Multivariate Bernstein polynomials and convexity

by

Thomas Sauer

Abstract

It is well known that in two or more variables Bernstein polynomials do not preserve convexity. Here we introduce two variations, one stronger than the classical notion, the other one weaker, which are preserved. Moreover, a weaker sufficient condition for the monotony of subsequent Bernstein polynomials is given.

§1 Introduction

Consider m+1 points $\mathbf{p}_0, \ldots, \mathbf{p}_m \in \mathbb{R}^d$ in general position; i.e., the vectors $\mathbf{p}_k - \mathbf{p}_0, k = 0, \ldots, m$, are linearly independent, in the course of which d has to be greater than or equal to m. A point \mathbf{p} in the affine hull of $\mathbf{p}_0, \ldots, \mathbf{p}_m$ can be uniquely written as

$$\mathbf{p} = \sum_{k=0}^{m} u_k \mathbf{p}_k, \quad \text{where } u_0 + \dots + u_m = 1.$$

The coefficients of $\mathbf{u} = (u_0, \ldots, u_m) \in \mathbb{R}^{m+1}$ are called the *barycentric coor*dinates of \mathbf{p} with respect to $\mathbf{p}_0, \ldots, \mathbf{p}_m$. Moreover, the points of the simplex $[\mathbf{p}_0, \ldots, \mathbf{p}_m]$, spanned by the vertices $\mathbf{p}_0, \ldots, \mathbf{p}_m$ have nonnegative barycentric coordinates and vice versa. In other words, we can identify $[\mathbf{p}_0, \ldots, \mathbf{p}_m]$ with

$$S_m = \{ \mathbf{u} = (u_0, \dots, u_m) : u_k \ge 0, \, u_0 + \dots + u_m = 1 \}$$

by using the 1–1 mapping which associates each point of the simplex with its uniquely defined barycentric coordinates.

We consider $\mathbb{S}_m = \{(x_1, \ldots, x_m) : x_k \ge 0, x_1 + \cdots + x_m \le 1\}$ to be the standard unit simplex in \mathbb{R}^m . If f is a function defined on an arbitrary m-dimensional simplex, especially on \mathbb{S}_m , it will prove quite useful to consider f as a function defined over S_m by the use of barycentric coordinates.

Let $\mathbf{i} = (i_0, \ldots, i_m) \in \mathbb{N}_0^{m+1}$ be a multiindex with $|\mathbf{i}| = i_0 + \cdots + i_m = n$, and let $\mathbf{u} \in S_m$. The **i**-th *Bernstein (base-) polynomial* of degree *m* is defined by

$$B_{\mathbf{i}}^{n}(\mathbf{u}) = \frac{n!}{\mathbf{i}!}\mathbf{u}^{\mathbf{i}} = \frac{n!}{i_{0}!\cdots i_{m}!}u_{0}^{i_{0}}\cdots u_{m}^{i_{m}}.$$

Given any control polyhedron $\Phi = {\mathbf{b}_{\mathbf{i}} : |\mathbf{i}| = n} \subset \mathbb{R}^d$ we define the associated Bernstein - Bézier - polynomial $\mathbf{b}^n : S_m \to \mathbb{R}^d$ of degree n via

$$\mathbf{u} \mapsto \mathbf{b}^n(\mathbf{u}) = \sum_{|\mathbf{i}|=n} \mathbf{b}_{\mathbf{i}} B_{\mathbf{i}}^n(\mathbf{u}).$$

In this paper we will be especially interested in polynomials, associated with functions $f: S_m \to \mathbb{R}$, as follows:

$$\mathbf{b}^{n}[f](\mathbf{u}) = \sum_{|\mathbf{i}|=n} f\left(\frac{\mathbf{i}}{n}\right) B_{\mathbf{i}}^{n}(\mathbf{u})$$

where $\frac{\mathbf{i}}{n} = (\frac{i_0}{n}, \dots, \frac{i_m}{n})$ are properly defined barycentric coordinates. This polynomial, known as the *Bernstein polynomial* of f of degree n, was introduced by *Dinghas* [4] and *Lorentz* [6] in 1951, independently.

Dealing with functions defined over an arbitrary simplex, it has shown to be convenient to make use of directional derivatives because no information is needed about the exact position of the simplex and its vertices. A *direction* \mathbf{d} in S_m , is given by the difference of two points in S_m ; i.e., $\mathbf{d} = \mathbf{u} - \mathbf{v}$, $\mathbf{u}, \mathbf{v} \in S_m$. According to this the *directional derivative* of a function f with respect to \mathbf{d} is defined in the following way:

$$D_{\mathbf{d}}f(\mathbf{u}) = \lim_{t \to 0} \frac{f(\mathbf{u} + t\mathbf{d}) - f(\mathbf{u})}{t} = \sum_{k=0}^{m} d_k \frac{\partial}{\partial u_k} f(\mathbf{u}).$$

Some special, but nevertheless important directions are given by $\hat{\mathbf{e}}_k = \mathbf{e}_k - \mathbf{e}_0$, where $\mathbf{e}_0, \ldots, \mathbf{e}_m$ denote the unit vectors in \mathbb{R}^{m+1} ; considered as barycentric coordinates they form the vertices of the simplex S_m . If we identify S_m with the *m*-dimensional unit simplex \mathbb{S}_m , the directional derivatives $D_{\hat{\mathbf{e}}_k}$ coincide with the standard partial derivatives $\frac{\partial}{\partial x_k}$. As a consequence we can replace multiple partial derivatives by

$$D^{\mathbf{r}}f(\mathbf{u}) = D^{r_1}_{\hat{\mathbf{e}}_1} \cdots D^{r_m}_{\hat{\mathbf{e}}_m} f(\mathbf{u}),$$

where $\mathbf{r} = (r_1, \ldots, r_m) \in \mathbb{N}_0^m$.

The main tool managing these partial derivatives of Bernstein polynomials is given by the forward difference operator, which is well-known, but remains overlooked in the CAGD-literature up to now. Forward differences, according to a nonnegative multiindex $\mathbf{r} = (r_1, \ldots, r_m) \in \mathbb{N}_0^m$ are inductively defined as follows:

$$\begin{aligned} \Delta^{0}_{\hat{\mathbf{e}}_{k}} \mathbf{b}_{\mathbf{i}} &= \mathbf{b}_{\mathbf{i}}, \\ \Delta^{r}_{\hat{\mathbf{e}}_{k}} \mathbf{b}_{\mathbf{i}} &= \Delta^{r-1}_{\hat{\mathbf{e}}_{k}} \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_{k}} - \Delta^{r-1}_{\hat{\mathbf{e}}_{k}} \mathbf{b}_{\mathbf{i}}, \\ \Delta^{\mathbf{r}} \mathbf{b}_{\mathbf{i}} &= \Delta^{r_{1}}_{\hat{\mathbf{e}}_{1}} \cdots \Delta^{\mathbf{r}_{m}}_{\hat{\mathbf{e}}_{m}} \mathbf{b}_{\mathbf{i}} \end{aligned}$$

where **i** is a multiindex with $i_0 \geq |\mathbf{r}|$ to guarantee that $\Delta^{\mathbf{r}}$ is well defined. Sometimes we will use the abbreviation $\Delta_{j,k}$ to mean $\Delta^{\mathbf{e}_j+\mathbf{e}_k}$ for $1 \leq j,k \leq m$. Now we are in position to represent the derivatives $D^{\mathbf{r}}$ of a Bernstein polynomial in terms of differences $\Delta^{\mathbf{r}}$; indeed,

$$D^{\mathbf{r}}\mathbf{b}^{n}(\mathbf{u}) = \frac{n!}{(n-r)!} \sum_{|\mathbf{i}|=n-|\mathbf{r}|} \Delta^{\mathbf{r}}\mathbf{b}_{\mathbf{i}+|\mathbf{r}|\mathbf{e}_{0}} B_{\mathbf{i}}^{n-|\mathbf{r}|}(\mathbf{u})$$
(1)

which is proved by using the well-known identity

$$D_{\mathbf{d}}B_{\mathbf{i}}^{n}(\mathbf{u}) = n \sum_{k=0}^{m} d_{k}B_{\mathbf{i}-\mathbf{e}_{k}}^{n-1}(\mathbf{u})$$

as well as

$$D_{\hat{\mathbf{e}}_{k}} \mathbf{b}^{n}(\mathbf{u}) = \sum_{|\mathbf{i}|=n} \mathbf{b}_{\mathbf{i}} n \left(B_{\mathbf{i}-\mathbf{e}_{k}}^{n-1}(\mathbf{u}) - B_{\mathbf{i}-\mathbf{e}_{0}}^{n-1}(\mathbf{u}) \right)$$
$$= n \sum_{|\mathbf{i}|=n-1} \left(\mathbf{b}_{\mathbf{i}+\mathbf{e}_{k}} - \mathbf{b}_{\mathbf{i}+\mathbf{e}_{0}} \right) B_{\mathbf{i}}^{n-1}(\mathbf{u})$$
$$= n \sum_{|\mathbf{i}|=n-1} \Delta_{\hat{\mathbf{e}}_{k}}^{1} \mathbf{b}_{\mathbf{i}+\mathbf{e}_{0}} B_{\mathbf{i}}^{n-1}(\mathbf{u}).$$

It is sometimes nescessary and helpful to write a Bernstein polynomial of degree $\leq n$ in terms of the base polynomials $B_{\mathbf{i}}^{n+1}$ of degree n + 1. The formula, needed in this context is known as *degree raising*. We have

$$\sum_{\mathbf{i}|=n} \mathbf{b}_{\mathbf{i}} B_{\mathbf{i}}^{n} = \sum_{|\mathbf{j}|=n+1} \hat{\mathbf{b}}_{\mathbf{j}} B_{\mathbf{j}}^{n+1}, \text{ where } \hat{\mathbf{b}}_{\mathbf{j}} = \sum_{k=0}^{m} \frac{j_{k}}{n+1} \mathbf{b}_{\mathbf{j}-\mathbf{e}_{k}};$$
(2)

for a proof see [5].

§2 Convexity

A function $f : S_m \to \mathbb{R}$ is said to be *convex* if for every two points $\mathbf{u}, \mathbf{v} \in S_m$ and every $\lambda \in [0, 1]$

$$f(\lambda \mathbf{u} + (1 - \lambda)\mathbf{v}) \le \lambda f(\mathbf{u}) + (1 - \lambda) f(\mathbf{v});$$

it is known that for twice differentiable functions this is equivalent to the statement that the Hessian $H[f](\mathbf{u})$ is *positive semi-definite* for every $\mathbf{u} \in S_m$; i.e., for each direction $\mathbf{d} = (d_1, \ldots, d_m)$ (remember: every direction can be written as linear combination $d_1\hat{\mathbf{e}}_1 + \cdots + d_m\hat{\mathbf{e}}_m$)

$$\mathbf{d} H[f](\mathbf{u}) \mathbf{d}^T \ge 0$$

Using this characterization and the endpoint interpolation property of Bernstein polynomials; i.e., $\mathbf{b}^{n}(\mathbf{e}_{k}) = \mathbf{b}_{n\mathbf{e}_{k}}$ (see e.g. [5]), we obtain at once that for every convex Bernstein polynomial the matrices

$$H_{\mathbf{i}} = \begin{pmatrix} \Delta^{(2,0,\dots,0)} \mathbf{b}_{\mathbf{i}} & \dots & \Delta^{(1,0,\dots,0,1)} \mathbf{b}_{\mathbf{i}} \\ \vdots & \ddots & \vdots \\ \Delta^{(1,0,\dots,0,1)} \mathbf{b}_{\mathbf{i}} & \dots & \Delta^{(0,\dots,0,2)} \mathbf{b}_{\mathbf{i}} \end{pmatrix}$$
(3)

have to be positive semi-definite for all $\mathbf{i} = (n-2)\mathbf{e}_k + 2\mathbf{e}_0$, $k = 0, \dots, m$, since $H[\mathbf{b}^n](\mathbf{e}_k) = H_{(n-2)\mathbf{e}_k+2\mathbf{e}_0}$.

On the other hand, we are also in a position to give a sufficient condition for convexity, which is due to the *Chang and Feng* [2]: the Bernstein polynomial \mathbf{b}^n is convex if the matrices H_i are positive semi-definite for every \mathbf{i} with $i_0 \geq 2$.

The proof is quite easy: for a direction $\mathbf{d} = (d_1, \ldots, d_m)$ we consider

$$\mathbf{d} H[\mathbf{b}^n](\mathbf{u}) \mathbf{d}^T = \sum_{|\mathbf{i}|=n-2} \left(\mathbf{d} H_{\mathbf{i}+2\mathbf{e}_0} \mathbf{d}^T \right) B_{\mathbf{i}}^{n-2}(\mathbf{u}).$$

Due to the positive semi-definiteness of the H_i , all coefficients of the righthand polynomial are nonnegative from which at once follows that $H[\mathbf{b}^n]$ is positive semi-definite, hence that \mathbf{b}^n is convex. In his classical book on Bernstein polynomials [6] Lorentz pointed out that in one variable Bernstein polynomials preserve many properties of the associated functions, among others convexity. This does not remain valid in two or more variables as the simple function $|u_1 - u_2|$ shows where for m = 2the matrix H_{ne_0} takes on the form

$$\frac{1}{n} \left(\begin{array}{cc} 0 & -2 \\ -2 & 0 \end{array} \right)$$

which is not positive semi-definite, in contradiction to the nescessary condition stated above.

This counterexample was first given in 1975 by Schmid¹ in [7], a similar example was later also considered by *Chang and Davis* [1].

§3 Axial convexity

As convexity seemed somehow inappropriate for dealing with multivariate Bernstein polynomials, the concept of axial convexity was introduced in [7] as a variant of classical convexity. Stepping from one to higher dimensions, it proves to be the right choice: indeed, axial convexity is preserved by Bernstein polynomials and, although much weaker than convexity, it also proves to be a sufficient condition for the monotony of subsequent Bernstein polynomials; i.e., $\mathbf{b}^{n-1}[f] \geq \mathbf{b}^n[f]$.

A function $f: S_m \to \mathbb{R}$ is called *axially convex*, if f is convex with respect to the directions $\mathbf{e}_k - \mathbf{e}_j$, $0 \le j < k \le m$, on S_m , that is

$$f(\lambda \mathbf{u} + (1 - \lambda)\mathbf{v}) \le \lambda f(\mathbf{u}) + (1 - \lambda) f(\mathbf{v})$$

for $\lambda \in [0, 1]$ and all \mathbf{u}, \mathbf{v} with $\mathbf{u} - \mathbf{v} = \mu (\mathbf{e}_k - \mathbf{e}_j)$ and suitable $\mu, 0 \leq j < k \leq m$. This property is obviously weaker than convexity. As an example for a function which is axially convex, but not convex serves $\mathbf{u} \mapsto -u_1 u_2$.

¹At the conference on multivariate approximation in Oberwolfach, Black Forest, in April 1976 H. Berens gave a talk about multidimensional Bernstein polynomials in which he discussed various results of H.J. Schmid and himself. Among others he pointed out that two- and higher- dimensional Bernstein polynomials do not preserve convexity, but that axially convex functions, see the section below, have axially convex Bernstein polynomials of all orders. The definition of and the results on axial convexity, given below, go back to Schmid in the case of two variables.

First we will give a characterization of axially convex functions, generalizing Schmid's two-dimensional results.

Proposition 1. A continuous function $f : S_m \to \mathbb{R}$ is axially convex if and only if the following inequalities hold for all $n \in \mathbb{N}$

$$\Delta_{k,k} f\left(\frac{\mathbf{i}}{n}\right) \ge 0 \tag{4}$$

and

$$\Delta_{j,k}^{s} f\left(\frac{\mathbf{i}}{n}\right) = \left(\Delta_{j,j} + \Delta_{k,k} - 2\Delta_{j,k}\right) f\left(\frac{\mathbf{i}}{n}\right) \ge 0 \tag{5}$$

where $|\mathbf{i}| = n$ and $1 \le j < k \le m$.

Proof: Convexity in the direction of $\mathbf{e}_k - \mathbf{e}_0$ yields

$$\frac{1}{2}f\left(\frac{\mathbf{i}+2\hat{\mathbf{e}}_k}{n}\right) + \frac{1}{2}f\left(\frac{\mathbf{i}}{n}\right) \ge f\left(\frac{\mathbf{i}+\hat{\mathbf{e}}_k}{n}\right),\tag{6}$$

which can easily be rewritten as (4). Moreover (6) is equivalent to convexity in the direction of $\mathbf{e}_k - \mathbf{e}_0$, since, due to the continuity of f, the curve $t \mapsto f(\mathbf{u} + t\hat{\mathbf{e}}_k)$ is convex if and only if it is *midpoint convex*; i.e., $\frac{1}{2}f(\mathbf{u} + 2t\hat{\mathbf{e}}_k) + \frac{1}{2}f(\mathbf{u}) \ge f(\mathbf{u} + t\hat{\mathbf{e}}_k)$ which yields (6) by choosing $\mathbf{u} = \frac{\mathbf{i}}{n}$ and $t = \frac{1}{n}$. The same argumentation applied to $\mathbf{e}_k - \mathbf{e}_j$ will produce the equivalence to

The same argumentation applied to $\mathbf{e}_k - \mathbf{e}_j$ will produce the equivalence to (5), respectively.

For twice differentiable f another characterization is obtained by the following

Proposition 2. A C^2 – function $f : S_m \to \mathbb{R}$ is axially convex if and only if for all $\mathbf{u} \in S_m$ the following inequalities hold:

$$D_{k,k}f(\mathbf{u}) \ge 0 \tag{7}$$

and

$$D_{j,j}f(\mathbf{u}) + D_{k,k}f(\mathbf{u}) - 2D_{j,k}f(\mathbf{u}) \ge 0$$
(8)

where $D_{j,k}f$ denotes $D_{\hat{\mathbf{e}}_j}D_{\hat{\mathbf{e}}_k}f$ for $1 \leq j < k \leq m$.

The proof is based on the identity

$$\Delta^{\mathbf{r}} f\left(\frac{\mathbf{i}}{n}\right) = \int_{0}^{\frac{1}{n}} \cdots \int_{0}^{\frac{1}{n}} D^{\mathbf{r}} f\left(\frac{\mathbf{i}}{n} + \sum_{k=0}^{m} \sum_{j=1}^{r_{k}} \hat{\mathbf{e}}_{k} t_{j}^{k}\right) dt_{1}^{1} \cdots dt_{m}^{r_{m}}$$
(9)

which is obtained by noting that

$$\frac{\partial}{\partial t}f\left(\frac{\mathbf{i}}{n}+t\hat{\mathbf{e}}_{k}\right)=D_{\hat{\mathbf{e}}_{k}}f\left(\frac{\mathbf{i}}{n}+t\hat{\mathbf{e}}_{k}\right)$$

and

$$\int_{0}^{\frac{1}{n}} D_{\hat{\mathbf{e}}_{k}} f\left(\frac{\mathbf{i}}{n} + t\hat{\mathbf{e}}_{k}\right) dt = \int_{0}^{\frac{1}{n}} \frac{\partial}{\partial t} f\left(\frac{\mathbf{i}}{n} + t\hat{\mathbf{e}}_{k}\right) dt = \Delta_{\hat{\mathbf{e}}_{k}} f\left(\frac{\mathbf{i}}{n}\right),$$

the fundamental step for induction.

Using (9), the equations (4) and (7) are easily proved to be equivalent, as are (5) and (8), too. Combining these results, we get

Theorem 3. If $f : S_m \to \mathbb{R}$ is a continuous and axially convex function, then all Bernstein polynomials $\mathbf{b}^n[f]$, $n \in \mathbb{N}$, are axially convex, too.

Proof: Since f is axially convex, the inequality $\Delta_{k,k} f\left(\frac{\mathbf{i}}{n}\right) \geq 0$ holds for all \mathbf{i} with $i_0 \geq 2$, due to Proposition 1; we have

$$D_{k,k}\mathbf{b}^{n}[f] = \sum_{|\mathbf{i}|=n-2} \Delta_{k,k} f\left(\frac{\mathbf{i}+2\mathbf{e}_{0}}{n}\right) B_{\mathbf{i}}^{n-2} \ge 0;$$

i.e., inequality (7) holds. Similarly, (8) can be deduced, and an application of Proposition 2 completes the proof. $\hfill \Box$

Given an arbitrary vector $\mathbf{h} = (h_0, \ldots, h_m) \in \mathbb{R}^{m+1}$, where $|\mathbf{h}| = h_0 + \cdots + h_m$, we define the difference

$$\nabla_{\mathbf{h}} f(\mathbf{u}) = \sum_{k=0}^{m} \frac{h_k}{|\mathbf{h}|} f\left(\mathbf{u} + \mathbf{h} - |\mathbf{h}|\mathbf{e}_k\right) - f(\mathbf{u})$$
(10)

where the *increment vector* \mathbf{h} is called *permissible*, if $\mathbf{u} + \mathbf{h} - |\mathbf{h}|\mathbf{e}_k \in S_m$ for $0 \le k \le m$. We notice that simple calculations yield the identity

$$\sum_{k=0}^{m} \frac{h_k}{|\mathbf{h}|} \left(\mathbf{u} + \mathbf{h} - |\mathbf{h}| \mathbf{e}_k \right) = \mathbf{u}, \tag{11}$$

where we are facing a special barycentric combination. This difference leads us to another characterization of axial convexity: **Proposition 4.** The function $f: S_m \to \mathbb{R}$ is axially convex iff $\nabla_{\mathbf{h}} f(\mathbf{u}) \geq 0$ for all $\mathbf{u} \in S_m$ and all permissible $\mathbf{h} \in \mathbb{R}^{m+1}$.

Proof: We set $\mathbf{u}_k^0 = \mathbf{u} + \mathbf{h} - |\mathbf{h}|\mathbf{e}_k$ and deal with the following recursion:

$$\mathbf{u}_{k}^{j+1} = \frac{h_{j}}{|\mathbf{h}| - h_{0} - \dots - h_{j-1}} \mathbf{u}_{j}^{j} + \left(1 - \frac{h_{j}}{|\mathbf{h}| - h_{0} - \dots - h_{j-1}}\right) \mathbf{u}_{k}^{j}, \quad (12)$$

where j = 0, ..., m - 1 and k = j + 1, ..., m. For the sake of brevity we set

$$\lambda_j = \frac{h_j}{|\mathbf{h}| - h_0 - \dots - h_{j-1}}$$

so that (12) now reads $\mathbf{u}_k^{j+1} = \lambda_j \mathbf{u}_j^j + (1 - \lambda_j) \mathbf{u}_j^k$. First we notice that $\mathbf{u}_k^0 - \mathbf{u}_l^0 = |\mathbf{h}| (\mathbf{e}_k - \mathbf{e}_l)$, and obtain in addition

$$\mathbf{u}_{k}^{j+1} - \mathbf{u}_{l}^{j+1} = \lambda_{j}\mathbf{u}_{j}^{j} + (1 - \lambda_{j})\mathbf{u}_{k}^{j} - \lambda_{j}\mathbf{u}_{j}^{j} - (1 - \lambda_{j})\mathbf{u}_{l}^{j} = (1 - \lambda_{j})\left(\mathbf{u}_{k}^{j} - \mathbf{u}_{l}^{j}\right)$$

so that two points of the same level, say \mathbf{u}_k^j and \mathbf{u}_l^j , differ only in a multiple of $\mathbf{e}_k - \mathbf{e}_l$, and hence

$$\lambda_j f\left(\mathbf{u}_j^j\right) + (1 - \lambda_j) f\left(\mathbf{u}_k^j\right) \ge f\left(\mathbf{u}_k^{j+1}\right),\tag{13}$$

due to the axial convexity of f.

In the second step we calculate

$$\mathbf{u}_{m}^{m} = \lambda_{m-1} \mathbf{u}_{m-1}^{m-1} + (1 - \lambda_{m-1}) \mathbf{u}_{m}^{m-1} = \\ \vdots \\ = \lambda_{0} \mathbf{u}_{0}^{0} + (1 - \lambda_{0}) \lambda_{1} \mathbf{u}_{1}^{0} + \dots + (1 - \lambda_{0}) \dots (1 - \lambda_{m-1}) \mathbf{u}_{m}^{0},$$

and notice that

$$(1 - \lambda_0) \cdots (1 - \lambda_{k-1}) \lambda_k = \\ = \frac{|\mathbf{h}| - h_0}{|\mathbf{h}|} \cdot \frac{|\mathbf{h}| - h_0 - h_1}{|\mathbf{h}| - h_0} \cdots \frac{h_k}{|\mathbf{h}| - h_0 - \cdots - h_{k-1}} = \frac{h_k}{|\mathbf{h}|}$$

as well as

$$(1 - \lambda_0) \cdots (1 - \lambda_{m-1}) = \frac{|\mathbf{h}| - h_0 - \cdots - h_{m-1}}{|\mathbf{h}|} = \frac{h_m}{|\mathbf{h}|}$$

which yields, in combination with (11),

$$\mathbf{u}_m^m = \sum_{k=0}^m rac{h_k}{|\mathbf{h}|} \mathbf{u}_k^0 = \mathbf{u}_k$$

Using the same argumentation for (13), we finally get

$$f(\mathbf{u}) = f(\mathbf{u}_m^m) \le \sum_{k=0}^m \frac{h_k}{|\mathbf{h}|} f(\mathbf{u} + \mathbf{h} - |\mathbf{h}|\mathbf{e}_k).$$

Thus we proved that for every axially convex f, every $\mathbf{u} \in S_m$ and every permissible \mathbf{h} the difference $\nabla_{\mathbf{h}} f(\mathbf{u})$ is always nonnegative. The equivalence is simply completed by setting $\mathbf{u} = \frac{\mathbf{i} + \hat{\mathbf{e}}_k}{n}$ and choosing $h_k = h_0 = \frac{1}{n}$ which yields (4), or $\mathbf{u} = \frac{\mathbf{i} + \hat{\mathbf{e}}_j + \hat{\mathbf{e}}_k}{n}$ and $h_j = h_k = \frac{1}{n}$ to get (5), respectively. \Box

Finally, we use a special case of (10), namely,

$$\nabla f\left(\frac{\mathbf{i}}{n}\right) = \sum_{k=0}^{m} \frac{i_k}{n} f\left(\frac{\mathbf{i} - \mathbf{e}_k}{n-1}\right) - f\left(\frac{\mathbf{i}}{n}\right)$$

to establish

Theorem 5. If f is axially convex, then $\mathbf{b}^{n-1}[f] \ge \mathbf{b}^n[f]$. **Proof:** According to (2) we only have to calculate

$$\mathbf{b}^{n-1}[f](\mathbf{u}) - \mathbf{b}^{n}[f](\mathbf{u}) = \sum_{|\mathbf{i}|=n} \left(\sum_{k=0}^{m} \frac{i_{k}}{n} f\left(\frac{\mathbf{i} - \mathbf{e}_{k}}{n-1}\right) - f\left(\frac{\mathbf{i}}{n}\right) \right) B_{\mathbf{i}}^{n}(\mathbf{u})$$
$$= \sum_{|\mathbf{i}|=n} \nabla f\left(\frac{\mathbf{i}}{n}\right) B_{\mathbf{i}}^{n}(\mathbf{u}) \ge 0$$

to get the monotone behavior of $\mathbf{b}^{n}[f]$ for axially convex functions.

 \Box

§4 Polyhedral convexity

Polyhedral convexity is another way of dealing with the non-preservation of convexity by Bernstein polynomials, using a stronger variation instead of a weaker one, as done introducing axial convexity. The notion is strongly geometrically motivated, because the convexity of the control polyhedron is something the human eye can perceive, at least in two variables where a twodimensional surface in \mathbb{R}^3 is formed. Otherwise it seemed nothing but natural that the *de Casteljau* – *Algorithm*, using only convex combinations, should produce nothing else but a convex patch, when applied to a convex polyhedron, as it was pointed out by *Chang and Davis* [1] for bivariate Bernstein polynomials at least. We will see that *m*-dimensional polyhedrons provide some hidden traps for someone who wishes to get close to their convexity.

Given a function $f: S_m \to \mathbb{R}$, the evaluation polyhedron associated to f is defined as the piecewise linear function $\mathcal{L}[f]: S_m \to \mathbb{R}$, given by the vertices

$$\mathcal{L}[f]\left(\frac{\mathbf{i}}{n}\right) = f\left(\frac{\mathbf{i}}{n}\right).$$

Since a piecewise linear function has to be considered over a simplicial dissection of the parameter space, this definition fails to be sufficient for the case $m \geq 3$, as the points $\frac{\mathbf{i}}{n}$, defining the vertices (or knots) of the dissection of S_m , where each vertex is joined to its neighbours $\frac{\mathbf{i}+\mathbf{e}_k-\mathbf{e}_j}{n}$, $0 \leq j < k \leq m$, leave "holes"; i.e., *m*-dimensional polyhedrons with vertices $\frac{\mathbf{i}+\hat{\mathbf{e}}_j+\hat{\mathbf{e}}_k}{n}$, $0 \leq j, k \leq m$. These polyhedrons have $\frac{m(m+1)}{2}$ vertices, so that they are no simplices for $m \geq 3$. The construction of triangulations of S_m , completing the definition of $\mathcal{L}[f]$, is discussed by *Dahmen and Micchelli* [3], but we shall see that there is no real need for it.

Nevertheless it is clear that a polyhedron given by $\Phi = \{\mathbf{b_i} \in \mathbb{R} : |\mathbf{i}| = n\}$ is convex, independently of any triangulations, iff all the subpolyhedrons, given by the vertices $\mathbf{b_i}$, $\mathbf{b_{i+\hat{e}_k}}$ and $\mathbf{b_{i+\hat{e}_j+\hat{e}_k}}$, where $|\mathbf{i}| = n$, $i_0 \ge 2$, $1 \le j, k \le m$ are convex. So it is sufficient to examine only polyhedrons of that type, the vertices denoted by $\mathbf{b}_{j,k}$, $0 \le j, k \le m$ where $\mathbf{b}_{j,k} = \mathbf{b_{i+\hat{e}_j+\hat{e}_k}}$, and $\hat{\mathbf{e}}_0 = \mathbf{e}_0 - \mathbf{e}_0 = 0$. Similarly, we will denote by $\mathbf{u}_{j,k} \in S_m$ the points $\frac{\mathbf{i+\hat{e}_j+\hat{e}_k}}{n}$ in the parameter space, obtaining $\mathbf{b}_{j,k} = f(\mathbf{u}_{j,k})$, so that we can associate $\mathbf{b}_{j,k}$ with $\mathbf{u}_{j,k}$.

Indepently of any triangulation, the vertices $\mathbf{u}_{j,k}$, $0 \leq j \leq m$, form m + 1 simplices σ_k in S_m , defining affine linear functions $\psi_k : S_m \to \mathbb{R}$ by $\psi_k(\mathbf{u}_{j,k}) = \mathbf{b}_{j,k}$. So convexity yields that all $\mathbf{b}_{j,l}$ have to be positioned atop of ψ_k ; i.e., $\mathbf{b}_{j,l} \geq \psi_k(\mathbf{u}_{j,l}), 0 \leq j, k, l \leq m$. From this we obtain

$$\mathbf{b}_{j,j} + \mathbf{b}_{k,l} \ge \mathbf{b}_{j,k} + \mathbf{b}_{j,l},\tag{14}$$

which can be transformed into

$$\Delta_{k,k} \mathbf{b}_{\mathbf{i}} \ge \Delta_{i,k} \mathbf{b}_{\mathbf{i}} \ge 0, \tag{15}$$

being true for $0 \le j < k \le m$ and **i** with $i_0 \ge 2$.

But how to proceed with the "hole"? Since all triangulations are equivalent, there are two possibilities to call Φ convex: either if there exists a triangulation such that the resulting polyhedron is convex, or we want Φ to process convex polyhedrons under arbitrary triangulations. The first way was recommended by *Dahmen and Micchelli* [3], who then prove that this property remaines valid under degree elavation and who proclaime that convex polyhedrons in this first sense would guarantee convexity of the Bernstein polynomials, due to the uniform convergence of degree elavated polyhedrons against \mathbf{b}^n . We shall see below that this does not hold even for m = 3, and that we will have to deal with the second, much more restrictive form of a convex polyhedron to avoid contradictions.

To give a counterexample to the claim in [3] we will consider the case m = 3 more thoroughly where for n = 2 the following figure appears:



Here we have, with respect to symmetry, three possibilities to triangulate the "hole" by introducing an additional edge, that is

$$\mathbf{u}_{0,1} \leftrightarrow \mathbf{u}_{2,3}, \quad \mathbf{u}_{0,2} \leftrightarrow \mathbf{u}_{1,3}, \quad \text{and} \ \mathbf{u}_{0,3} \leftrightarrow \mathbf{u}_{1,2}.$$

Convexity with respect to the first introduced vertex leads to

$$\frac{\mathbf{b}_{0,2} + \mathbf{b}_{1,3}}{\mathbf{b}_{0,3} + \mathbf{b}_{1,2}} \ge \mathbf{b}_{0,1} + \mathbf{b}_{2,3},$$

which can be transformed in the following way $(\mathbf{i} = (2, 0, \dots, 0))$:

$$\begin{array}{rcl} \mathbf{b}_{0,2} + \mathbf{b}_{1,3} & \geq & \mathbf{b}_{0,1} + \mathbf{b}_{2,3}, \\ \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_1+\hat{\mathbf{e}}_3} - \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_1} & \geq & \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_2+\hat{\mathbf{e}}_3} - \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_2}, \\ & \Delta_{\hat{\mathbf{e}}_3}^1 \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_1} & \geq & \Delta_{\hat{\mathbf{e}}_3}^1 \mathbf{b}_{\mathbf{i}+\hat{\mathbf{e}}_2}, \\ & & \Delta_{1,3} \mathbf{b}_{\mathbf{i}} & \geq & \Delta_{2,3} \mathbf{b}_{\mathbf{i}}. \end{array}$$

Applying this idea to all triangulations, convexity with respect to the first case turns out to be equivalent to

$$\frac{\Delta_{1,2} \mathbf{b}_{\mathbf{i}}}{\Delta_{1,3} \mathbf{b}_{\mathbf{i}}} \ge \Delta_{2,3} \mathbf{b}_{\mathbf{i}}, \tag{16}$$

introducing a vertex of the second type leads to

$$\begin{array}{l} \Delta_{1,2} \mathbf{b}_{\mathbf{i}} \\ \Delta_{2,3} \mathbf{b}_{\mathbf{i}} \end{array} \ge \Delta_{1,3} \mathbf{b}_{\mathbf{i}}, \tag{17}$$

while the last case means

$$\begin{array}{l} \Delta_{1,3} \mathbf{b}_{\mathbf{i}} \\ \Delta_{2,3} \mathbf{b}_{\mathbf{i}} \end{array} \ge \Delta_{1,2} \mathbf{b}_{\mathbf{i}}.$$
 (18)

We now consider the control polyhedron spanned by the vertices

$$\mathbf{b}_{(0,2,0,0)} = \mathbf{b}_{(0,1,1,0)} = \mathbf{b}_{(0,0,2,0)} = \mathbf{b}_{(0,0,1,1)} = \mathbf{b}_{(0,0,0,2)} = 1$$

and all remaining $\mathbf{b_i} = \mathbf{0}$. This results in the matrix

$$\Delta = \left(\Delta_{j,k} \mathbf{b}_{(2,0,0,0)}\right)_{j,k=1}^{m} = \begin{pmatrix} 1 & 1 & 0\\ 1 & 1 & 1\\ 0 & 1 & 1 \end{pmatrix}$$

which shows that the control polyhedron satisfies the inequalities (15) and (17) and is therefore convex with respect to the second triangulation. According to [3]

$$\mathbf{b}^2(\mathbf{u}) = \sum_{|\mathbf{i}|=2} \mathbf{b}_{\mathbf{i}} B_{\mathbf{i}}^2(\mathbf{u})$$

is a convex function.

On the other hand one evaluates $H[\mathbf{b}^2]$, to be equal to $\frac{1}{2}\Delta$; not being positive semi-definite, since $det \Delta = -1$, in contradiction to the nescessary condition for convexity stated above. Hence it follows, that convexity with respect to *one* single triangulation does not lead to consistent results.

Due to these results we call a polyhedron Φ convex, if it is convex with respect to all triangulations; that means, Φ is convex in the sense of (15) and the vertices of the "hole" are complanar in an *m*-dimensional meaning; i.e., these points lie on an *m*-dimensional hyperplane, thus forming one face of Φ , which seems to be a nice geometric interpretation. In other words, Φ is convex if and only if the inequalities

$$\Delta_{k,k} \mathbf{b}_{\mathbf{i}} \ge \Delta_{j,k} \mathbf{b}_{\mathbf{i}} \ge 0 \tag{19}$$

and

$$\Delta_{j,k} \mathbf{b}_{\mathbf{i}} = \Delta_{k,l} \mathbf{b}_{\mathbf{i}} \tag{20}$$

hold for $1 \leq j, k, l \leq m$ and $i_0 \geq 2$, where (20) is the numerical expression for complanarity.

A function $f: S_m \to \mathbb{R}$ is said to be *polyhedrally convex*, if all evaluation polyhedrons $\Phi = \left\{ f(\frac{\mathbf{i}}{n}) : |\mathbf{i}| = n \right\}$ are convex for $n \ge 1$. Using (9) once again, we can give the characterization of polyhedrally convex functions:

Proposition 6. A C^2 – function $f: S_m \to \mathbb{R}$ is polyhedrally convex if and only if the inequalities

$$D_{k,k}f(\mathbf{u}) \ge D_{j,k}f(\mathbf{u}) \ge 0 \tag{21}$$

and

$$D_{j,k}f(\mathbf{u}) = D_{k,l}f(\mathbf{u}) \tag{22}$$

hold for $1 \leq j, k, l \leq m$ and $\mathbf{u} \in S_m$.

We notice that polyhedral convexity (especially for $m \geq 3$) is a rather restrictive property, but nevertheless Bernstein polynomials do preserve it. Indeed, we prove in analogy to Theorem 3

Theorem 7. If $f : S_m \to \mathbb{R}$ is a polyhedrally convex function, then all Bernstein polynomials $\mathbf{b}^n[f]$, $n \in \mathbb{N}$ are polyhedrally convex, too.

Let us finally remark that from the point of view given above, the twodimensional case, as discussed in [1], can now be simply transferred to the m-dimensional one without any further difficulties. Indeed, one can easily show that the conditions stated in (19) and (20) are more than sufficient for the positive semi-definiteness of H_i and thus for the convexity of \mathbf{b}^n .

§5 Conclusions

In the paper we included the notion of convexity between two other notions, one stronger than classical convexity, the other one weaker, so that they are all equivalent in the univariate case. We further showed that they are preserved by the Bernstein poylnomials in the multivariate case, according to the following scheme

 $\begin{array}{cccc} f & \text{polyhedrally convex} \Rightarrow & \text{convex} \Rightarrow & \text{axially convex}, \\ & & & & & & & \\ & & & & & & & \\ \mathbf{b}^n[f] & \text{polyhedrally convex} \Rightarrow & \text{convex} \Rightarrow & \text{axially convex}. \end{array}$

Moreover, we were able to proof one further property of axially convex functions, namely the monotony of subsequent Bernstein polynomials, as well as to point out some hidden traps in dimensions lying out of our normal range of imagination.

Acknowlegdements

The author wishes to thank Prof. Berens and Prof. Schmid for their helpful suggestions and for making me aquainted with their unpublished manuscripts.

References

- G. Chang and P.J. Davis. The convexity of Bernstein polynomials over triangles. J. Approx. Theory, 40, 1984.
- [2] G. Chang and Y.Y. Feng. An improved condition for the convexity of Bernstein – Bézier patches. Comput. Aided Geom. Design, 1, 1984.

- [3] W. Dahmen and C.A. Micchelli. Convexity of multivariate Bernstein polynomials and box spline surfaces. *Studia Sci. Math. Hungar.*, 23, 1988.
- [4] A. Dinghas. Über einige Identitäten vom Bernsteinschen Typus. Det Koneglige Norske Videnskabers Selskab, 24(21), 1951.
- [5] G. Farin. Curves and surfaces in CAGD. Academic Press Inc., 1988.
- [6] G. G. Lorentz. Bernstein Polynomials. University of Toronto Press, 1953.
- [7] H.J. Schmid. Bernsteinpolynome. Unpublished manuscript. 1975.

Mathematical Institute University of Erlangen – Nuremberg Bismarckstr. $1\frac{1}{2}$ 8520 Erlangen, F.R.G.