B-form basics

C. de Boor

0. Introduction; notation This paper lists the essential facts about the representation of polynomials in m variables as **Bernstein polynomials**. An expanded version may appear elsewhere.

While univariate Bernstein polynomials are well studied - see, e.g., Lorentz' classical book Lorentz (1953), - the multivariate version has only attracted attention sporadically. Lorentz' book devotes just one page to the two most direct generalizations: the tensor product or coordinate degree generalization, and the total degree generalization which is the topic of the present paper.

Motivation for the paper comes from computer-aided geometric design where, through the initiative of de Casteljau and Bézier, the Bernstein polynomials of mostly one variable have become the main tool for the representation and computational use of **pp** (:= piecewise polynomial) functions. Farin's work Farin (1979), Farin (1980) brought popularity and understanding to the use of bivariate Bernstein polynomials, and my own understanding starts from that work. My own interest has been started and repeatedly reinforced by work with smooth pp functions in two or more variables (de Boor and Höllig (1983), (1986)), in which their representation in terms of Bernstein polynomials, i.e., their **B-net**, for short, plays an essential role, since it reflects so nicely, and far better than other standard representations, the interplay between the geometry of the underlying triangular partition and the smoothness requirements.

For the sake of brevity, and since there are several people and ideas responsible, I am proposing here the term **B-form** (and correspondingly, B-net) for what would, more properly, be called the **barycentric-Bernstein-de Casteljau-Bézier-Farin-** $\cdots$ **-form**. I apologize to de Casteljau and Farin and $\cdots$ for the slight they might feel.

While the bivariate and trivariate situation is of most practical interest, I have chosen here to record the facts in the general m-dimensional context. This forces careful consideration of notation and brings out the essential mathematical aspects and surprising beauty of the B-form.

I will adhere to the following notational conventions: I won't bother with boldface, arrows, or underlines to distinguish points in $\mathbb{R}^m$ from other objects. The j-th component of a point $x \in \mathbb{R}^m$ I will denote by $x(j)$ (rather than x_j). I will use standard multi-index notation throughout. Thus

$$x^\alpha := x(1)^{\alpha(1)} x(2)^{\alpha(2)} \cdots x(m)^{\alpha(m)}$$

with $\alpha \in \mathbb{Z}^m$, i.e., α an m-vector with integer entries. The **normalized monomial** is so handy a function that it deserves a special symbol:

$$[\![x]\!]^\alpha := x^\alpha/\alpha! = \prod_{j=1}^{m} x(j)^{\alpha(j)}/\alpha(j)!, \tag{0.1}$$

hence $[\![x]\!]^\alpha = [\![x(1)]\!]^{\alpha(1)} [\![x(2)]\!]^{\alpha(2)} \cdots [\![x(m)]\!]^{\alpha(m)}$, with the conventions

$$[\![z]\!]^n = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n < 0. \end{cases}$$

In these terms, the **multinomial theorem** (for the power of a vector sum) takes the very simple form

$$[\![x+y+\cdots+z]\!]^{\alpha} = \sum_{\xi+\upsilon+\cdots+\zeta=\alpha} [\![x]\!]^{\xi}[\![y]\!]^{\upsilon}\cdots[\![z]\!]^{\zeta},$$

whose proof by induction on the **length**

$$|\alpha| := \sum_{j=1}^{m} \alpha(j)$$

of α is immediate. It is possible (and a useful exercise) to build the entire discussion of the B-form on this identity.

The normalization of the monomials used here also makes differentiation neat. With

$$D^{\alpha} := D_1^{\alpha(1)} D_2^{\alpha(2)} \cdots D_m^{\alpha(m)}$$

and $D_j f$ the partial derivative of $f : \mathbb{R}^m \to \mathbb{R}$ with respect to its j-th argument, we have

$$D^{\alpha}[\![\ ]\!]^{\beta} = [\![\ ]\!]^{\beta-\alpha}.$$

Hence

$$D^{\alpha}[\![0]\!]^{\beta} = \delta_{\alpha\beta}, \tag{0.2}$$

showing that $\{[\![\ ]\!]^{\alpha} : |\alpha| \le k, \alpha \in \mathbb{Z}_+^m\}$ is linearly independent. Their span

$$\pi_k := \operatorname{span}\{[\![\ ]\!]^{\alpha} : |\alpha| \le k, \alpha \in \mathbb{Z}_+^m\}$$

is, by definition, the collection of all polynomials of degree $\le k$. We conclude that

$$\dim \pi_k = \#\{[\![\ ]\!]^{\alpha} : |\alpha| \le k, \alpha \in \mathbb{Z}_+^m\} = \binom{m+k}{k}. \tag{0.3}$$

While the **power form**

$$p = \sum_{|\alpha| \le k} [\![\ ]\!]^{\alpha} c(\alpha)$$

(with $c(\alpha) = \big(D^{\alpha}p\big)(0)$, by (0.2)) is the standard mathematical representation for $p \in \pi_k$, it is not suited for work with pp functions since it provides explicit information only about the behavior of p near 0. By contrast, the B-form (with respect to some $(m+1)$-set $V \subset \mathbb{R}^m$) provides explicit information about the behavior of p at all the faces of the **convex hull**

$$[V]$$

of V. This makes it appropriate for the representation of smooth multivariate pp functions over a triangular partition.

1. Linear interpolation A set V of $m+1$ points in IR^m is said to be **in general position** in case every linear polynomial on IR^m can be written in terms of its values on V, i.e.,

$$\forall p \in \pi_1 \qquad p \;=\; \sum_{v\in V} \xi_v p(v)\ . \tag{1.1}$$

There are equivalent definitions of this term. E.g., V is in general position in case its affine hull is all of IR^m, or, in case the simplex $[V]$ is proper, or if the $m+1$ vectors $(v|1), v \in V$ in IR^{m+1} are linearly independent, etc. But I stick with the above definition since it uses the property of immediate interest here. In this way we associate with each such V $m+1$ linear polynomials $\xi_v, v \in V$, characterized by the fact that (1.1) holds. In particular, with p the constant polynomial, we find that

$$1 \;=\; \sum_{v\in V} \xi_v, \tag{1.2}$$

while, with $p : x \mapsto x(j)$, we get

$$x(j) \;=\; \sum_{v\in V} \xi_v(x) v(j)$$

for $j = 1, \cdots, m$, hence

$$\forall x \qquad x \;=\; \sum_{v\in V} \xi_v(x) v. \tag{1.3}$$

We conclude that the vector

$$\xi(x) \;:=\; \big(\xi_v(x)\big)_{v\in V} \tag{1.4}$$

provides the **barycentric coordinates** for x with respect to V. Note that I have chosen here to use the points in V (rather than the integers from 0 to m, or from 1 to $m+1$) to index the components of the vector $\xi(x)$. This unorthodox notation is more to the point since it does not impose some artificial order on the **vertices** v; it also simplifies notation.

Since dim $\pi_1 = m+1$, we conclude from (1.1) that $\big(\xi_v\big)_{v\in V}$ is a basis for π_1. In particular, the representation (1.1) is unique. Therefore

$$p \;=\; \sum_{v\in V} \xi_v a(v) \;\implies\; \forall v \in V \quad p(v) = a(v). \tag{1.5}$$

We conclude that

$$\xi_w(v) \;=\; \delta_{wv}, \tag{1.6}$$

hence

$$\mathrm{affine}(V\backslash w) \;=\; \ker\xi_w := \{x \in \mathrm{IR}^m : \xi_w(x) = 0\}. \tag{1.7}$$

2. Definition of the B-form The B-form for $p \in \pi_k$ is a somewhat unexpected generalization of the linear interpolation formula (1.1), viz.

$$p \;=\; (\xi E)^k c(0). \tag{2.1}$$

Here, c is a **mesh function** and ξE is a **difference operator**, and the formula is to be read as an instruction: "Apply the difference operator ξE k times, starting with the mesh function c, then evaluate the resulting mesh function at the mesh point 0."

This definition of the B-form is unorthodox. In effect, I propose here to use de Casteljau's algorithm (de Casteljau (1959)) as the definition, and to derive the other properties of the B-form from this algorithm. Implicit in this is the claim that this provides a more efficient path to these properties than standard approaches.

Consider now this definition more explicitly. The c appearing in (2.1) is a **meshfunction**, i.e., defined on the mesh of nonnegative integer points

$$\alpha := \big(\alpha(v)\big)_{v\in V} \in \mathbb{Z}_+^V \ ,$$

and the difference operator ξE acts on mesh functions by the rule

$$(\xi E)c(\alpha) := \sum_{v\in V} \xi_v c(\alpha + e_v), \tag{2.2}$$

with e_v the unit vector given by $e_v(w) := \delta_{vw}$.

If $k = 0$, then we are to apply the difference operator no times, i.e., then

$$p = c(0).$$

If $k = 1$, then we are to apply the difference operator one time, i.e, then

$$p = \sum_{v\in V} \xi_v c(e_v).$$

This is just (1.1) again, in slightly changed notation, i.e., $c(e_v) = p(v), v \in V$.

If $k = 2$, then we are to apply the difference operator two times, i.e., then

$$p = \sum_{v\in V}\sum_{w\in V} \xi_v \xi_w c(e_v + e_w).$$

For general k,

$$p = \sum_{u\in V}\sum_{v\in V}\cdots\sum_{w\in V} \xi_u\xi_v\cdots\xi_w \ \ c(e_u + e_v + \cdots + e_w), \tag{2.3}$$

and this shows that **the function p given by (2.1) is a polynomial of degree $\le k$** since it shows that p is a linear combination of products of k linear polynomials. This also shows that, for the purpose of the definition (2.1), we only need to know c at meshpoints of the form

$$e_u + e_v + \cdots + e_w$$

involving exactly k summands, i.e., at all mesh points $\alpha \in \mathbb{Z}_+^V$ with

$$|\alpha| := \sum_{v\in V} \alpha(v) = k.$$

Writing (2.3) in terms of distinct meshpoints $\alpha \in \mathbb{Z}_+^V$, we get

$$p = \sum_{|\alpha|=k} B_\alpha c(\alpha), \tag{2.4}$$

with

$$B_\alpha := |\alpha|! [\![\xi^\alpha]\!] = \binom{|\alpha|}{\alpha} \xi^\alpha$$

and

$$[\![\xi]\!]^\alpha := \prod_{v\in V} [\![\xi_v]\!]^{\alpha(v)} = \prod_{v\in V} {\xi_v}^{\alpha(v)}/\alpha(v)!.$$

The **multinomial coefficient** $\binom{|\alpha|}{\alpha} = |\alpha|!/\prod_{v\in V}\alpha(v)!$ appears here in accordance with the multinomial theorem, but its precise value is not important here. The only thing that matters is that there are exactly

$$\#\{\alpha \in \mathbb{Z}_+^V : |\alpha| = k\} = \#\{\beta \in \mathbb{Z}_+^m : |\beta| \le k\} = \dim \pi_k$$

summands in (2.4) . Hence, necessarily

$$\bigl(B_\alpha\bigr)_{|\alpha|=k} \textit{ and } \bigl(\xi^\alpha\bigr)_{|\alpha|=k} \textit{ are both bases for } \pi_k,$$

provided we can convince ourselves that **every** $p \in \pi_k$ **can be written in the form (2.3) or (2.4)**. But that is not hard to do. Observe that every $p \in \pi_k$ can be written as a sum of products of k linear polynomials, – e.g.,

$$\forall |\beta| \le k \quad x^\beta = \Bigl(\prod_{\beta(j)>0} \bigl(x(j)\bigr)^{\beta(j)}\Bigr)(1)^{k-|\beta|},$$

– and each linear polynomial can be written as a linear combination of the ξ_v as in (1.1). Thus, for every $p \in \pi_k$, there is a collection of k-tuples $(q, r, \cdots, s)$ of linear polynomials so that

$$\begin{aligned} p &= \sum_{(q,r,\cdots,s)} qr\cdots s \\ &= \sum_{(q,r,\cdots,s)} \bigl(\sum_{u\in V} \xi_u q(u)\bigr)\bigl(\sum_{v\in V} \xi_v r(v)\bigr)\cdots\bigl(\sum_{w\in V} \xi_w s(w)\bigr) \\ &= \sum_{u,v,\cdots,w} \xi_u\xi_v\cdots\xi_w \sum_{(q,r,\cdots,s)} q(u)r(v)\cdots s(w) \end{aligned}$$

which shows our claim.

In particular, the representation (2.4) is unique, i.e., for each $p \in \pi_k$, there is exactly one choice for $c(\alpha), |\alpha| = k$, so that (2.1) (or, equivalently (2.4)) holds.

Since $\sum_{|\alpha|=k} B_\alpha = \bigl(\sum_{v\in V} \xi_v\bigr)^k = 1$, the basis $\bigl(B_\alpha\bigr)_{|\alpha|=k}$ forms a **partition of unity**, and this partition of unity is nonnegative on $[V]$ since the barycentric coordinates are nonnegative there.

Further, B_α **is unimodal on** $[V]$**, and the coefficient** $c(\alpha)$ **has its biggest influence on** p **in** $[V]$ **at the point**

$$v_\alpha := \sum_{v \in V} v\alpha(v)/|\alpha| \tag{2.5}$$

since B_α, and equivalently ξ^α, **takes its maximum over** $[V]$ **at** v_α. For this, and many other, reasons, the pointset

$$\{\big(v_\alpha, c(\alpha)\big) \in \mathbb{R}^{m+1} : |\alpha| = k, \alpha \in \mathbb{Z}_+^V\} \tag{2.6}$$

in $\mathbb{R}^{m+1}$ is given special attention. It is called the **B-net** for p. The points which make up the B-net are often called the **Bézier control points** for p.

The point of the formulation (2.1) is to avoid having to deal with expressions like (2.3) and to operate, calculate and reason directly with the simple expression (2.1) according to the operations it prescribes. The next two sections may make this clearer.

3. Evaluation of the B-form Evaluation of the B-form (2.1) at some point x requires k-fold application of the difference operator $\xi(x)E$. Since we are only interested in $\big(\xi(x)E\big)^k c$ at the meshpoint 0, we only need to apply the difference operator "at" certain meshpoints. Precisely, we calculate

$$c_1(\alpha) := \big(\xi(x)E\big)c(\alpha) = \sum_{v \in V} \xi_v(x) c(\alpha + e_v) \quad \text{for} \quad |\alpha| = k-1$$

and this only requires knowledge of $c(\alpha)$ for $|\alpha| = k$. Then we calculate

$$c_2(\alpha) := \big(\xi(x)E\big)c_1(\alpha) = \sum_{v \in V} \xi_v(x) c_1(\alpha + e_v) \quad \text{for} \quad |\alpha| = k-2$$

and this only requires knowledge of $c_1(\alpha)$ for $|\alpha| = k-1$. In this way, we calculate

$$c_j(\alpha) := \big(\xi(x)E\big)c_{j-1}(\alpha) = \sum_{v \in V} \xi_v(x) c_{j-1}(\alpha + e_v) \quad \text{for} \quad |\alpha| = k-j$$

for $j = 1, \cdots, k$ (and with $c_0 = c$), and the final calculation gives

$$p(x) = c_k(0) = \big(\xi(x)E\big)c_{k-1}(0) = \sum_{v \in V} \xi_v(x) c_{k-1}(e_v).$$

In this description, I used different meshfunctions c_j to avoid confusion. But, since c_j is only considered and generated at meshpoints α with $|\alpha| = k - j$, we might as well use the same letter c for all of them. Thus, the evaluation amounts to generating the whole $(m+1)$-simplex

$$c(\alpha), |\alpha| \le k,$$

of numbers from its base

$$c(\alpha), |\alpha| = k,$$

by the calculation

$$\begin{aligned} &\text{for } j = 1, \cdots, k, \text{ do:} \\ &\quad c(\alpha) = \big(\xi(x)E\big)c(\alpha) = \sum_{v\in V} \xi_v(x)c(\alpha + e_v), \ |\alpha| = k - j. \end{aligned} \tag{3.1}$$

While the base of this discrete simplex never changes, the layers built upon it do depend on x. Its apex, $c(0)$, provides the desired number $p(x)$. We will see later that the numbers $c(\alpha), |\alpha| < k$, generated here also provide useful information about p.

Figure 3 shows the meshpoints of interest for the case $m = 2$, i.e., the **bivariate** case. In this case, the mesh points have **three** components, corresponding to the fact that a two-dimensional simplex has three vertices. Correspondingly, each of the mesh point layers $|\alpha| = k - j$ of interest forms a triangle in this case, and the total set forms a tetrahedron. For $m = 1$, the meshpoints of interest would form a triangle. This is quite familiar from the evaluation of univariate polynomials, e.g., from their Newton form in which one also generates a triangular array of numbers.

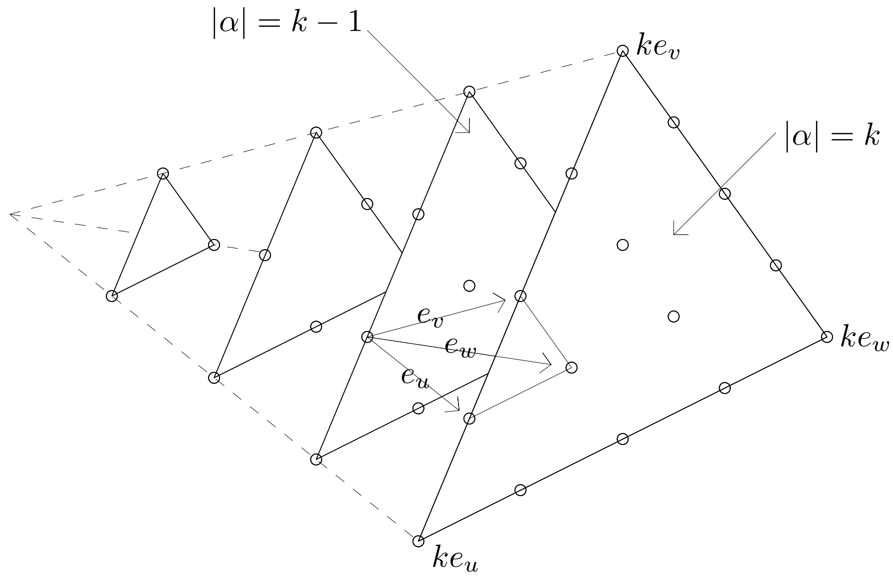

Figure 3. The meshpoint simplex for evaluation

Figure 3 also shows the typical stencil of the difference operator which, for the calculation of

$$\big(\xi(x)E\big)c(\beta) = \sum_{v\in V} \xi_v(x)c(\beta + e_v),$$

requires one to go from meshpoint β into each of the $m+1$ coordinate directions, picking up the value of c at the $m + 1$ meshpoints $\beta + e_v, v \in V$ reached, multiplying that value with the number $\xi_v(x)$ and then summing these products over v to obtain the value of $\big(\xi(x)E\big)c$ at β. Recall that $\sum_{v\in V} \xi_v(x) = 1$, hence application of the difference operator $\xi(x)E$ always amounts to **averaging**. If $x \in [V]$, then this average is proper since then $\xi(x) \geq 0$. Thus,

$$p(x) \textit{ is a convex combination of } \{c(\alpha), |\alpha| = k\} \textit{ in case } x \in [V]. \tag{3.2}$$

As an example, consider the calculation of $p(w)$ for some $w \in V$. Since $\xi(w) = e_w$, the difference operator simplifies in this case to

$$\big(\xi(w)E\big)c(\beta) = c(\beta + e_w),$$

i.e., we merely pick up the value at the next meshpoint in the w-direction. We conclude that therefore

$$p(w) = c(ke_w) \quad \text{for} \quad w \in V. \tag{3.3}$$

As another example, consider the calculation of $p(x)$ for some $x \in [V \backslash u]$. Now $\xi_u(x) = 0$, hence

$$\big(\xi(x)E\big)c(\beta) = \sum_{v \in (V \backslash u)} \xi_v(x)c(\beta + e_v).$$

We conclude that, in this case, $p(x)$ is a convex combination of just the coefficients $c(\alpha)$ with $|\alpha| = k$ and $\alpha(u) = 0$. More generally,

$$x \in [V \setminus U] \quad \Longrightarrow \quad p(x) \in [c(\alpha) : |\alpha| = k, \text{ supp } \alpha \subset (V \setminus U)] \tag{3.4}$$

To put it differently: **For $W \subset V$, the coefficients for the B-form** (with respect to W) **of the restriction of p to the affine hull of W are provided by the restriction of c to the corresponding mesh "simplex"** $\{\alpha \in \mathbb{Z}_+^V : |\alpha| = k, \text{ supp } \alpha \subset W\}$. In these terms, (3.3) provides the extreme case $W = \{w\}$.

4. Differentiation of the B-form The **directional derivative** $D_y f$ of the function f on $\mathbb{R}^m$ in the direction y is, by definition, given by the rule

$$\big(D_y f\big)(x) := \lim_{t \to 0} \frac{f(x + ty) - f(x)}{t}.$$

Hence, in terms of the partial derivatives,

$$D_y = \sum_{j=1}^{m} y(j)D_j.$$

(This would suggest the alternative notation yD for the operator D_y or of E_ξ for ξE.)

It requires nothing more than the chain rule to differentiate the B-form:

$$D_y p = D_y \big(\xi E\big)^k c(0) = \big(\xi E\big)^{k-1} k \big(D_y \xi E\big) c(0). \tag{4.1}$$

Since $\xi(x)$ is the unique solution of the linear system

$$\sum_v \xi_v(x) = 1, \qquad \sum_v \xi_v(x)v = x,$$

we have $\xi(x+ty)-\xi(x)=t\eta(y)$, with $\eta(y)$ the unique solution of

$$\sum_v \eta_v(y)=0, \qquad \sum_v \eta_v(y)v=y.$$

E.g.,

$$\eta(w-v)=e_w-e_v, \quad w\in V\backslash v. \tag{4.2}$$

Thus, **with**

$$c_y := \big(\eta(y)E\big)c, \tag{4.3a}$$

we obtain explicitly the B-form

$$D_y p = \big(\xi E\big)^{k-1} k c_y(0) \tag{4.3b}$$

for the polynomial $D_y p$. Therefore, at a vertex,

$$D_y p(v) = kc_y\big((k-1)e_v\big) = k\big(\eta(y)E\big)c\big((k-1)e_v\big), \quad v\in V.$$

In particular, from (4.2),

$$D_{w-v}p(v) = \frac{c\big((k-1)e_v+e_w\big)-c(ke_v)}{1/k}, \quad w\in V\backslash v,$$

hence **the** $m+1$ **distinct points**

$$\{\big(v_\alpha, c(\alpha)\big)\in \mathbb{R}^{m+1} : \alpha=(k-1)e_v+e_w, w\in V\}$$

with

$$v_\alpha := \sum_{v\in V} v\alpha(v)/|\alpha| \tag{4.4}$$

all lie on the tangent plane to p **at** v and therefore determine that plane. This is a further indication of the usefulness of the **B-net** for p, which, to recall from (2.6), is the collection of points

$$\{\big(v_\alpha, c(\alpha)\big)\in \mathbb{R}^{m+1} : |\alpha|=k, \alpha\in \mathbb{Z}_+^V\}.$$

Higher derivatives are obtained by iteration of this process. If $Y\subset \mathbb{R}^m\setminus 0$ contains r points, then

$$D_Y p := \big(\prod_{y\in Y} D_y\big)p = \frac{k!}{(k-r)!}\big(\xi E\big)^{k-r} c_Y(0) \tag{4.5a}$$

with

$$c_Y(\alpha) := \prod_{y\in Y}\big(\eta(y)E\big)c(\alpha), \quad |\alpha|=k-r. \tag{4.5b}$$

The evaluation of such a derivative at some point proceeds just as the evaluation of p itself, except that different difference operators are to be used at different stages. The order in which these are applied is immaterial since all (constant coefficient) difference operators commute.

5. A Taylor formula By the binomial theorem and the fact that $\xi(x+y) = \xi(x)+\eta(y)$,

$$p(x+y) = \big(\xi(x+y)E\big)^k c(0) = \sum_{r=0}^{k} \binom{k}{r} \big(\xi(x)E\big)^{k-r} \big(\eta(y)E\big)^r c(0). \qquad (5.1)$$

Since

$$\binom{k}{r} \big(\xi(x)E\big)^{k-r} \big(\eta(y)E\big)^r c(0) = \frac{1}{r!} \big({D_y}^r p\big)(x)$$

by (4.5), we recognize in (5.1) the standard Taylor formula

$$p(x+y) = p(x) + \big(D_y p\big)(x) + \frac{1}{2}\big({D_y}^2 p\big)(x) + \cdots$$

6. Invariance under an affine change of variables Any affine change of variables leaves the B-form unchanged. More precisely, if also W is an $(m+1)$-point set in $\mathbb{R}^m$ in general position, and f is an affine map so that $f(W) = V$, then the B-form for the polynomial $q := p \circ f$ with respect to W has again c as its coefficient sequence, in the following sense. With $\xi'(x)$ the barycentric coordinates of x with respect to W,

$$q(x) = p(f(x)) = \big(\xi(x)E\big)^k c(0) = \big(\xi'(x)E\big)^k c^f(0),$$

where

$$c^f(\alpha) := c(\alpha \circ f^{-1}).$$

7. Integration of the B-form It is possible to express the integral of p over $[V]$ quite simply in terms of the B-form:

$$\int_{[V]} p = \frac{\mathrm{vol}_m[V]}{\binom{k+m}{k}} \sum_{|\alpha|=k} c(\alpha). \qquad (7.1)$$

In effect, this identity claims that $\int_{[V]} B_\alpha$ only depends on $|\alpha|$. The quickest way to see this is to realize that

$$\int_{[V]} B_\alpha = \int_{[W]} 1,$$

with W the $(m+1+k)$-set in $\mathbb{R}^{m+k}$ generated from V and α by thinking of V as a subset of $\mathbb{R}^{m+k}$ and adjoining to it, for each $v \in V$, $\alpha(v)$ points of the form (v, e_j), with the unit vectors $e_j \in \mathbb{R}^k$ different from point to point.

8. Product of two B-forms The B-form of the product of two polynomials is obtainable from their B-forms with the aid of a few factorials:

$$\big(\xi E\big)^k c(0)\big(\xi E\big)^h d(0) = \big(\xi E\big)^{k+h} c * d(0)$$

with

$$c * d(\gamma) := \sum_{\alpha+\beta=\gamma} c(\alpha)d(\beta)C_{\alpha,\beta} \tag{8.1}$$

and

$$C_{\alpha,\beta} := \frac{\binom{|\alpha|}{\alpha}\binom{|\beta|}{\beta}}{\binom{|\alpha+\beta|}{\alpha+\beta}},$$

since

$$B_\alpha B_\beta = \binom{|\alpha|}{\alpha}\xi^\alpha \binom{|\beta|}{\beta}\xi^\beta = C_{\alpha,\beta}B_{\alpha+\beta}.$$

The special choice $d(\beta) = 1, |\beta| = h$, gives the formula

$$\big(\xi E\big)^k c(0) = \big(\xi E\big)^{k+h} c'(0)$$

with

$$c'(\gamma) := \sum_{\alpha+\beta=\gamma} c(\alpha)C_{\alpha,\beta}, \quad |\gamma| = k+h.$$

The particular case $h = 1$ of such **degree raising** has received special attention.

9. Degree raising Observe that

$$[\![\xi]\!]^\alpha = \big(\sum_v \xi_v\big)[\![\xi]\!]^\alpha = \sum_v \big(\alpha(v)+1\big)[\![\xi]\!]^{\alpha+e_v}.$$

Hence

$$\begin{aligned}\sum_{|\alpha|=k} B_\alpha c(\alpha) &= k! \sum_{|\alpha|=k}\sum_v \big(\alpha(v)+1\big)[\![\xi]\!]^{\alpha+e_v} c(\alpha)\\ &= k! \sum_{|\alpha|=k+1} [\![\xi]\!]^\alpha \sum_v \alpha(v)c(\alpha-e_v).\end{aligned}$$

Conclude that

$$\sum_{|\alpha|=k} B_\alpha c(\alpha) \;=\; \sum_{|\alpha|=k+1} B_\alpha\big(Rc\big)(\alpha) \tag{9.1}$$

with

$$\big(Rc\big)(\alpha) \;:=\; \sum_{v\in V} c(\alpha-e_v)\alpha(v)/|\alpha|. \tag{9.2}$$

Note that (9.1-2) requires knowledge of $c(\beta)$ for β with $|\beta| = k$ and with, possibly, a negative entry. This presents no difficulty, though, since all such values are multiplied by zero, hence are not really needed.

Since $v_\alpha = \sum_{v\in V} v_{\alpha-e_v}\alpha(v)/|\alpha|$, formula (9.2) can be interpreted as **linear interpolation** at the point v_α by the plane or linear polynomial through the points

$$\bigl(v_\beta, c(\beta)\bigr) \quad \text{with} \quad \beta = \alpha - e_v, v \in V.$$

This draws further attention to the **control polytope** for p, i.e., the piecewise linear function obtained from the B-net

$$\{\bigl(v_\alpha, c(\alpha)\bigr) : |\alpha| = k\}$$

by local linear interpolation.

10. The Bernstein polynomial The **Bernstein polynomial for f of order k with respect to V** is, by definition, the particular polynomial

$$B_k f := \sum_{|\alpha|=k} B_\alpha f(v_\alpha). \tag{10.1}$$

The Bernstein polynomial provides an approximation to f which, on $[V]$, converges uniformly to f as $k \to \infty$ in case f is continuous (cf. Lorentz (1953), p.51). The convergence is monotone in case f is **V-convex** in the sense that each of the univariate functions $t \mapsto f(x + t(v - w))$ is convex (Berens (1976)). Moreover, in this case, $B_k f$ is also V-convex. But $B_k f$ need not be convex even if f is (Stancu (1959), Berens (1976), Chang and Davis (1984)).

The B-form $\bigl(\xi E\bigr)^k c(0)$ for $p \in \pi_k$ (with respect to V) provides the essential information about any f for which $p = B_k f$. We have

$$p = B_k f \quad \Longleftrightarrow \quad \forall |\alpha| = k \quad f(v_\alpha) = c(\alpha). \tag{10.2}$$

The simplest such functions f are the **control polytopes** for p, i.e., the piecewise linear interpolants to the data

$$\bigl(v_\alpha, c(\alpha)\bigr), |\alpha| = k. \tag{10.3}$$

For this reason, we denote any such control polytope by

$$B_k^{-1} p. \tag{10.4}$$

I have used the plural here advisedly since, for $m > 2$, there are several equally reasonable piecewise linear interpolants, as has been rightfully stressed and detailed by Dahmen and Micchelli in [13]. Different interpolants differ in how the points $\{v_\alpha : |\alpha| = k\}$ are connected to produce a **triangulation**, i.e., a partition into simplices, for the simplex $[V]$. The typical triangulation is obtained by choosing an ordering $v_0, v_1, \ldots, v_m$ of the vertex set V, thus obtaining the directions $d_i := v_i - v_{i-1}$, $i = 1, \ldots, m$. The corresponding triangulation with meshpoints $\{v_\alpha : |\alpha| = k\}$ for V consists of all simplices in $[V]$ of the form

$$\sigma_{\alpha,q} := v_\alpha + [0, d_{q(1)}, d_{q(1)} + d_{q(2)}, \ldots, d_{q(1)} + \cdots + d_{q(m)}]/k$$

with $|\alpha| = k$ and q a permutation of the first m integers. Thus,

$$\sigma_{\alpha,q} = [v_{\alpha_0}, \ldots, v_{\alpha_m}],$$

with

$$\alpha_0 := \alpha,$$
$$\alpha_j := \alpha_{j-1} + e_{v_{q(j)}} - e_{v_{q(j)-1}}, \quad j = 1, \ldots, m.$$

A simplex may appear in the triangulation only for certain orderings of V and not for others. To see this, observe that two points v_α and v_β will be vertices for the same simplex if and only if their difference can be written as a sum of some of the vectors d_j/k. If we order the entries of α and β to correspond to the ordering of the vertex set used, writing, e.g., $\alpha(j)$ instead of $\alpha(v_j)$, this means that either $\beta - \alpha$ or else $\alpha - \beta$ must be writable as a sum of distinct vectors of the form $e_j - e_{j-1}$. For example, for $m = 3$ and $k = 2$, the two vertices $v_{(1,0,0,1)}$ and $v_{(0,1,1,0)}$ are not connected by a meshline (in the triangulation corresponding to the ordering used), while the two vertices $v_{(0,1,0,1)}$ and $v_{(1,0,1,0)}$ are. This shows that the reordering v_1, v_0, v_2, v_3 would connect the former and disconnect the latter.

On the other hand, the simplices

$$[v_{\beta - e_v} : v \in V] \quad \text{with } |\beta| = k + 1$$

involved in degree raising are part of any such triangulation since, in terms of the particular ordering used,

$$[v_{\beta - e_v} : v \in V] = \sigma_{\alpha,q}$$

with $\alpha = \beta - e_{v_m}$ and $q(j) = m + 1 - j,\ j = 1, \ldots, m$. Thus, regardless of the particular ordering of the vertex set V used, the resulting piecewise linear interpolant $B_k^{-1}p$ to the data (10.3) will agree with $B_{k+1}^{-1}p$ at the basepoints $v_\beta, |\beta| = k + 1$, as we saw in §9. This implies, by induction, that the Lipschitz constant (over $[V]$) for any $B_{k+n}^{-1}p, n > 0$, is no bigger than that for $B_k^{-1}p$, hence the sequence $\left(B_{k+n}^{-1}p\right)$ has uniform limit points. It implies further that $B_{k+n}^{-1}p$ converges pointwise to some function f as $n \to \infty$, hence f is the uniform limit of $B_{k+n}^{-1}p$. But this limit function is necessarily p since

$$p = B_{k+n}B_{k+n}^{-1}p = B_{k+n}f + B_{k+n}(B_{k+n}^{-1}p - f) \longrightarrow f, \text{ as } n \to \infty$$

using the facts that $B_{k+n}f$ converges to f and $\|B_{k+n}(B_{k+n}^{-1}p - f)\| \le \|B_{k+n}^{-1}p - f\| \to 0$.

Since local linear interpolation preserves convexity, we conclude that p is convex in case its control polytope $B_k^{-1}p$ is.

11. Boundary behavior We now come to the heart of the B-form. We consider how to extract from the B-form of p information about its behavior on the boundary of the simplex $[V]$.

The boundary of $[V]$ is made up of faces, i.e., of convex hulls of subsets of V. For any $W \subset V$, we call $[W]$ the W**-face** of $[V]$. The (proper) faces of highest dimension are the **facets** of $[V]$. We find it convenient to call the $(V\backslash w)$-face of $[V]$ the w**-facet** of $[V]$. In other words, we identify the faces of the simplex by the set of vertices contained in them, but identify a facet by the sole vertex not contained in it.

Recall from Section 3 that, on the W-face, p is entirely determined by $c(\alpha)$ with supp $\alpha \subset W$. Recall from Section 4 that the tangent plane for p at the vertex w is entirely determined by $c(\alpha)$ with $\alpha = (k-1)e_w + e_v, v \in V$. We can describe this last set also as

$$\{\alpha : |\alpha|_{V\backslash w} \le 1\},$$

using the abbreviation

$$|\alpha|_W := \sum_{v\in W} \alpha(v).$$

This makes it easy to recognize both of these facts as special cases of the following theorem.

Theorem *All derivatives of p of order $\le \rho$ on the W-face are determined by*

$$c(\alpha), |\alpha|_{\backslash W} \le \rho. \tag{11.1}$$

If the W-face in question is a facet, say the w-facet, then the coefficients involved are those $c(\alpha)$ with $\alpha(w) \le \rho$. In the whole coefficient-"simplex", these occupy layers $0, 1, \cdots, \rho$ "parallel" to the w-facet, i.e., the layers $c(\alpha), \alpha(w) = j$, with $j = 0, 1, \cdots, \rho$.

For the general W-face, the relevant coefficients are those no more than ρ steps away from the corresponding coefficient "facet" $c(\alpha)$, supp $\alpha \subset W$.

For the **proof**, observe that the Theorem's claim is equivalent to the statement that p vanishes $\rho+1$-fold on $[W]$ iff $c(\alpha) = 0$ for $|\alpha|_{\backslash W} \le \rho$. But this follows from the fact that $|\alpha|_{\backslash W} + 1$ gives the order to which ξ^α vanishes on $[W]$.

In terms of the B-net

$$b_p := \{(v_\alpha, c(\alpha)) : \alpha \in \mathbb{Z}_+^V\}$$

for p introduced in Section 4, the Theorem states that knowing all derivatives of order $\le \rho$ on the W-face is the same as knowing b_p on all v_α within ρ steps from that face. It is part of the attraction of the B-net that it makes such neat geometric statements possible.

For the application of this theorem to the problem of smoothly fitting together polynomial pieces, we must be prepared to express two such polynomial pieces in B-form with respect to the same simplex. In approaching this problem, we give another proof of the Theorem. The approach makes use of the polynomials whose B-form with respect to V is part of the B-form for p.

12. The subpolynomials The evaluation of $p \in \pi_k$ and its derivatives from the B-form proceeds by repeated **differencing**. It is a remarkable fact that this differencing is **uniform**. Regardless of the meshpoint at which it is applied, the difference operator is the same. This implies that, **during the calculation of some information about p, we are simultaneously computing the same information for** a whole host of polynomials, viz. **all polynomials whose B-form coefficients** (with respect to V) **form a subsimplex of those for p.** These are the polynomials

$$p_\alpha := (\xi E)^{k-|\alpha|} c(\alpha), \quad |\alpha| \le k. \tag{12.1}$$

For $|\alpha| = k$, p_α is the constant $c(\alpha)$, while, at the other extreme, $p_0 = p$.

Since the B-form coefficients for such a p_α form a subsimplex of the coefficient simplex for p, the B-form coefficients of its derivative $D_Y p_\alpha$ form a corresponding subsimplex of those for $D_Y p$, up to a scalar factor. Precisely, from (4.5), with $Y \subset \mathrm{IR}^m \backslash 0$ and $r := \#Y$, and $|\alpha| \le k - r$,

$$D_Y p_\alpha \;=\; \frac{(k-|\alpha|)!}{(k-|\alpha|-r)!} \big(\xi E\big)^{k-|\alpha|-r} c_Y(\alpha) \tag{12.2}$$

with

$$c_Y := \Big(\prod_{y\in Y} \eta(y)E\Big)c, \tag{12.3}$$

while

$$D_Y p \;=\; D_Y p_0 \;=\; \frac{k!}{(k-r)!}\big(\xi E\big)^{k-r} c_Y(0).$$

This shows that, on the W-face, $D_Y p_\alpha$ is determined by the numbers

$$c_Y(\alpha+\beta),\ \operatorname{supp} \beta \subset W,\ |\beta| = k - |\alpha| - r, \tag{12.4}$$

while $D_Y p$ is determined there by the numbers

$$c_Y(\gamma),\ \operatorname{supp} \gamma \subset W,\ |\gamma| = k - r. \tag{12.5}$$

Now note that (12.4) is a subset of (12.5) exactly when $\operatorname{supp} \alpha \subset W$. Since p_α is of degree $\le k - |\alpha|$, this implies that, for any α with $\operatorname{supp} \alpha \subset W$, we know **all** derivatives of p_α on the W-face, as soon as we know there all derivatives of p of order $\le k - |\alpha|$. But, knowing all the derivatives of a polynomial even at just one point determines that polynomial entirely. This proves the following

Proposition *Each p_α depends linearly on p and its derivatives of order $\le k - |\alpha|$ on* $[\operatorname{supp} \alpha]$.

Conversely, (12.5) is the union of all the sets (12.4) with $\operatorname{supp} \alpha \subset W$. This proves the following restatement of Theorem 11.

Theorem *Let $p, q \in \pi_k$, $W \subset V$. Then*

$$\forall (Y \subset \mathrm{IR}^m\backslash 0, \#Y \le r) \quad D_Y p = D_Y q \text{ on the } W\text{-face} \quad \Longleftrightarrow \quad \forall (|\alpha| = k - r, \operatorname{supp} \alpha \subset W) \quad p_\alpha = q_\alpha.$$

13. Change of V The subpolynomials p_α introduced in the preceding section depend on V. This is reflected in the notation since, after all, α is defined on V. But, by Proposition 12, p_α depends, more precisely, only on the points in $\operatorname{supp} \alpha$. To say it differently:

Proposition *If also V' is an $(m+1)$-set in IR^m in general position, and $\alpha \in \mathbb{Z}_+^V$ has its support in $V \cap V'$, then*

$$p_\alpha = p_{\alpha'},$$

with

$$\alpha' : V' \to \mathbb{Z}_+ : v \mapsto \begin{cases} \alpha(v), & \text{if } v \in V \cap V'; \\ 0, & \text{otherwise.} \end{cases}$$

This is so because, by the proposition, p_α only depends on p and its derivatives on $[\operatorname{supp} \alpha]$. This suggests the identification of any two α, α' which agree on their support, and we will follow this suggestion from now on. In effect, we think of α as defined at all the vertices that might enter the discussion, but to be zero on all but at most $m+1$ of them.

This makes it easy to describe the change of V, i.e., the derivation of the B-form $p =: \big(\xi' E\big) c'(0)$ for p with respect to V' from the B-form with respect to V. It is sufficient to consider the case

$$V' = (V \backslash w) \cup w',$$

since an arbitrary V' can be reached from V as the $(m+1)$-st in a chain of $(m+1)$-sets whose neighbors only differ by one point.

The crucial observation is the following. The coefficient $c'(\alpha)$ is the extreme coefficient (associated with the vertex w') for the subpolynomial p_β with $\beta = \alpha - \alpha(w') e_{w'}$. In other words,

$$c'(\alpha) = p_\beta(w'), \text{ with } \beta := \alpha - \alpha(w') e_{w'}. \tag{13.1}$$

On the other hand, $\operatorname{supp} \beta \subset (V \backslash w)$, hence

$$p_\beta = \big(\xi E\big)^{\alpha(w')} c(\beta). \tag{13.2}$$

This implies that p_β is evaluated during the course of evaluation of p from its B-form with respect to V. Specifically, we find $c'(\alpha) = p_\beta(w')$ at position β in the $(m+1)$-simplex $c(\beta)$, $|\beta| \le k$, generated during the evaluation of p at w', i.e.,

$$c'(\alpha) = c(\alpha - \alpha(w') e_{w'}), \quad \text{for } |\alpha| = k. \tag{13.3}$$

In fact, since the evaluation of p at w' from the B-form with respect to V proceeds without any special attention paid to the vertex w, it follows that we are generating simultaneously the B-form coefficients for p with respect to every one of the $(m+1)$ sets V' obtainable from V by an exchange of some $w \in V$ for w'. This provides a **subdivision algorithm**. Choosing w' somewhere in $[V]$, we obtain a triangulation of $[V]$ into at most $m+1$ nontrivial simplices $[V_w]$, with $V_w := (V \backslash w) \cup w'$, and the coefficient simplex for the B-form of p with respect to V_w is to be found in the w-facet of the $(m+1)$-simplex $c(\beta), |\beta| \le k$ generated during the evaluation of p at w'.

14. Smoothness across an interface The matching of derivatives of polynomial pieces across an interface between two simplices is easily described in terms of the subpolynomials associated with that interface, since these describe completely the behavior of a polynomial and its derivatives on that interface, by Theorem 12. The precise statement of the smoothness conditions is made quite simple by our agreement to think of meshpoints α as defined on all vertices that might appear in the discussion, with its value usually zero, with at most $m+1$ exceptions.

Theorem *Let $p, q \in \pi_k$, let $\rho \le k$, and let V, V' be the vertex sets of two simplices in a triangulation. Then the pp function*

$$f : [V] \cup [V'] \to \mathrm{IR} : x \mapsto \begin{cases} p(x), & \text{if } x \in [V]; \\ q(x), & \text{if } x \in [V'], \end{cases}$$

is in C^ρ if and only if

$$\forall (\mathit{supp}\ \alpha \subset V \cap V', |\alpha| = k - \rho) \quad p_\alpha = q_\alpha. \tag{14.1}$$

If V and V' differ by just one point,

$$V' = (V \backslash w) \cup w',$$

say, and $q = (\xi' E) c'(0)$ is the B-form for q with respect to V', then the condition (14.1) reads more explicitly

$$\forall (\text{supp}\ \alpha \subset (V \backslash w), k - \rho \le |\alpha| \le k) \quad \big(\xi(w')E\big)^{k-|\alpha|} c(\alpha) = c'(\alpha + (k - |\alpha|) e_{w'}). \tag{14.2}$$

Note that **these conditions are independent of** k and depend on α only in the sense that the weights in the linear relations between c and c' in (14.2) depend on ρ or $k - |\alpha|$, i.e., on the order of the derivatives being constrained to be continuous. This means that, in studying a linear system of such conditions across one or more neighboring facets, we can choose k at will, e.g., $k = \rho$.

In general, C^ρ-continuity across the w-facet of $[V]$ imposes conditions which connect $c(\alpha)$ for $\alpha(w) \le \rho$ with $c'(\alpha)$ for $\alpha(w') \le \rho$. The form (14.2) makes explicit that this involves exactly

$$\#\{\alpha \in \mathbb{Z}_+^V : \alpha(w) = 0, k - \rho \le |\alpha| \le k\}$$

linearly independent conditions, i.e., exactly as many conditions as there are degrees of freedom in p and its directional derivatives of order $\le \rho$ on the w-facet in some fixed direction transversal to that facet.

15. The B-net Let Δ be a triangulation of some domain in IR^m. This means that Δ consists of simplices δ, with the intersection $\delta \cap \delta'$ of any two always a face (possibly the empty face) of both of them.

I denote by V_δ the vertex set of the simplex $\delta \in \Delta$, and by

$$V := \cup_{\delta \in \Delta} V_\delta$$

the totality of the vertices of simplices of Δ. Denote by

$$A = A_{k,\Delta} := \{\alpha \in \mathbb{Z}_+^V : |\alpha| = k, \exists \delta \in \Delta \quad \text{supp}\ \alpha \subset V_\delta\}$$

the corresponding set of index meshpoints of interest. In words, these are elements of $\mathbb{Z}_+^V$, i.e., defined on V and with nonnegative integer entries. In addition, each $\alpha \in A$ has support only on some V_δ and has length $|\alpha| = k$.

Consider now the space

$$S := \pi^{\rho}_{k,\Delta}$$

of pp functions of degree $\le k$ on the triangulation Δ and in C^ρ. This means that each $f \in S$ agrees on each simplex in Δ with some polynomial of degree $\le k$, and these polynomial pieces fit together to form a function with r continuous derivatives.

Consider specifically

$$S_0 := \pi^{0}_{k,\Delta},$$

the space of continuous pp functions (of degree $\le k$) on Δ. Since two of its polynomial pieces on neighboring simplices fit together continuously exactly when their B-form coefficients associated with the common face coincide, it is possible to describe an element f of S_0 by the meshfunction c defined on the mesh A and providing in

$$c(\alpha),\ \operatorname{supp}\alpha \subset V_\delta,$$

the B-form coefficients for the polynomial piece $f_{|\delta}$ with respect to the vertex set V_δ of δ.

The **B-net** for such f is, by definition, the collection of points

$$\bigl(v_\alpha, c(\alpha)\bigr), \quad \alpha \in A,$$

with

$$v_\alpha := \sum_{v\in V} v\alpha(v)/|\alpha|, \quad \alpha \in A.$$

While it is satisfactory to deal with the meshfunction c, the B-net reflects more explicitly the geometry of the situation. We think of the B-net as the function

$$b_f : V_A \to \mathbb{R} : v_\alpha \mapsto c(\alpha)$$

on the discrete set

$$V_A := \{v_\alpha : \alpha \in A\},$$

which is a subset of the domain of $f \in S_0$. The values of this discrete function at all the points in some face of some δ determine f on that face. In particular,

$$f(v) = c(ke_v) = b_f(v), \quad v \in V.$$

Further, C^ρ-continuity of f is equivalent to certain linear relations involving b_f on points at most ρ layers away from the facets of the δ. For example, C^1-continuity is equivalent to having each $(m+2)$-tuple

$$\bigl(v_\beta, b_f(v_\beta)\bigr),\ \beta = \alpha + e_v,\ v \in W$$

lie on a plane, with W the vertices of any two simplices having a facet in common, and $|\beta| = k-1$ with support only on the vertices common to both simplices. This localizes the effect of such continuity conditions as much as possible.

30nov09 Corrected some misprints (in (3.3) and in the last display before Section 6), and provided the missing Figure 3.

References

H. Berens (1976), Über Bernsteinpolynome in mehreren Veränderlichen, talk given at the April 1976 Oberwolfach meeting organized by Zeller and Schempp.

S. Bernstein (1912-3), Démonstration du théorème de Weierstrass, fondée sur le calcul des probabilités, Comm. Soc. Math. Kharkov(2) **13** , 1-2.

P. Bézier (1972), **Numerical Control, - Mathematics and Applications**, John Wiley & Sons, London, xvi+240 pages.

W. Böhm, G. Farin and J. Kahmann (1984), A survey of curve and surface methods in CAGD, Computer Aided Geometric Design **1**, 1-60.

C. de Boor and K. Höllig (1983), Approximation order from bivariate C^1-cubics: A counterexample, Proc. Amer. Math. Soc. **87** , 649-655.

C. de Boor and K. Höllig (1986), Approximation order from smooth bivariate pp functions, ms.

P. de Casteljau (1959), Courbes et surfâces a pôles, André Citroen Automobiles SA, Paris.

G.-z. Chang and P. J. Davis (1984), A new proof for the convexity of Bernstein-Bézier surfaces over triangles, J.Approximation Theory **40** , 11-28.

W. Dahmen and C. A. Micchelli (1985), Convexity of multivariate Bernstein polynomials and box spline surfaces, Res.Report RC 11176 (#50344), IBM, Yorktown Heights, NY.

G. Farin (1979), Subsplines über Dreiecken, Dissertation, Braunschweig, West Germany.

G. Farin (1980), Bézier polynomials over triangles and the construction of piecewise C^r-polynomials, TR/91, Department of Mathematics, Brunel University, Uxbridge, Middlesex, UK.

G. G. Lorentz (1953), **Bernstein Polynomials**, University of Toronto Press, Toronto, Canada, x+130 pages.

D. D. Stancu (1959), De l'approximation, par des polynômes du type Bernstein, des fonctions de deux variables, Comm.Akad. R.P.Romine **9**, 773-777.