1}{4}\sum_{i+j+k=n-1}(A_1^i A_2^j A_3^k X)_{n-1-i,n-1-j,n-1-k}. \tag{73}$$

From (70)–(73) and Lemma 4 (with b replaced by X and n by $n-1$) we have

$$\sum_{r=1}^{4} V(\hat{g}_n^r, U_n(\Delta_r)) \leqslant \sum_{i+j+k=n-1} X_{2i,2j,2k} = V(\hat{f}_n, U_n(T))$$

by (68). ∎

Proof of Theorem 3. Proceeding in a manner similar to that in the proof of Theorem 1, we see that there is a constant C such that

$$V(\hat{g}_n, \Delta) \leqslant CV(B_n(g), \Delta) \tag{74}$$

for all triangles Δ similar to T and all functions g on Δ.

Now take $m \geqslant 1$ and any triangle Δ in Ω_m. We write $B_n(f)|\Delta = q(\Delta) + r(\Delta)$, where $q(\Delta) \in \Pi_1$ and all first-order derivatives of $r(\Delta)$ vanish at some point in Δ. Then

$$V(B_n(f), U_n(\Delta)) = V(q(\Delta), U_n(\Delta)) + o(1/m^2). \tag{75}$$

Now choose a function g on Δ such that $B_n(f)|\Delta = B_n(g, \Delta)$. Then, putting $q(\Delta) = q$,

$$\begin{aligned} V(q(\Delta), U_n(\Delta)) &= V(\hat{q}_n, U_n(\Delta)) \\ &\leqslant V(\hat{g}_n, U_n(\Delta)) + V(q-g)\hat{}_n, \Delta) \\ &\leqslant V(\hat{g}_n, U_n(\Delta)) + CV(B_n(q-g), \Delta) \end{aligned} \tag{76}$$

by (74). Now

$$B_n(q-g, \Delta) = q - B_n(g, \Delta) = q - B_n(f) = -r(\Delta). \tag{77}$$

Combining (75), (76), and (77) gives

$$V(B_n(f), U_n(\Delta)) \leqslant V(\hat{g}_n(\Delta), U_n(\Delta)) + o(1/m^2). \tag{78}$$

Now

$$V(B_n(f), \Delta) = \frac{2n}{n+1} V(B_n(f), U_n(\Delta)) + o\left(\frac{1}{m^2}\right). \tag{79}$$

So by (78) and (79)

$$V(B_n(f), T) = \sum_{\Delta \in \Omega_m} V(B_n(f), \Delta) \leqslant \frac{2n}{n+1} \sum_{\Delta \in \Omega_m} V(\hat{g}_n(\Delta), U_n(\Delta)) + o(1). \tag{80}$$

If $m = 2^s$ we can apply Lemma 9 successively to give

$$\sum_{\Delta \in \Omega_m} V(\hat{g}_n(\Delta), U_n(\Delta)) \leq V(\hat{f}_n, U_n(T)). \tag{81}$$

Combining (80) and (81) and letting $s \to \infty$ then completes the proof of Theorem 3. ∎

We remark that Theorem 3 can also be proved by a similar method to the proof of (9) in Section 1. However for consistency we have proved all three results by the same method of subdivision.

Finally we give an example to show that in general we do not have $V(B_n(f), T) \leq V(\hat{f}_n, T)$. Let T be an equilateral triangle with vertices $T_1 = (0, \sqrt{3})$, $T_2 = (-1, 0)$, $T_3 = (1, 0)$. Take $n = 2$ and let f be any function on T satisfying $f(T_2) = 1$, $f(T_3) = -1$, $f(T_1) = f(\frac{1}{2}(T_1 + T_2)) = f(\frac{1}{2}(T_2 + T_3)) = f(\frac{1}{2}(T_3 + T_1)) = 0$. Then $B_n(f)(u, v, w) = v^2 - w^2$ and $B_n(f)(x, y) = xy/\sqrt{3} - x$. A straightforward calculation then shows that $V(B_n(f), T) = \sinh^{-1}(1/\sqrt{3}) + \frac{2}{3}$. It is easily seen that $V(\hat{f}_n, T) = 1$ and so $V(B_n(f), T) > V(\hat{f}_n, T)$.

References

1. G. Chang and P. J. Davis, The convexity of Bernstein polynomials over triangles, *J. Approx. Theory* **40** (1984), 11–28.
2. G. Farin, "Bézier Polynomials over Triangles and the Construction of Piecewise C^r Polynomials," TR/91, Dept. of Mathematics, Brunel University, Uxbridge, Middlesex, England, 1980.
3. G. G. Lorentz, "Bernstein Polynomials," Univ. of Toronto Press, Toronto, 1953.
4. C. A. Micchelli, On a numerically efficient method for computing multivariate B-splines, *in* "Multivariate Approximation Theory" (W. Schempp and K. Zeller, Eds.), pp. 211–248, Birkhaüser, Basel, 1979.
5. G. Pólya and I. J. Schoenberg, Remark on the de la Vallée–Poussin means and convex conformal maps of the circle, *Pacific J. Math.* **8** (1958), 295–334.