

Total positivity in CAGD

Carla Manni and Hendrik Speleers

Abstract We provide some insights into the fundamental properties of total positivity and geometric optimality. These properties allow us to explain why B-splines are such popular and efficient tools in CAGD. Actually, we show that they form the best basis in a geometric sense to represent piecewise polynomials. The presented material is an extended version of [15, Section 2.2]

1 Totally positive matrices

In this section we introduce stochastic and totally positive matrices [13, 19], and some of their properties that are of interest in connection with geometric features of bases of vector spaces [10].

Definition 1. A matrix is **stochastic** if all its elements are nonnegative and the elements in each row sum up to one.

Definition 2. A matrix $A := (a_{ij}) \in \mathbb{R}^{m \times n}$ is **totally positive (TP)** if

$$\det \begin{pmatrix} a_{i_1, j_1} & \cdots & a_{i_1, j_k} \\ \vdots & & \vdots \\ a_{i_k, j_1} & \cdots & a_{i_k, j_k} \end{pmatrix} \geq 0, \quad 1 \leq i_1 < \cdots < i_k \leq m, \quad 1 \leq j_1 < \cdots < j_k \leq n, \quad (1)$$

for $k = 1, \dots, \min(m, n)$. Moreover, we say that A is **strictly totally positive (STP)** if (1) is strictly positive. The submatrix in (1) is denoted by $A_{\mathbf{i}, \mathbf{j}}$ where $\mathbf{i} := (i_1, \dots, i_k)$, $\mathbf{j} := (j_1, \dots, j_k)$.

A direct consequence of this definition is that a square TP matrix has a nonnegative determinant and all its minors are nonnegative. The set of submatrices to be

Department of Mathematics, University of Rome Tor Vergata, Italy
e-mail: manni@mat.uniroma2.it, speleers@mat.uniroma2.it

checked for being an STP matrix can be reduced as follows; see [1, Theorem 2.5] and [9].

Theorem 1. *A matrix is STP if and only if any square submatrix consisting of consecutive rows/columns is STP.*

It is clear that any TP matrix can only have nonnegative elements and that the opposite is not true. It turns out that any TP matrix can be expressed as a product of very simple matrices with nonnegative elements. The remainder of this section is devoted to prove this factorization result and to derive an important consequence.

Theorem 2. *The product of (stochastic) TP matrices is a (stochastic) TP matrix.*

Proof. Let $A := (a_{i,j}) \in \mathbb{R}^{m \times n}$ and $B := (b_{i,j}) \in \mathbb{R}^{n \times p}$. Let us start by recalling the Cauchy–Binet formula: for any $1 \leq r \leq \min\{m, n, p\}$,

$$\det((AB)_{i,j}) = \sum_{\mathbf{k} \in K_{r,n}} \det((A)_{i,\mathbf{k}}) \det((B)_{\mathbf{k},j}), \quad \mathbf{i} \in K_{r,m}, \quad \mathbf{j} \in K_{r,p}, \quad (2)$$

where

$$K_{r,s} := \{\mathbf{k} := (k_1, \dots, k_r) \in \mathbb{N}^r : 1 \leq k_1 < k_2 < \dots < k_r \leq s\}.$$

Suppose that A and B are TP matrices. The determinant of any submatrix of the product $C = AB$ is nonnegative, because by (2) it can be expressed as the sum of products of determinants of submatrices of A and B , which are nonnegative values. Hence, C is a TP matrix.

Suppose now that A and B are stochastic. Then, the elements $c_{i,j}$ of their product C are clearly nonnegative, and

$$\sum_{j=1}^p c_{i,j} = \sum_{j=1}^p \sum_{k=1}^n a_{i,k} b_{k,j} = \sum_{k=1}^n a_{i,k} \sum_{j=1}^p b_{k,j} = \sum_{k=1}^n a_{i,k} = 1.$$

This means that C is stochastic. □

We also need to introduce the concept of bidiagonal matrices.

Definition 3. *A matrix $A := (a_{ij})$ is **bidiagonal** if $a_{i,j} \neq 0$ implies $k \leq j - i \leq k + 1$ for some k .*

The structure of some bidiagonal matrices is illustrated in Figure 1.

Lemma 1. *Any bidiagonal matrix with nonnegative elements is TP.*

Proof. It is clear that any submatrix of a bidiagonal matrix with nonnegative elements is a lower- or upper-triangular matrix with nonnegative diagonal elements. Thus, its determinant is nonnegative, and Definition 2 concludes the proof. □

Theorem 2 and Lemma 1 imply that the product of (stochastic) bidiagonal matrices with nonnegative elements is a (stochastic) TP matrix. Actually, also the opposite is true as stated in the next theorem; see [11, Theorems 2.2 and 2.3]. The main properties of (stochastic) TP matrices follow from this fundamental factorization result.

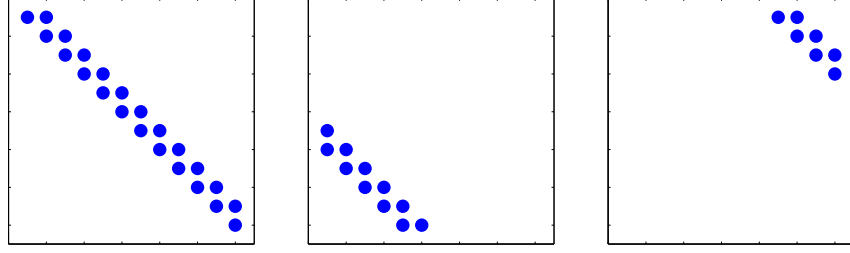


Fig. 1 Bidiagonal matrices. Left: $k = 0$. Center: $k < 0$. Right: $k > 0$.

Theorem 3. *A matrix is (stochastic) TP if and only if it is the product of (stochastic) bidiagonal matrices with nonnegative elements.*

Proof. Thanks to Theorem 2 and Lemma 1, it just remains to prove that any (stochastic) TP matrix is the product of (stochastic) bidiagonal matrices with nonnegative elements.

If A is the zero matrix there is nothing to prove. Let $A \in \mathbb{R}^{m \times n}$ be a nonzero TP matrix. If the k -th row of A vanishes identically, then we consider the bidiagonal matrix $B \in \mathbb{R}^{m \times (m-1)}$, where

$$\begin{aligned} b_{i,i} &= 1, & i &= 1, \dots, k-1, \\ b_{i,i-1} &= 1, & i &= k+1, \dots, m, \\ b_{i,j} &= 0, & & \text{otherwise,} \end{aligned}$$

and we can factorize $A = B\hat{A}$, where \hat{A} is the submatrix obtained from A by removing the k -th row. We repeat the process till we get

$$A = B_1 \cdots B_p \tilde{A},$$

where B_1, \dots, B_p are bidiagonal matrices with nonnegative elements and $\tilde{A} \in \mathbb{R}^{(m-p) \times n}$ is the submatrix obtained from A by removing all the zero rows. So, \tilde{A} is a TP matrix and has no zero rows.

Let $\{\tilde{a}_{i,j}, i-j=l\}$ be the lowest diagonal of \tilde{A} which does not vanish identically, i.e., $\tilde{a}_{i,j} = 0$ whenever $i-j > l$ and there is some i such that $\tilde{a}_{i,i-l} > 0$. Consider the first nonzero element in this diagonal, say $\tilde{a}_{q,q-l}$. This means that $\tilde{a}_{q,q-l} > 0$ and $\tilde{a}_{i,i-l} = 0$ for $i < q$. If $\tilde{a}_{q-1,q-l} = 0$ then for $j > q-l$,

$$\det \begin{pmatrix} \tilde{a}_{q-1,q-l} & \tilde{a}_{q-1,j} \\ \tilde{a}_{q,q-l} & \tilde{a}_{q,j} \end{pmatrix} = -\tilde{a}_{q,q-l} \tilde{a}_{q-1,j},$$

implying that $\tilde{a}_{q-1,j} = 0$. This contradicts the fact that \tilde{a} has no zero rows. Hence, $\tilde{a}_{q-1,q-l} > 0$ and we can apply one step of Gauss elimination to eliminate $\tilde{a}_{q,q-l}$.

More precisely, $\tilde{A} = \tilde{B}S$ where the matrix $\tilde{B} \in \mathbb{R}^{(m-p) \times (m-p)}$ consists of

$$\begin{aligned} \tilde{b}_{i,i} &= 1, \quad i = 1, \dots, m-p, \\ \tilde{b}_{q,q-1} &= \frac{\tilde{a}_{q,q-l}}{\tilde{a}_{q-1,q-l}}, \\ \tilde{b}_{i,j} &= 0, \quad \text{otherwise,} \end{aligned} \quad (3)$$

and the matrix $S \in \mathbb{R}^{(m-p) \times n}$ consists of

$$\begin{aligned} s_{i,j} &= \tilde{a}_{i,j}, \quad i \neq q, \\ s_{q,j} &= 0, \quad j = 1, \dots, q-l. \end{aligned}$$

We see that \tilde{B} is a bidiagonal matrix with nonnegative elements, and from [2, Lemma B] it follows that S is a TP matrix. Set $B_{p+1} := \tilde{B}$. Then, by successively repeating this procedure we can write

$$A = B_1 \cdots B_r U,$$

where B_1, \dots, B_r are bidiagonal matrices with nonnegative elements and U is an upper-triangular TP matrix. Now, we apply the same procedure to U^T , resulting in

$$U^T = \check{B}_1 \cdots \check{B}_s D,$$

where $\check{B}_1, \dots, \check{B}_s$ are bidiagonal matrices with nonnegative elements and D is a diagonal matrix with nonnegative diagonal elements. Summarizing,

$$A = B_1 \cdots B_r D \check{B}_s^T \cdots \check{B}_1^T,$$

which gives the required factorization.

Finally, we address the case where A is a stochastic TP matrix. First, we notice that if two rows in A are linearly dependent then they must be equal because the elements of both of them sum up to 1. So, assuming that two consecutive rows of A are linearly dependent, say k and $k+1$, we write $A = B\hat{A}$ where $B \in \mathbb{R}^{m \times (m-1)}$ is the bidiagonal matrix given by

$$\begin{aligned} b_{i,i} &= 1, \quad i = 1, \dots, k, \\ b_{i,i-1} &= 1, \quad i = k+1, \dots, m, \\ b_{i,j} &= 0, \quad \text{otherwise,} \end{aligned}$$

and \hat{A} is obtained from A by removing the k -th row. Applying the same procedure as before we get

$$A = B_1 \cdots B_p \tilde{A},$$

where $\tilde{A} \in \mathbb{R}^{(m-p) \times n}$ is a stochastic TP matrix with no two consecutive rows linearly dependent. Now, we proceed exactly as in the first part of the proof, and we write $\tilde{A} = \tilde{B}S$, where \tilde{B} is given in (3). Since the rows $q-1$ and q of \tilde{A} are linearly inde-

pendent, no rows of S vanishes identically. Thus, considering the diagonal matrix $D \in \mathbb{R}^{(m-p) \times (m-p)}$ with diagonal elements

$$d_{i,i} = \left(\sum_{j=1}^n s_{i,j} \right)^{-1} > 0$$

we have that DS is stochastic. Then, we arrive at $\tilde{A} = (\tilde{B}D^{-1})(DS)$. Since \tilde{A} is stochastic it follows that $\tilde{B}D^{-1}$ is stochastic. By successively applying this procedure, we can conclude the proof. \square

The factorization in Theorem 3 implies that any TP matrix is variation-diminishing in a certain sense.

Theorem 4 (Variation diminution). *Let $A \in \mathbb{R}^{m \times n}$ be a TP matrix and let $\mathbf{v} \in \mathbb{R}^n$ be any vector. Then,*

$$S^-(A\mathbf{v}) \leq S^-(\mathbf{v}), \quad (4)$$

where $S^-(\mathbf{v})$ denotes the number of sign changes in the components of \mathbf{v} .

Proof. Thanks to Theorem 3 we have $A = B_1 \cdots B_q$ where each B_j , $j = 1, \dots, q$ is a bidiagonal matrix with nonnegative elements. Hence, it suffices to prove (4) for bidiagonal matrices, because this ensures

$$S^-(A\mathbf{v}) = S^-(B_1 B_2 \cdots B_q \mathbf{v}) \leq S^-(B_2 \cdots B_q \mathbf{v}) \leq \cdots \leq S^-(B_q \mathbf{v}) \leq S^-(\mathbf{v}).$$

Suppose $B \in \mathbb{R}^{m \times n}$ is a TP bidiagonal matrix. We choose l such that $b_{i,j} \neq 0$ implies $l \leq j - i \leq l + 1$. For all $i = 1, \dots, m$ we have

$$(B\mathbf{v})_i = b_{i,i+l}v_{i+l} + b_{i,i+l+1}v_{i+l+1},$$

where undefined terms are taken to be zero. So, if v_{i+l} and v_{i+l+1} have the same sign then Bv shares the same sign. Now, we set $s := \min(1, l - 1)$ and $r := \max(m, n - l - 1)$ and we define the vector

$$\mathbf{w} := (v_{s+l}, (Bv)_s, v_{s+l+1}, (Bv)_{s+1}, \dots, v_{i+l}, (Bv)_i, v_{i+l+1}, \dots, v_{r+l}, (Bv)_r, v_{r+l+1}),$$

where again undefined terms are taken to be zero. Thus,

$$S^-(B\mathbf{v}) \leq S^-(\mathbf{w}) = S^-(\mathbf{v}),$$

which concludes the proof. \square

Finally, we focus on square nonsingular TP matrices. The next theorem can be proved by using the same line of arguments as in the proof of Theorem 3; see [12, Theorem 1].

Theorem 5. *If $A \in \mathbb{R}^{n \times n}$ is a nonsingular stochastic TP matrix then it can be factorized as*

$$A = L^{(1)} \cdots L^{(n-1)} U^{(1)} \cdots U^{(n-1)}, \quad (5)$$

where

$$L^{(l)} := \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & 1 & 0 & & 0 \\ 0 & & \lambda_{n-l+1}^{(l)} & 1 - \lambda_{n-l+1}^{(l)} & & 0 \\ & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & \lambda_n^{(l)} & 1 - \lambda_n^{(l)} \end{pmatrix}, \quad \lambda_i^{(l)} \in [0, 1), \quad (6)$$

and

$$U^{(l)} := \begin{pmatrix} 1 - \mu_1^{(l)} & \mu_1^{(l)} & 0 & \cdots & \cdots & 0 \\ 0 & 1 - \mu_2^{(l)} & \mu_2^{(l)} & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & & 1 - \mu_l^{(l)} & \mu_l^{(l)} & 0 \\ 0 & & & 0 & 1 & 0 \\ & & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \quad \mu_i^{(l)} \in [0, 1). \quad (7)$$

2 Totally positive bases

In this section we connect TP matrices with bases of vector spaces.

Definition 4. A basis $\{\varphi_1, \dots, \varphi_n\}$ of a space \mathbb{U}_n is **totally positive (TP)** on an interval $I \subset \mathbb{R}$ if each collocation matrix

$$\begin{pmatrix} \varphi_1(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & & \vdots \\ \varphi_1(x_m) & \cdots & \varphi_n(x_m) \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (8)$$

is TP, where

$$x_1 < x_2 < \cdots < x_m, \quad x_i \in I, \quad i = 1, \dots, m.$$

Moreover, the basis is **normalized** if $\sum_{j=1}^n \varphi_j = 1$.

By Theorem 2 we immediately obtain the following results.

Theorem 6. Let $\{\varphi_1, \dots, \varphi_n\}$ be TP on I .

- (i) If $f : J \rightarrow I$ is increasing then $\{\varphi_1 \circ f, \dots, \varphi_n \circ f\}$ is TP on J .
- (ii) If g is nonnegative on I then $\{g\varphi_1, \dots, g\varphi_n\}$ is TP on I .
- (iii) If $A := (a_{i,j})$ is a TP matrix then $\{\sum_{j=1}^n a_{1,j}\varphi_j, \dots, \sum_{j=1}^n a_{n,j}\varphi_j\}$ is TP on I .

Note that evaluating $\sum_{j=1}^n c_j \varphi_j$ at any sequence of points $x_1 < x_2 < \dots < x_m$, $x_i \in I$, $i = 1, \dots, m$, is nothing else than multiplying the collocation matrix (8) by the coefficient vector $(c_1, \dots, c_n)^T$. The next theorem is a direct consequence.

Theorem 7 (Variation diminution). *If $\{\varphi_1, \dots, \varphi_n\}$ is a TP basis of \mathbb{U}_n and $c_1, \dots, c_n \in \mathbb{R}$, then the number of sign changes of the element $(\sum_{j=1}^n c_j \varphi_j) \in \mathbb{U}_n$ is less than or equal to the number of sign changes of (c_1, \dots, c_n) .*

Let $\mathbf{c}_j \in \mathbb{R}^d$ be given. Then,

$$\mathcal{C}(t) := \sum_{j=1}^n \mathbf{c}_j \varphi_j(t), \quad t \in I$$

defines a curve in \mathbb{R}^d with components in \mathbb{U}_n . We say that $\mathbf{c}_1, \dots, \mathbf{c}_n$ are the **control points** of \mathcal{C} (with respect to the basis $\{\varphi_1, \dots, \varphi_n\}$) and the polygonal line they form is the **control polygon** of \mathcal{C} .

The following result is a more geometric consequence of the variation-diminishing property in Theorem 7.

Theorem 8. *Let $\{\varphi_1, \dots, \varphi_n\}$ be a normalized TP basis. Define the planar curve $\mathcal{C}(t) = \sum_{j=1}^n \mathbf{c}_j \varphi_j(t)$, $t \in I$, $\mathbf{c}_j \in \mathbb{R}^2$. Then the number of times the curve \mathcal{C} crosses any straight line ℓ is bounded by the number of times its control polygon crosses ℓ .*

Proof. Let $ax + by + c = 0$ be the equation of ℓ , and define $\mathbf{c}_j := (c_{j,x}, c_{j,y})$ and $\mathcal{C}(t) := (C_x(t), C_y(t))$. By Theorem 7, if the basis is normalized and TP then the number of changes in sign in

$$aC_x(t) + bC_y(t) + c = \sum_{j=1}^n (ac_{j,x} + bc_{j,y} + c) \varphi_j(t)$$

is bounded above by the number of changes in sign in the sequence

$$ac_{j,x} + bc_{j,y} + c, \quad j = 1, \dots, n.$$

Of course, any change in sign in the sequence above corresponds to a cross of the control polygon and the straight line ℓ . \square

The previous result ensures that TP bases provide shape-preserving representations, as illustrated in Figure 2. For example, if the control polygon is convex then the corresponding curve is convex as well. We refer to [11, Section 8] for further generalizations of the variation diminishing properties of TP bases.

TP matrices and bases have been extensively used in constructing shape preserving interpolating schemes for planar and spatial curves, see for example [5, 6, 7] and references therein.

We now give some examples of TP bases of the polynomial space \mathbb{P}_p of degree p .

Example 1. The monomial basis $\{1, x, \dots, x^p\}$ is TP on $[0, +\infty)$. This can be easily seen as follows. By using the relation

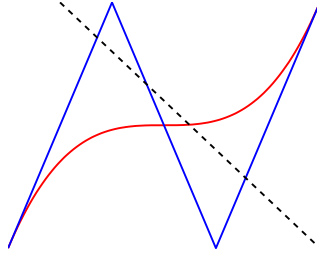


Fig. 2 Variation diminution and shape preservation.

$$\det \begin{pmatrix} 1 & x_1 & \cdots & x_1^k \\ \vdots & \vdots & & \vdots \\ 1 & x_{k+1} & \cdots & x_{k+1}^k \end{pmatrix} = \prod_{1 \leq i < j \leq k+1} (x_j - x_i),$$

we conclude from Theorem 1 that any collocation matrix is STP on $(0, +\infty)$. Then, by continuity, any collocation matrix is TP on $[0, +\infty)$.

Example 2. Consider Theorem 6 with the monomial basis $\varphi_i(x) := x^i$, $i = 0, \dots, p$,

$$f : [0, 1) \rightarrow [0, +\infty), \quad f(x) := \frac{x}{1-x},$$

and

$$g(x) := (1-x)^p.$$

It follows that

$$g(x)[\varphi_i \circ f](x) = (1-x)^p \left(\frac{x}{1-x} \right)^i.$$

Since we know from the previous example that the monomial basis is TP on $[0, +\infty)$, we have that

$$x^i (1-x)^{p-i}, \quad i = 0, \dots, p$$

is a TP basis for \mathbb{P}_p on $[0, 1)$. Since the diagonal matrix with diagonal entries $\binom{p}{i}$ is TP, and by applying again Theorem 6, we obtain that the Bernstein polynomials

$$B_{j,p}(x) := \binom{p}{j} x^j (1-x)^{p-j}, \quad j = 0, \dots, p. \quad (9)$$

are a TP basis on $[0, 1)$. By continuity, the TP property extends to $[0, 1]$.

3 Geometrically optimal bases

In this section we introduce the concept of geometrically optimal bases.

Definition 5. Let $(\vartheta_1, \dots, \vartheta_n)$ be a normalized TP basis of the space \mathbb{U}_n . We say that $(\vartheta_1, \dots, \vartheta_n)$ is an **optimal normalized TP (ONTP)** basis of the space \mathbb{U}_n , in a geometric sense, if any other normalized TP (NTP) basis $(\varphi_1, \dots, \varphi_n)$ of \mathbb{U}_n can be written as

$$(\varphi_1, \dots, \varphi_n) = (\vartheta_1, \dots, \vartheta_n)K, \quad (10)$$

where the matrix K is stochastic and TP.

Definition 5 has a geometric interpretation. The matrix K in (10) is nonsingular and stochastic, so it can be factorized in the form (5). A direct computation shows that $U^{(l)}$ in (7) can be factorized as

$$U^{(l)} = \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & & 1 - \mu_i^{(l)} & \mu_i^{(l)} & 0 \\ 0 & & & 0 & 1 & 0 \\ & & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 - \mu_1^{(l)} & \mu_1^{(l)} & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & & 1 & 0 & 0 \\ 0 & & & 0 & 1 & 0 \\ & & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix},$$

and a similar factorization holds for any $L^{(l)}$ in (6). Hence, we have that any nonsingular stochastic TP matrix can be factorized in terms of matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & & 1 & 0 & 0 \\ 0 & & & \lambda_i & 1 - \lambda_i & 0 \\ & & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & & 1 - \mu_i & \mu_i & 0 \\ 0 & & & 0 & 1 & 0 \\ & & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \quad \lambda_i, \mu_j \in [0, 1). \quad (11)$$

We say that each matrix in (11) describes an **elementary corner cutting**. This is motivated by the fact that, see also Figure 3,

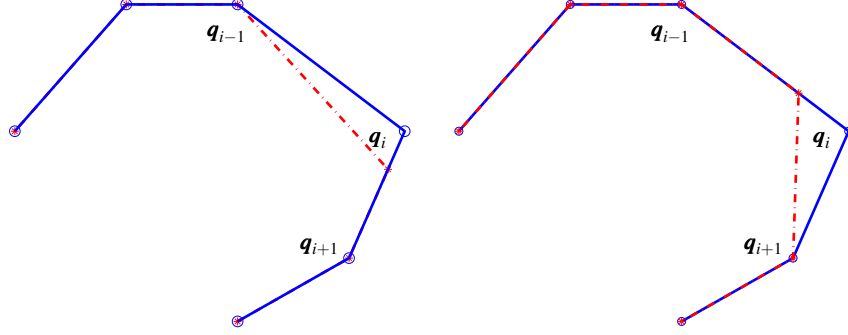


Fig. 3 Elementary corner cuttings.

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & 0 \\ & & \ddots & & & \\ 0 & & 1 - \mu_i & \mu_i & & 0 \\ 0 & & 0 & 1 & & 0 \\ & & & & \ddots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ \vdots \\ \mathbf{q}_n \end{pmatrix} = \begin{pmatrix} \mathbf{q}_1 \\ \mathbf{q}_2 \\ (1 - \mu_i)\mathbf{q}_i + \mu_i\mathbf{q}_{i+1} \\ \vdots \\ \mathbf{q}_n \end{pmatrix}.$$

Let $\{\vartheta_1, \dots, \vartheta_n\}$ be the ONTP basis of \mathbb{U}_n , and let $\{\varphi_1, \dots, \varphi_n\}$ be an NTP basis. Then, we can represent the same curve \mathcal{C} with respect to these two bases,

$$\mathcal{C}(t) := \sum_{j=1}^n \mathbf{q}_j \varphi_j(t) = \sum_{j=1}^n \mathbf{q}_j \sum_{i=1}^n k_{ij} \vartheta_i(t) = \sum_{i=1}^n \left(\sum_{j=1}^n k_{ij} \mathbf{q}_j \right) \vartheta_i(t) = \sum_{i=1}^n \mathbf{p}_i \vartheta_i(t),$$

where

$$\begin{pmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_n \end{pmatrix} = K \begin{pmatrix} \mathbf{q}_1 \\ \vdots \\ \mathbf{q}_n \end{pmatrix}.$$

It follows that the control polygon of \mathcal{C} with respect to the ONTP basis is obtained by successive elementary corner cuttings from the control polygon with respect to any other NTP basis. Hence, the control polygon with respect to the ONTP basis is the closest one to \mathcal{C} ; see also Figure 4. It lies “between” \mathcal{C} and the control polygon with respect to any other NTP basis. As a result, the coefficients of an element of \mathbb{U}_n with respect to the ONTP basis provide the most accurate description from a geometric point view of the element itself.

The next Theorem ensures that whenever a space has an NTP basis then it has an ONTP basis; see [4, Theorem 4.2].

Theorem 9. *If the space \mathbb{U}_p has an NTP basis, then it has a unique ONTP basis.*

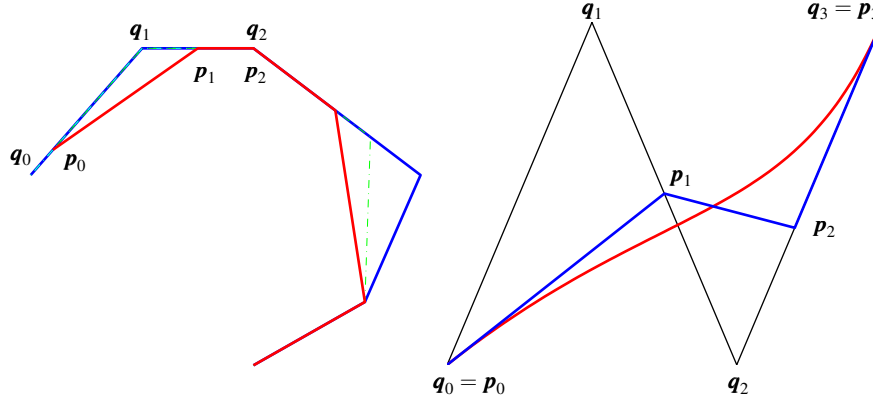


Fig. 4 Left: corner cutting as successive elementary corner cuttings. Right: control polygons with respect to an NTP basis (black) and to the ONTP basis (blue) for the same curve (red).

ONTP bases can be characterized in terms of the behavior of the basis elements at the two ends of their support (with respect to I) as stated in the following theorem; see [17, Proposition 3.2].

Theorem 10. *Suppose that the space \mathbb{U}_n has an NTP basis on a close set I , and let $(\vartheta_1, \dots, \vartheta_n)$ be a basis of \mathbb{U}_n formed by nonnegative functions. Then, the following properties are equivalent.*

- (i) $(\vartheta_1, \dots, \vartheta_n)$ is the ONTP basis of \mathbb{U}_n .
- (ii) Let $I_j := \{x \in I : \vartheta_j(x) \neq 0\}$ and $a_j := \inf I_j$, $b_j := \sup I_j$, for $j = 1, \dots, n$. Then, I_j is an interval and

$$a_k \leq a_j, b_k \leq b_j, \quad \lim_{x \rightarrow a_k^+} \frac{\vartheta_j(x)}{\vartheta_k(x)} = 0, \quad \lim_{x \rightarrow b_j^-} \frac{\vartheta_k(x)}{\vartheta_j(x)} = 0 \text{ for } k < j.$$

Example 3. Let $\mathbb{U}_p = \mathbb{P}_p$ and let $\vartheta_j = B_{j-1,p}$, $j = 1, \dots, p+1$ be the Bernstein polynomials as in (9). The case $p = 3$ is illustrated in Figure 5. These functions form a partition of unity and span the space \mathbb{P}_p on $[0, 1]$. Then, example 2 ensures that they are an NTP basis of \mathbb{P}_p on $[0, 1]$. Consider $I_j = [0, 1]$, $j = 1, \dots, p+1$ in Theorem 10, and since

$$D^r B_{j-1,p}(0) = 0, \quad r = 0, \dots, j-2,$$

and

$$D^r B_{j-1,p}(1) = 0, \quad r = 0, \dots, p-j,$$

we immediately obtain that Bernstein polynomials are the ONTP basis of \mathbb{P}_p on $[0, 1]$. Therefore, Bernstein polynomials are the geometrically optimal way to represent any polynomial of degree p ; see also [3]. It turns out that, besides this geometric

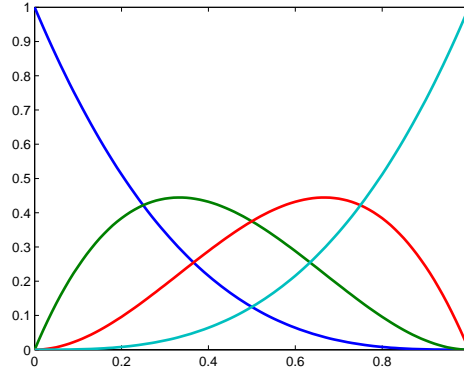


Fig. 5 Cubic Bernstein polynomials ($p = 3$).

optimality property, Bernstein polynomials also possess optimal numerical stability properties. We refer to [8] for further details.

4 Spline examples of TP bases

B-splines and NURBS are some of the most important tools in computer-aided geometric design. In this section we prove that they are NTP bases.

4.1 B-splines

We first show that B-splines form an NTP basis, actually the ONTP basis, of the space they span. In order to define B-splines we need the concept of knot sequences.

Definition 6. A **knot sequence** ξ is a nondecreasing sequence of real numbers,

$$\xi := \{\xi_i\}_{i=1}^m = \{\xi_1 \leq \xi_2 \leq \dots \leq \xi_m\}, \quad m \in \mathbb{N}.$$

The elements ξ_i are called **knots**.

Provided that $m \geq p + 2$ we can define B-splines of degree p over the knot-sequence ξ .

Definition 7. Suppose for a nonnegative integer p and some integer j that $\xi_j \leq \xi_{j+1} \leq \dots \leq \xi_{j+p+1}$ are $p+2$ real numbers taken from a knot sequence ξ . The j -th **B-spline** $B_{j,p,\xi} : \mathbb{R} \rightarrow \mathbb{R}$ of degree p is identically zero if $\xi_{j+p+1} = \xi_j$ and otherwise defined recursively by

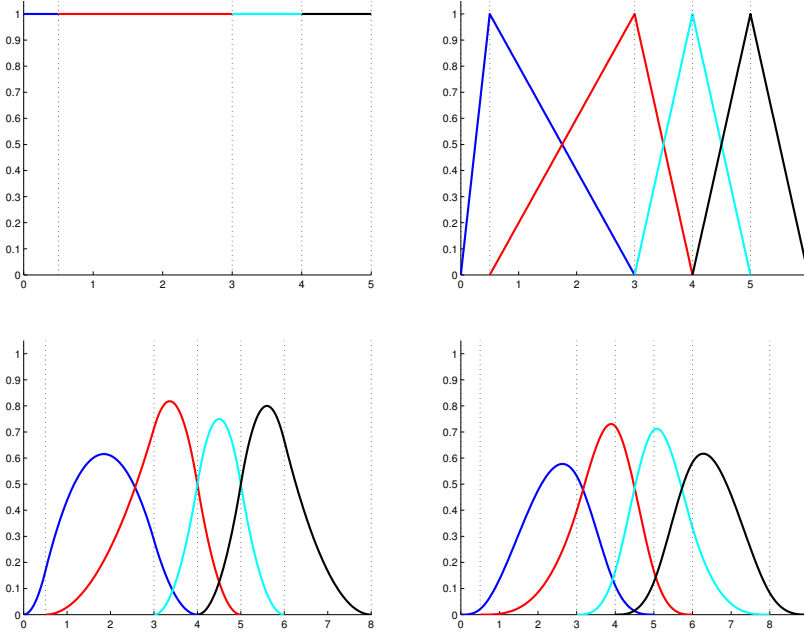


Fig. 6 B-splines of degree 0, 1, 2 and 3.

$$B_{j,p,\xi}(x) := \frac{x - \xi_j}{\xi_{j+p} - \xi_j} B_{j,p-1,\xi}(x) + \frac{\xi_{j+p+1} - x}{\xi_{j+p+1} - \xi_{j+1}} B_{j+1,p-1,\xi}(x),$$

starting with

$$B_{i,0,\xi}(x) := \begin{cases} 1, & \text{if } x \in [\xi_i, \xi_{i+1}), \\ 0, & \text{otherwise.} \end{cases}$$

Here we used the convention that fractions with zero denominator have value zero.

Figure 6 illustrates some B-splines of different degree. A B-spline enjoys several nice properties; see, e.g., [14]. In particular, the following list is of interest for our TP purpose.

- **Local support.** A B-spline is locally supported on the interval given by the extreme knots used in its definition, i.e.,

$$B_{j,p,\xi}(x) = 0, \quad x \notin [\xi_j, \xi_{j+p+1}). \quad (12)$$

- **Nonnegativity.** A B-spline is nonnegative everywhere, and positive inside its support, i.e.,

$$B_{j,p,\xi}(x) \geq 0, \quad x \in \mathbb{R}, \quad \text{and} \quad B_{j,p,\xi}(x) > 0, \quad x \in (\xi_j, \xi_{j+p+1}). \quad (13)$$

- **Smoothness.** If ξ is a knot of $B_{j,p,\xi}$ of multiplicity $\mu \leq p+1$, then

$$B_{j,p,\xi} \in C^{p-\mu}(\xi), \quad (14)$$

i.e., its derivatives of order $0, 1, \dots, p-\mu$ are continuous at ξ . Moreover, the derivative of order $p-\mu+1$ has a nonzero jump at ξ .

- **Knot insertion.** Let the knot sequence $\hat{\xi}$ be obtained from ξ by inserting a single knot $\hat{\xi}$. Then,

$$B_{j,p,\xi} = \hat{\omega}_{j,p} B_{j,p,\hat{\xi}} + (1 - \hat{\omega}_{j+1,p}) B_{j+1,p,\hat{\xi}}, \quad (15)$$

where

$$\hat{\omega}_{j,0} := \begin{cases} 1, & \text{if } \xi_j \leq \hat{\xi}, \\ 0, & \text{if } \hat{\xi} < \xi_j, \end{cases}$$

and

$$\hat{\omega}_{j,p} := \begin{cases} 1, & \text{if } \xi_{j+p} \leq \hat{\xi}, \\ \frac{\hat{\xi} - \xi_j}{\xi_{j+p} - \xi_j}, & \text{if } \xi_j < \hat{\xi} < \xi_{j+p}, \quad p \geq 1. \\ 0, & \text{if } \hat{\xi} \leq \xi_j, \end{cases}$$

Suppose for integers $n > p \geq 0$ that a knot sequence

$$\xi := \{\xi_i\}_{i=1}^{n+p+1} = \{\xi_1 \leq \xi_2 \leq \dots \leq \xi_{n+p+1}\}, \quad n \in \mathbb{N}, \quad p \in \mathbb{N}_0,$$

is given. This knot sequence allows us to define a set of n B-splines of degree p , namely

$$\{B_{1,p,\xi}, \dots, B_{n,p,\xi}\}. \quad (16)$$

We consider the space

$$\mathbb{S}_{p,\xi} := \left\{ s : [\xi_{p+1}, \xi_{n+1}] \rightarrow \mathbb{R} : s = \sum_{j=1}^n c_j B_{j,p,\xi}, c_j \in \mathbb{R} \right\}.$$

This is the space of **splines** spanned by the B-splines in (16) over the interval $[\xi_{p+1}, \xi_{n+1}]$, which is called the **basic interval**.

- A knot sequence ξ is called **$(p+1)$ -regular** if $\xi_j < \xi_{j+p+1}$ for $j = 1, \dots, n$. By the local support such a knot sequence ensures that all the B-splines in (16) are not identically zero.
- A knot sequence ξ is called **$(p+1)$ -basic** if it is $(p+1)$ -regular with $\xi_{p+1} < \xi_{p+2}$ and $\xi_n < \xi_{n+1}$. The B-splines in (16) defined on a $(p+1)$ -basic knot sequence are linearly independent on the basic interval $[\xi_{p+1}, \xi_{n+1}]$.
- A knot sequence ξ is called **$(p+1)$ -open** on an interval $[a, b]$ if it is $(p+1)$ -regular and it has end knots of multiplicity $p+1$, i.e.,

$$a := \xi_1 = \dots = \xi_{p+1} < \xi_{p+2} \leq \dots \leq \xi_n < \xi_{n+1} = \dots = \xi_{n+p+1} =: b. \quad (17)$$

This sequence is $(p+1)$ -basic and is often used in practice. In particular, it turns out to be natural to construct open curves, clamped at two given points

The set (16) of B-splines has several nice features; see, e.g., [14]. In particular, we mention the following properties.

- **Linear independence.** If ξ is $(p+1)$ -basic, then the B-splines in (16) are linearly independent on the basic interval. Thus, the spline space $\mathbb{S}_{p,\xi}$ is a vector space of dimension n .
- **Partition of unity.** We have

$$\sum_{j=1}^n B_{j,p,\xi}(x) = 1, \quad x \in [\xi_{p+1}, \xi_{n+1}]. \quad (18)$$

Since the B-splines are nonnegative it follows that they form a **nonnegative partition of unity** on $[\xi_{p+1}, \xi_{n+1}]$.

Let $\mathbf{c}_j \in \mathbb{R}^d$ be given. The parametric curve

$$\mathcal{C}(t) := \sum_{j=1}^n \mathbf{c}_j B_{j,p,\xi}(t), \quad t \in [\xi_{p+1}, \xi_{n+1}], \quad (19)$$

is called a **B-spline curve** in \mathbb{R}^d , and the points \mathbf{c}_j are its **control points**.

In the following we assume that the knot sequence ξ is $(p+1)$ -basic.

Theorem 11. *The B-spline basis $\{B_{1,p,\xi}, \dots, B_{n,p,\xi}\}$ is NTP on the basic interval $[\xi_{p+1}, \xi_{n+1}]$.*

Proof. By (18) the B-spline basis is a normalized basis on the basic interval. Hence, it suffices to prove that any collocation matrix

$$A := \begin{pmatrix} B_{1,p,\xi}(x_1) & \cdots & B_{n,p,\xi}(x_1) \\ \vdots & & \vdots \\ B_{1,p,\xi}(x_m) & \cdots & B_{n,p,\xi}(x_m) \end{pmatrix}, \quad (20)$$

where

$$x_1 < x_2 < \cdots < x_m \in [\xi_{p+1}, \xi_{n+1}], \quad (21)$$

is TP. We will prove this in three steps. Let $[\dots, B_{j,p,\xi}, \dots]$ be any square submatrix of A containing (a part of) the j -th column of A .

In the first step, we consider the knot sequence $\hat{\xi}$ obtained by inserting a single knot $\hat{\xi}$ in the sequence ξ . We denote by \hat{A} the collocation matrix of the basis $\{B_{1,p,\hat{\xi}}, \dots, B_{n+1,p,\hat{\xi}}\}$ at the points in (21). By using (15) and the linearity of the determinant we have

$$\begin{aligned} & \det([\dots, B_{j,p,\xi}, \dots]) \\ &= \hat{\omega}_{j,p} \det([\dots, B_{j,p,\hat{\xi}}, \dots]) + (1 - \hat{\omega}_{j+1,p}) \det([\dots, B_{j+1,p,\hat{\xi}}, \dots]). \end{aligned}$$

Therefore, for any increasing sequence of row indices $\mathbf{i} := (i_1, \dots, i_k)$ and any increasing sequence of column indices $\mathbf{j} := (j_1, \dots, j_k)$, we have

$$\det((A)_{\mathbf{i}, \mathbf{j}}) = \sum_{\mathbf{r}} \hat{\gamma}_{\mathbf{r}} \det((\hat{A})_{\mathbf{i}, \mathbf{r}}),$$

where $\hat{\gamma}_{\mathbf{r}} \geq 0$ and any $\mathbf{r} := (r_1, \dots, r_k)$ is an increasing sequence of column indices, because only two consecutive columns of \hat{A} contribute to express a given column of A . Moreover, we can confine ourselves to increasing sequences of column indices only because otherwise the corresponding determinants would vanish.

In the second step, we successively insert each x_i in the knot sequence till we obtain a new sequence $\check{\xi} := \{\check{\xi}_i\}_{i=1}^{\check{n}+p+1}$ such that x_i is a knot of multiplicity $p+1$ in $\check{\xi}$. We denote by \check{A} the collocation matrix of the B-spline basis corresponding to $\check{\xi}$ at the points in (21). By applying repeatedly the arguments of the previous step, we obtain

$$\det((A)_{\mathbf{i}, \mathbf{j}}) = \sum_{\mathbf{r}} \check{\gamma}_{\mathbf{r}} \det((\check{A})_{\mathbf{i}, \mathbf{r}}), \quad (22)$$

where $\check{\gamma}_{\mathbf{r}} \geq 0$ and any $\mathbf{r} := (r_1, \dots, r_k)$ is an increasing sequence of column indices. Moreover, we have

$$\check{\xi}_{m-1} < x_i = \check{\xi}_m = \dots = \check{\xi}_{m+p},$$

for some m . Therefore, from the local support and the partition of unity we deduce

$$B_{j,p,\check{\xi}}(x_i) = \begin{cases} \delta_{j,m}, & \text{if } m \leq \check{n}, \\ \delta_{j,m-1}, & \text{if } m = \check{n} + 1, \end{cases}$$

so that in the i -th row of \check{A} only one element is different from zero.

Finally, as last step, we consider any submatrix $(\check{A})_{\mathbf{i}, \mathbf{r}}$ in (22). Let us consider a row of such submatrix, say the one with index l . Since the sequence of indices \mathbf{r} is increasing, if this row is different from zero then it contains exactly one nonzero element, say in the column r , and this element is equal to 1. Moreover, if $\check{a}_{l,\check{r}} \neq 0$ with $\check{l} > l$ then $\check{r} > r$, because the sequence of points in (21) is strictly increasing. Summarizing, $\check{A}_{\mathbf{i}, \mathbf{r}}$ has nonzero rows if and only if it is the identity matrix. In other words, either $\det((\check{A})_{\mathbf{i}, \mathbf{r}}) = 0$ or $\det((\check{A})_{\mathbf{i}, \mathbf{r}}) = 1$. In view of (22), this implies $\det((A)_{\mathbf{i}, \mathbf{j}}) \geq 0$, and it concludes the proof. \square

The next result, see [4], shows that the B-spline basis is the ONTP basis for any space of piecewise polynomials.

Theorem 12. *The B-spline basis $\{B_{1,p,\xi}, \dots, B_{n,p,\xi}\}$ is the ONTP basis of the space $\mathbb{S}_{p,\xi}$ on the basic interval $[\xi_{p+1}, \xi_{n+1}]$.*

Proof. Let us consider Theorem 10 with $\mathbb{U}_n = \mathbb{S}_{p,\xi}$ and $\vartheta_j = B_{j,p,\xi}$ for $j = 1, \dots, n$. By Theorem 11 the set of nonnegative functions $\{B_{1,p,\xi}, \dots, B_{n,p,\xi}\}$ is an NTP basis of $\mathbb{S}_{p,\xi}$. Therefore, Theorem 9 implies that the space $\mathbb{S}_{p,\xi}$ possesses a unique ONTP basis. From (12) and (13) we deduce that $I_j := \{x : B_{j,p,\xi}(x) \neq 0\}$ is an interval such that $\inf I_j = \xi_j$ and $\sup I_j = \xi_{j+p+1}$. Moreover, for $k < j$ we have $\xi_k \leq \xi_j$. Note that

if $\xi_k < \xi_j$ then $B_{j,p,\xi}$ vanishes in a neighborhood of ξ_k . On the other hand, if $\xi_k = \xi_j$ then both $B_{k,p,\xi}$ and $B_{j,p,\xi}$ vanish at ξ_k but $B_{j,p,\xi}$ has a zero of higher multiplicity than $B_{k,p,\xi}$ at ξ_k because of (14). Therefore,

$$\lim_{x \rightarrow \xi_k^+} \frac{B_{j,p,\xi}(x)}{B_{k,p,\xi}(x)} = 0.$$

Similarly, for $k < j$ we have that $\xi_{k+p+1} \leq \xi_{j+p+1}$ and

$$\lim_{x \rightarrow \xi_{j+p+1}^-} \frac{B_{k,p,\xi}(x)}{B_{j,p,\xi}(x)} = 0.$$

From Theorem 10 we conclude that $\{B_{1,p,\xi}, \dots, B_{n,p,\xi}\}$ is the ONTP basis of the space $\mathbb{S}_{p,\xi}$. \square

The above results ensure that the B-spline basis is the best basis from the geometrical point of view for a given space of piecewise polynomials. It is worth mentioning that the total positivity properties of B-splines can be alternatively derived by means of the so-called blossoming principle [16, 20].

4.2 NURBS

We now show that NURBS, the rational extension of B-splines, also form an NTP basis of the space they span.

Definition 8. Given a set of B-splines $\{B_{j,p,\xi}(x), j = 1, \dots, n\}$ and a corresponding set of positive weights,¹

$$\mathbf{w} := \{w_j > 0, j = 1, \dots, n\},$$

the j -th NURBS (non-uniform rational B-spline), $j = 1, \dots, n$ is defined by

$$R_{j,p,\xi,\mathbf{w}}(x) := \frac{w_j B_{j,p,\xi}(x)}{\sum_{i=1}^n w_i B_{i,p,\xi}(x)}.$$

NURBS immediately inherit from B-splines the following properties.

- **Local support.**
- **Nonnegativity.**
- **Smoothness.**
- **Linear independence.**

¹ Negative and zero weights can be used as well. Here, we just consider positive weights because they ensure nice properties of the resulting functions.

- **Partition of unity.**

Similar to (19), a **NURBS curve** is given by

$$\sum_{j=1}^n \mathbf{c}_j R_{j,p,\xi,\mathbf{w}}(t), \quad \mathbf{c}_j \in \mathbb{R}^d.$$

It can be seen as a projective transformation of a B-spline curve in \mathbb{R}^{d+1} ,

$$\left(\sum_{j=1}^n \mathbf{c}_j w_j B_{j,p,\xi}(t), \sum_{j=1}^n w_j B_{j,p,\xi}(t) \right).$$

For further details, we refer to [18].

For NURBS we have the following result.

Theorem 13. *The NURBS basis is NTP.*

Proof. Suppose the values $x_1 < x_2 < \dots < x_m$ are given, and set

$$\mathcal{W}_{p,\xi,\mathbf{w}}(x) := \left(\sum_{j=1}^n w_j B_{j,p,\xi}(x) \right)^{-1}.$$

The NURBS collocation matrix

$$\begin{pmatrix} R_{1,p,\xi,\mathbf{w}}(x_1) & \cdots & R_{n,p,\xi,\mathbf{w}}(x_1) \\ \vdots & & \vdots \\ R_{1,p,\xi,\mathbf{w}}(x_m) & \cdots & R_{n,p,\xi,\mathbf{w}}(x_m) \end{pmatrix} \quad (23)$$

is equal to the product $D_{\mathcal{W}} B D_{\mathbf{w}}$ of the matrices

$$B := \begin{pmatrix} B_{1,p,\xi}(x_1) & \cdots & B_{n,p,\xi}(x_1) \\ \vdots & & \vdots \\ B_{1,p,\xi}(x_m) & \cdots & B_{n,p,\xi}(x_m) \end{pmatrix},$$

and

$$D_{\mathcal{W}} := \begin{pmatrix} \mathcal{W}_{p,\xi,\mathbf{w}}(x_1) & 0 & \cdots & 0 \\ 0 & \mathcal{W}_{p,\xi,\mathbf{w}}(x_2) & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & \mathcal{W}_{p,\xi,\mathbf{w}}(x_m) \end{pmatrix}, \quad D_{\mathbf{w}} := \begin{pmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & w_n \end{pmatrix}.$$

From Theorem 11 we know that the B-spline basis is TP. Thus, the matrix (23) is the product of TP matrices, and so it is TP as well (see Theorem 2). Normalization comes from the partition of unity. \square

References

1. Ando, T.: Totally positive matrices. *Linear Algebra and its Applications* **90**, 165–219 (1987)
2. de Boor, C., Pinkus, A.: The approximation of a totally positive band matrix by a strictly banded totally positive one. *Linear Algebra and its Applications* **42**, 81–98 (1982)
3. Carnicer, J.M., Peña, J.M.: Shape preserving representations and optimality of the Bernstein basis. *Advances in Computational Mathematics* **1**, 173–196 (1993)
4. Carnicer, J.M., Peña, J.M.: Totally positive bases for shape preserving curve design and optimality of B-splines. *Computer Aided Geometric Design* **6**, 633–654 (1994)
5. Costantini, P., Manni, C.: Shape-preserving C^3 interpolation: the curve case. *Advances in Computational Mathematics* **18**, 41–63 (2003)
6. Costantini, P., Manni, C.: A geometric approach for Hermite subdivision. *Numerische Mathematik* **115**, 333–369 (2010)
7. Cravero, I., Manni, C.: Shape-preserving interpolants with high smoothness. *Journal of Computational and Applied Mathematics* **153**, 383–405 (2003)
8. Farouki, R.T., Goodman, T.N.T.: On the optimal stability of the Bernstein basis. *Mathematics of Computation* **65**, 1553–1566 (1996)
9. Fekete, M., Pólya, G.: Über ein Problem von Laguerre. *Rendiconti del Circolo Matematico di Palermo* **34**, 89–120 (1912)
10. Gasca, M., Micchelli, C.A.: *Total Positivity and its Applications*. Kluwer Academic Publishers, Dordrecht (1996)
11. Goodman, T.N.T.: Total positivity and shape of curves. In: M. Gasca, C.A. Micchelli (eds.) *Total Positivity and its Applications*, pp. 157–186. Kluwer Academic Publishers, Dordrecht (1996)
12. Goodman, T.N.T., Micchelli, C.A.: Corner cutting algorithms for the Bézier representation of free form curves. *Linear Algebra and its Applications* **99**, 225–252 (1988)
13. Karlin, S.: *Total Positivity*. Stanford University Press, Stanford (1968)
14. Lyche, T., Manni, C., Speleers, H.: Foundations of spline theory: B-splines, spline approximation, and hierarchical refinement. In: T. Lyche, C. Manni, H. Speleers (eds.) *Splines and PDEs: From Approximation Theory to Numerical Linear Algebra*, *Lecture Notes in Mathematics*, vol. 2219, pp. 1–76. Springer International Publishing, Cham (2018)
15. Manni, C., Speleers, H.: Standard and non-standard CAGD tools for isogeometric analysis: A tutorial. In: A. Buffa, G. Sangalli (eds.) *IsoGeometric Analysis: A New Paradigm in the Numerical Approximation of PDEs*, *Lecture Notes in Mathematics*, vol. 2161, pp. 1–69. Springer International Publishing, Cham (2016)
16. Mazure, M.L.: Chebyshev splines beyond total positivity. *Advances in Computational Mathematics* **14**, 129–156 (2001)
17. Peña, J.M.: *Shape Preserving Representations in Computer-Aided Geometric Design*. Nova Science Publishers, New York (1999)
18. Piegl, L., Tiller, W.: *The NURBS Book*, second edn. Springer-Verlag, Berlin (1997)
19. Pinkus, A.: *Totally Positive Matrices*. Cambridge University Press, Cambridge (2010)
20. Ramshaw, L.: Blossoms are polar forms. *Computer Aided Geometric Design* **6**, 323–358 (1989)