MATHEMATICS OF COMPUTATION
VOLUME 55, NUMBER 192
OCTOBER 1990, PAGES 637-647

ON THE NUMERICAL CONDITION OF BERNSTEIN-BÉZIER SUBDIVISION PROCESSES

R. T. FAROUKI AND C. A. NEFF

ABSTRACT. The linear map $\mathbf{M}$ that takes the Bernstein coefficients of a polynomial $P(t)$ on a given interval $[a, b]$ into those on any subinterval $[\overline{a}, \overline{b}]$ is specified by a stochastic matrix which depends only on the degree n of $P(t)$ and the size and location of $[\overline{a}, \overline{b}]$ relative to $[a, b]$. We show that in the $\|\cdot\|_\infty$-norm, the condition number of $\mathbf{M}$ has the simple form $\kappa_\infty(\mathbf{M}) = [2f\max(u_{\overline{m}}, v_{\overline{m}})]^n$, where $u_{\overline{m}} = (\overline{m}-a)/(b-a)$ and $v_{\overline{m}} = (b-\overline{m})/(b-a)$ are the barycentric coordinates of the subinterval midpoint $\overline{m} = \frac{1}{2}(\overline{a}+\overline{b})$, and f denotes the "zoom" factor $(b-a)/(\overline{b}-\overline{a})$ of the subdivision map. This suggests a practical rule-of-thumb in assessing how far Bézier curves and surfaces may be subdivided without exceeding prescribed (worst-case) bounds on the typical errors in their control points. The exponential growth of $\kappa_\infty(\mathbf{M})$ with n also argues forcefully against the use of high-degree forms in computer-aided geometric design applications.

1. INTRODUCTION

Subdivision algorithms [1, 11] play a key role in a variety of geometric procedures dealing with parametric curves and surfaces: rendering, approximating, computing intersections, etc. Conceptually, the function of a subdivision algorithm is to compute explicit representations of successively smaller segments of a curve or surface until these subsegments are amenable to certain simplifying assumptions which will facilitate the geometric procedure at hand (the familiar "divide and conquer" strategy).

The *de Casteljau algorithm* [3] is perhaps the simplest and most popular of all subdivision algorithms (cf. [4]). Consider the degree-n Bézier curve with control points $\{\mathbf{p}_k\}$:

$$\mathbf{r}(t) = \sum_{k=0}^{n} \mathbf{p}_k \binom{n}{k} \frac{(b-t)^{n-k}(t-a)^k}{(b-a)^n} \quad \text{for } t \in [a, b]. \tag{1}$$

To subdivide the curve (1) at some prescribed parameter value $s \in [a, b]$, we set $\mathbf{p}_k^{(0)} = \mathbf{p}_k$ for $k = 0, 1, \ldots, n$, and then iterate the sequence of linear

Received June 26, 1989; revised November 1, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 15A12, 68U05.

interpolations [4]:

$$\mathbf{p}_k^{(r)} = \frac{(b-s)\mathbf{p}_{k-1}^{(r-1)} + (s-a)\mathbf{p}_k^{(r-1)}}{(b-a)}, \qquad k = r, r+1, \ldots, n, \tag{2}$$

for $r = 1, 2, \ldots, n$. The quantities $\{\mathbf{p}_k^{(r)}\}$ generated by the algorithm (2) form a triangular array, and the elements

$$\mathbf{p}_0^{(0)}, \mathbf{p}_1^{(1)}, \ldots, \mathbf{p}_n^{(n)} \quad \text{and} \quad \mathbf{p}_n^{(n)}, \mathbf{p}_n^{(n-1)}, \ldots, \mathbf{p}_n^{(0)}, \tag{3}$$

on the left- and right-hand sides of this array are the control points of $\mathbf{r}(t)$ on the parameter intervals $[a, s]$ and $[s, b]$ to the left and right of the split point s:

$$\mathbf{r}(t) = \sum_{k=0}^{n} \mathbf{p}_k^{(k)} \binom{n}{k} \frac{(s-t)^{n-k}(t-a)^k}{(s-a)^n} \quad \text{for } t \in [a, s], \tag{4a}$$

$$\mathbf{r}(t) = \sum_{k=0}^{n} \mathbf{p}_n^{(n-k)} \binom{n}{k} \frac{(b-t)^{n-k}(t-s)^k}{(b-s)^n} \quad \text{for } t \in [s, b]. \tag{4b}$$

To subdivide further, we may choose parameter values on the intervals $[a, s]$ and $[s, b]$, and apply the de Casteljau algorithm again to the above forms.

By such compounded application, the original parameter domain $[a, b]$ of $\mathbf{r}(t)$ may be dissected into a sequence of contiguous subintervals $[a_1, b_1], \ldots, [a_N, b_N]$ such that certain subdivision termination criteria are satisfied for each curve segment (these may be phrased in terms of the *variation diminishing* or *convex hull* properties of the Bézier form, or the convergence of the control polygon to the curve; cf. [4] for a review).

In certain circumstances the required subdivision factors $f_i = (b-a)/(b_i - a_i)$ can be very large, and if the de Casteljau algorithm (2) is invoked repeatedly in finite-precision arithmetic, the computed control points on the subintervals $[a_i, b_i]$ may suffer significant error accumulations. These may then induce erroneous decisions regarding the satisfaction of the subdivision termination criteria (such problems are likely to occur, for example, when subdivision methods are used to compute near-tangential intersections of plane curves).

The stability and error-propagation analysis of subdivision techniques is thus of great practical importance. For the de Casteljau algorithm, it is not difficult [5] to write down the *running error analysis* formulae appropriate to (2) and, by simplification, these even furnish (rather weak) a priori error bounds. However, we shall not be concerned in this paper with specific subdivision algorithms such as (2), but rather with the fundamental problem of the *intrinsic stability of the subdivision process*, i.e., the sensitivity of the representation of a curve or surface subsegment to perturbations in the representation of the parent curve or surface from which it is derived.

This intrinsic sensitivity is intimately (though rather subtly) related to the performance of actual subdivision algorithms by the method of *backward error analysis* [16]. Certainly, the accuracy of *any* subdivision algorithm will always be severely degraded under circumstances in which the intrinsic sensitivity is very high.

2. Bernstein-Bézier subdivision processes

For Bernstein-Bézier forms, we derive below a simple characterization of this intrinsic stability in terms of a *condition number* for the subdivision process. The dependence of this condition number on the size and location of the subinterval relative to the parent interval and on the polynomial degree is quite transparent, allowing a simple practical interpretation.

For conciseness we consider henceforth only the subdivision of scalar-valued, univariate polynomials in Bernstein-Bézier form; the extension of the ideas presented below to vector-valued univariate polynomials (parametric curves) or vector-valued tensor products of univariate polynomials (tensor-product parametric surfaces) is quite straightforward. Thus, we will speak simply of "Bernstein coefficients" rather than Bézier control points.

Consider the degree-n Bernstein basis functions:

$$b_k^n(t) = \binom{n}{k} \frac{(b-t)^{n-k}(t-a)^k}{(b-a)^n} \quad \text{for } k = 0, 1, \ldots, n, \tag{5a}$$

$$\overline{b}_j^n(t) = \binom{n}{j} \frac{(\overline{b}-t)^{n-j}(t-\overline{a})^j}{(\overline{b}-\overline{a})^n} \quad \text{for } j = 0, 1, \ldots, n, \tag{5b}$$

defined on two distinct proper intervals $[a, b]$ and $[\overline{a}, \overline{b}]$. The Bernstein coefficients of a polynomial $P(t)$ of degree n in these bases will be denoted by $\{c_k\}$ and $\{\overline{c}_j\}$, so that

$$P(t) = \sum_{k=0}^{n} c_k b_k^n(t) = \sum_{j=0}^{n} \overline{c}_j \overline{b}_j^n(t). \tag{6}$$

Although the expressions (6) for $P(t)$ are valid for all t, we shall be concerned primarily with the interval $t \in [a, b]$ for the first form, and $t \in [\overline{a}, \overline{b}]$ for the second.

Remark 1. We recall two well-known properties of the Bernstein basis functions (5) which will be of subsequent use: (i) they form a *partition of unity*:

$$1 = \sum_{k=0}^{n} b_k^n(t) = \sum_{j=0}^{n} \overline{b}_j^n(t) \quad \text{for all } t, \tag{7}$$

and (ii) they are *nonnegative* over their respective intervals:

$$b_k^n(t) \geq 0 \quad \text{for } t \in [a, b], \ k = 0, 1, \ldots, n, \tag{8a}$$

$$\overline{b}_j^n(t) \geq 0 \quad \text{for } t \in [\overline{a}, \overline{b}], \ j = 0, 1, \ldots, n. \tag{8b}$$

Together, these characteristics imply the well-known *convex hull* property of Bézier forms.

At present we are primarily interested in the transformation of the Bernstein coefficients of $P(t)$ on $[a, b]$ into those on $[\overline{a}, \overline{b}]$. In order to identify the square matrix $\mathbf{M}$ which defines this transformation, we rewrite the terms $b-t$ and $t-a$ in $b_k^n(t)$ as

$$b - t = \frac{(b-\overline{a})(\overline{b}-t) + (b-\overline{b})(t-\overline{a})}{(\overline{b}-\overline{a})}, \tag{9a}$$

$$t - a = \frac{(\overline{a}-a)(\overline{b}-t) + (\overline{b}-a)(t-\overline{a})}{(\overline{b}-\overline{a})}. \tag{9b}$$

Substituting these expressions into (5a) and performing binomial expansions then gives

$$\binom{n}{k}\sum_{r=0}^{n-k}\sum_{s=0}^{k}\frac{\binom{n-k}{r}\binom{k}{s}}{\binom{n}{r+s}}\,\frac{(b-\overline{b})^r(\overline{b}-a)^s(b-\overline{a})^{n-k-r}(\overline{a}-a)^{k-s}}{(b-a)^n}\overline{b}_{r+s}^n(t) \tag{10}$$

for $b_k^n(t)$. Now it may be verified by direct expansion that

$$\frac{\binom{n-k}{r}\binom{k}{s}}{\binom{n}{r+s}} = \frac{\binom{r+s}{s}\binom{n-r-s}{k-s}}{\binom{n}{k}}, \tag{11}$$

so (10) can be rewritten in the form

$$\begin{aligned}\sum_{r=0}^{n-k}\sum_{s=0}^{k}\binom{r+s}{s}\frac{(b-\overline{b})^r(\overline{b}-a)^s}{(b-a)^{r+s}}\binom{n-r-s}{k-s}\frac{(b-\overline{a})^{n-k-r}(\overline{a}-a)^{k-s}}{(b-a)^{n-r-s}}\overline{b}_{r+s}^n(t)\\ = \sum_{r=0}^{n-k}\sum_{s=0}^{k} b_s^{r+s}(\overline{b})\,b_{k-s}^{n-r-s}(\overline{a})\,\overline{b}_{r+s}^n(t).\end{aligned} \tag{12}$$

Thus, on setting $j = r+s$ and taking appropriate summation limits, we observe that

$$b_k^n(t) = \sum_{j=0}^{n}\overline{b}_j^n(t)M_{jk}, \tag{13}$$

where

$$M_{jk} = \sum_{i=\max(0,\,j+k-n)}^{\min(j,\,k)} b_i^j(\overline{b})\,b_{k-i}^{n-j}(\overline{a}) \quad \text{for } j, k = 0, 1, \ldots, n. \tag{14}$$

By substituting (13) into (6) we see that the Bernstein coefficients $\{\overline{c}_j\}$ on the interval $[\overline{a}, \overline{b}]$ are related to those $\{c_k\}$ on $[a, b]$ by

$$\overline{c}_j = \sum_{k=0}^{n} M_{jk}c_k \quad \text{for } j = 0, 1, \ldots, n. \tag{15}$$

We now highlight two important properties of the family of transformation matrices $\mathbf{M}$ whose elements are defined by (14).

Lemma 2. *The matrix elements* (14) *sum to unity across each row of* $\mathbf{M}$:

$$\sum_{k=0}^{n} M_{jk} = 1 \quad \text{for } j = 0, 1, \ldots, n. \tag{16}$$

Proof. The normalization (16) arises from the partition-of-unity property (7) of the basis functions on $[a, b]$ and $[\overline{a}, \overline{b}]$, since on summing over k in (13) we observe that

$$1 \equiv \sum_{k=0}^{n} \sum_{j=0}^{n} \overline{b}_j^n(t) M_{jk} = \sum_{j=0}^{n} \left[\sum_{k=0}^{n} M_{jk} \right] \overline{b}_j^n(t). \tag{17}$$

From (7) and (17) we may then deduce condition (16) by virtue of the linear independence of the basis functions $\{\overline{b}_j^n(t)\}$. □

Lemma 3. *If* $[\overline{a}, \overline{b}] \subset [a, b]$, *the matrix elements* (14) *are all nonnegative*: $M_{jk} \geq 0$ (*and hence* $|M_{jk}| = M_{jk}$) *for all* $0 \leq j, k \leq n$.

Proof. Each of the $n+1$ basis functions $\{b_k^n(t)\}$ is nonnegative over the entire interval $t \in [a, b]$. Since $[\overline{a}, \overline{b}] \subset [a, b]$ implies that $\overline{a} \in [a, b]$ and $\overline{b} \in [a, b]$, each term of the sum (14) for M_{jk} is clearly nonnegative, and thus the entire sum is nonnegative also, for all pairs of indices $0 \leq j, k \leq n$. □

Remark 4. When $[\overline{a}, \overline{b}] \subset [a, b]$, the properties described in Lemmas 2 and 3 indicate that $\mathbf{M}$ is a *stochastic matrix.* In this case, the coefficients $\{\overline{c}_j\}$ on the subinterval $[\overline{a}, \overline{b}]$ are *convex combinations* of those $\{c_k\}$ on the parent interval $[a, b]$.

Stochastic matrices or "Markov chains" arise in discrete probability theory [7]. For a detailed discussion of their pervasive role in computer-aided geometric design problems, the reader may refer to [8, 9].

Now let $\mathbf{M}^{-1}$ denote the inverse of the matrix $\mathbf{M}$ whose elements are defined by (14). We can write down the elements $\{M_{jk}^{-1}\}$ of $\mathbf{M}^{-1}$ directly by simply interchanging the roles of the basis functions $\{b_k^n(t)\}$ and $\{\overline{b}_j^n(t)\}$ on $[a, b]$ and $[\overline{a}, \overline{b}]$ in equations (9)–(14):

$$M_{jk}^{-1} = \sum_{i=\max(0, j+k-n)}^{\min(j,k)} \overline{b}_i^j(b) \overline{b}_{k-i}^{n-j}(a) \quad \text{for } j, k = 0, 1, \ldots, n. \tag{18}$$

The Bernstein coefficients of $P(t)$ on $[a, b]$ are thus given in terms of those on $[\overline{a}, \overline{b}]$ by:

$$c_j = \sum_{k=0}^{n} M_{jk}^{-1} \overline{c}_k \quad \text{for } j = 0, 1, \ldots, n. \tag{19}$$

The elements of $\mathbf{M}^{-1}$ also satisfy the unit row-sum property of Lemma 2, for the same reason, but when $[\overline{a}, \overline{b}] \subset [a, b]$ they clearly cannot share the nonnegativity property of Lemma 3, since $\mathbf{M}^{-1}\mathbf{M} = \mathbf{I}$ and the elements of $\mathbf{M}$ are all nonnegative. In this case, $\mathbf{M}^{-1}$ is *not* a stochastic matrix, and in fact we have:

Lemma 5. *If* $[\bar{a}, \bar{b}] \subset [a, b]$, *the matrix elements* (18) *are of alternating sign*:

$$|M_{jk}^{-1}| = (-1)^{j+k} M_{jk}^{-1} \quad \textit{for } j, k = 0, 1, \ldots, n. \tag{20}$$

Proof. Consider the term $\bar{b}_i^j(b)\bar{b}_{k-i}^{n-j}(a)$ in the sum (18) for the matrix element M_{jk}^{-1}. When $[\bar{a}, \bar{b}] \subset [a, b]$, we observe that $b - \bar{a} \geq 0$, $\bar{b} - a \geq 0$ and $\bar{b} - b \leq 0$, $a - \bar{a} \leq 0$, so by the definition (5b) of the basis functions we have

$$|\bar{b}_i^j(b)| = (-1)^{j-i}\bar{b}_i^j(b) \quad \text{and} \quad |\bar{b}_{k-i}^{n-j}(a)| = (-1)^{k-i}\bar{b}_{k-i}^{n-j}(a). \tag{21}$$

Thus, we may infer that

$$|\bar{b}_i^j(b)\bar{b}_{k-i}^{n-j}(a)| = |\bar{b}_i^j(b)||\bar{b}_{k-i}^{n-j}(a)| = (-1)^{j+k}\bar{b}_i^j(b)\bar{b}_{k-i}^{n-j}(a), \tag{22}$$

for each term $i = \max(0, j+k-n), \ldots, \min(j, k)$ of the sum (18). According to whether $j + k$ is even or odd, these terms are evidently all nonnegative or nonpositive, and hence $|M_{jk}^{-1}| = (-1)^{j+k} M_{jk}^{-1}$. □

The transformation matrices $\mathbf{M}$ and $\mathbf{M}^{-1}$ whose elements are given by (14) and (18) are quite general, applying to *arbitrary* distinct intervals $[a, b]$ and $[\bar{a}, \bar{b}]$. However, we shall henceforth be interested solely in cases where $[\bar{a}, \bar{b}] \subset [a, b]$, so that Lemmas 3 and 5 hold, and we refer to the matrices $\mathbf{M}$ corresponding to such cases as *subdivision matrices.*

3. The condition number $\kappa_\infty(\mathbf{M})$

Let $\mathbf{c} = (c_0, \ldots, c_n)^T$ and $\bar{\mathbf{c}} = (\bar{c}_0, \ldots, \bar{c}_n)^T$ denote column vectors representing the Bernstein coefficients of a polynomial $P(t)$ on the intervals $[a, b]$ and $[\bar{a}, \bar{b}]$. We are interested in the sensitivity of $\bar{\mathbf{c}}$ to perturbations in $\mathbf{c}$, i.e., in the condition of the linear map

$$\bar{\mathbf{c}} = \mathbf{M}\mathbf{c}. \tag{23}$$

It is well known (cf. [17]) that in terms of the vector norm and *subordinate* matrix norm

$$\|\mathbf{c}\|_p = \left[\sum_{i=0}^{n} |c_i|^p\right]^{1/p}, \qquad \|\mathbf{M}\|_p = \sup_{\mathbf{c}\neq\mathbf{0}} \frac{\|\mathbf{M}\mathbf{c}\|_p}{\|\mathbf{c}\|_p}, \tag{24}$$

this sensitivity may be characterized by the relation

$$\bar{\varepsilon}_p \leq \kappa_p(\mathbf{M})\varepsilon_p, \quad \text{where } \kappa_p(\mathbf{M}) = \|\mathbf{M}^{-1}\|_p\|\mathbf{M}\|_p, \tag{25}$$

and the scalar fractional error measures ε_p and $\bar{\varepsilon}_p$ for corresponding perturbations $\delta\mathbf{c} = (\delta c_0, \ldots, \delta c_n)^T$ and $\delta\bar{\mathbf{c}} = (\delta\bar{c}_0, \ldots, \delta\bar{c}_n)^T$ in $\mathbf{c}$ and $\bar{\mathbf{c}}$ are defined by:

$$\varepsilon_p = \frac{\|\delta\mathbf{c}\|_p}{\|\mathbf{c}\|_p} \quad \text{and} \quad \bar{\varepsilon}_p = \frac{\|\delta\bar{\mathbf{c}}\|_p}{\|\bar{\mathbf{c}}\|_p}. \tag{26}$$

The quantity $\kappa_p(\mathbf{M})$ is the *condition number* of the subdivision matrix $\mathbf{M}$, in the $\|\cdot\|_p$-norm.

We shall be concerned here primarily with the $\|\cdot\|_\infty$ or *maximum* norm [15]:

$$(27) \qquad \|\mathbf{c}\|_\infty = \max_{0\le i\le n} |c_i|, \qquad \|\mathbf{M}\|_\infty = \max_{0\le j\le n} \sum_{k=0}^{n} |M_{jk}|,$$

since this is most amenable to the derivation of a simple closed-form expression for the condition number. We claim also that $\kappa_\infty(\mathbf{M})$ may be regarded as offering a representative value for the general $\kappa_p(\mathbf{M})$ condition number:

Proposition 6. *The p-norm condition number of an* $(n+1)\times(n+1)$ *matrix* $\mathbf{M}$ *is bounded in terms of the* ∞*-norm value by*

$$(28) \qquad \kappa_p(\mathbf{M}) \le (n+1)^{2/p}\kappa_\infty(\mathbf{M}) \quad \textit{for all } \mathbf{M}.$$

Proof. By the *Riesz convexity theorem* [2, 14] we may bound the general matrix norm $\|\mathbf{M}\|_p$ in terms of the "simple" norms $\|\mathbf{M}\|_1$ and $\|\mathbf{M}\|_\infty$ as follows:

$$(29) \qquad \|\mathbf{M}\|_p \le \|\mathbf{M}\|_1^{1/p}\|\mathbf{M}\|_\infty^{(p-1)/p} \quad \text{for all } \mathbf{M}.$$

Now since $\|\mathbf{M}\|_\infty$ is the greatest sum of absolute values of the elements of $\mathbf{M}$ across rows (cf. (27)), and $\|\mathbf{M}\|_1$ is the greatest sum of absolute values across columns [15], we have $\|\mathbf{M}\|_1 \le (n+1)\|\mathbf{M}\|_\infty$ for all $\mathbf{M}$. Together with equation (29), this implies that $\|\mathbf{M}\|_p \le (n+1)^{1/p}\|\mathbf{M}\|_\infty$. The bound (28) follows on applying this inequality to $\mathbf{M}$ and $\mathbf{M}^{-1}$. □

Since condition numbers are in general extremely large, and one is primarily interested in their orders of magnitude rather than their precise numerical values, the factor $(n+1)^{2/p}$ in the bound (28) is not of great significance, and we may regard $\kappa_\infty(\mathbf{M})$ as being reasonably representative of the magnitude of the general $\kappa_p(\mathbf{M})$ condition number.

We are now ready to prove our main result:

Theorem 7. *For* $[\overline{a}, \overline{b}] \subset [a, b]$ *let* $\overline{m} = \frac{1}{2}(\overline{a}+\overline{b})$ *denote the midpoint of the subinterval* $[\overline{a}, \overline{b}]$, *let* $u_{\overline{m}} = (\overline{m}-a)/(b-a)$ *and* $v_{\overline{m}} = (b-\overline{m})/(b-a)$ *be the barycentric coordinates of that midpoint relative to* $[a, b]$, *and let* $f = (b-a)/(\overline{b}-\overline{a})$ *be the "zoom" factor of the subdivision map. Then for polynomials of degree* n, *the* $\|\cdot\|_\infty$*-norm condition number of the corresponding subdivision matrix* $\mathbf{M}$ *is given by*

$$(30) \qquad \kappa_\infty(\mathbf{M}) = [2f\max(u_{\overline{m}}, v_{\overline{m}})]^n.$$

Proof. As an immediate consequence of Lemmas 2 and 3, we observe that when $[\overline{a}, \overline{b}] \subset [a, b]$, the corresponding subdivision matrix $\mathbf{M}$ satisfies $\|\mathbf{M}\|_\infty = 1$, and its condition number is thus given by $\kappa_\infty(\mathbf{M}) = \|\mathbf{M}^{-1}\|_\infty$. To evaluate the latter, we note on substituting for $\overline{b}_i^j(b)$ and $\overline{b}_{k-i}^{n-j}(a)$ in (18) and making use of Lemma 5, that

(31)
$$|M_{jk}^{-1}| = (-1)^{j+k} M_{jk}^{-1}$$
$$= \sum_{i=\max(0,j+k-n)}^{\min(j,k)} \binom{j}{i} \binom{n-j}{k-i} \frac{(b-\overline{b})^{j-i}(b-\overline{a})^i(\overline{b}-a)^{n-j-k+i}(\overline{a}-a)^{k-i}}{(\overline{b}-\overline{a})^n},$$

where we have set $(-1)^{j+k}(\overline{b}-b)^{j-i}(a-\overline{a})^{k-i} = (b-\overline{b})^{j-i}(\overline{a}-a)^{k-i}$. The sum of the quantities (31) across row j of $\mathbf{M}^{-1}$ is thus given by

(32)
$$\sum_{k=0}^{n} \sum_{i=\max(0,j+k-n)}^{\min(j,k)} \binom{j}{i} \binom{n-j}{k-i} \frac{(b-\overline{b})^{j-i}(b-\overline{a})^i(\overline{b}-a)^{n-j-k+i}(\overline{a}-a)^{k-i}}{(\overline{b}-\overline{a})^n}.$$

Now for a fixed row index j and all column indices $0 \le k \le n$, let σ_{ik} be the summand of the double sum (32) if $0 \le i \le j$, and let $\sigma_{ik} = 0$ if $i < 0$ or $i > j$. Then we may simplify the summation limits on the inner sum of (32) by writing it in the form:

$$\sum_{k=0}^{n} |M_{jk}^{-1}| = \sum_{k=0}^{n} \sum_{i=j+k-n}^{k} \sigma_{ik}, \tag{33}$$

since this entails only adding spurious zeros in certain cases. Now (33) is more amenable to rearranging the order of summation:

$$\sum_{k=0}^{n} |M_{jk}^{-1}| = \sum_{i=j-n}^{n} \sum_{k=\max(0,i)}^{\min(n,n-j+i)} \sigma_{ik} = \sum_{i=0}^{j} \sum_{k=i}^{n-j+i} \sigma_{ik}, \tag{34}$$

the last step following from the fact that $\sigma_{ik} = 0$ for all $0 \le k \le n$ whenever $i < 0$ or $i > j$.

Since the final form in (34) has no zero contributions σ_{ik} with $i < 0$ or $i > j$, we may now substitute the original summand of (32) in place of σ_{ik}. Hence we have

(35)
$$\sum_{k=0}^{n} |M_{jk}^{-1}|$$
$$= \sum_{i=0}^{j} \sum_{k=i}^{n+i-j} \binom{j}{i} \binom{n-j}{k-i} \frac{(b-\overline{b})^{j-i}(b-\overline{a})^i(\overline{b}-a)^{n-j-k+i}(\overline{a}-a)^{k-i}}{(\overline{b}-\overline{a})^n}$$
$$= \frac{1}{(\overline{b}-\overline{a})^n} \sum_{i=0}^{j} \binom{j}{i} (b-\overline{b})^{j-i}(b-\overline{a})^i \sum_{l=0}^{n-j} \binom{n-j}{l} (\overline{b}-a)^{n-j-l}(\overline{a}-a)^l$$

for any row index j. Now if $\overline{m} = \frac{1}{2}(\overline{a}+\overline{b})$ denotes the midpoint of $[\overline{a}, \overline{b}]$, the sums in (35) are evidently just the binomial expansions of $[(b-\overline{b})+(b-\overline{a})]^j = [2(b-\overline{m})]^j$ and $[(\overline{b}-a)+(\overline{a}-a)]^{n-j} = [2(\overline{m}-a)]^{n-j}$, so we arrive at the simple expression

$$(36) \qquad \sum_{k=0}^{n} |M_{jk}^{-1}| = \frac{2^n(b-\overline{m})^j(\overline{m}-a)^{n-j}}{(\overline{b}-\overline{a})^n} = 2^n f^n \left[\frac{b-\overline{m}}{b-a}\right]^j \left[\frac{\overline{m}-a}{b-a}\right]^{n-j}$$

for the sum across row j, where $f = (b-a)/(\overline{b}-\overline{a})$.

Finally, if $m = \frac{1}{2}(a+b)$ denotes the midpoint of the parent interval $[a, b]$, the index j which maximizes the quantity $(b-\overline{m})^j(\overline{m}-a)^{n-j}$ in (36) over all rows is either $j=0$ or $j=n$, according to whether $\overline{m} > m$ or $\overline{m} < m$. Hence, on setting $u_{\overline{m}} = (\overline{m}-a)/(b-a)$ and $v_{\overline{m}} = (b-\overline{m})/(b-a)$, the $\|\cdot\|_\infty$-norm of $\mathbf{M}^{-1}$ is given by

$$(37) \qquad \|\mathbf{M}^{-1}\|_\infty = [2f\max(u_{\overline{m}}, v_{\overline{m}})]^n,$$

and since $\kappa_\infty(\mathbf{M}) = \|\mathbf{M}^{-1}\|_\infty$ we have the desired result (30). □

4. Practical remarks

In previous studies [5, 6] we have demonstrated the superior intrinsic stability of the Bernstein-Bézier representation for parametric polynomial curves and surfaces, as compared to the familiar power form, and discussed the feasibility of systematic computation with Bernstein-Bézier forms. This paper offers a concise characterization of the intrinsic stability of a fundamental computation which enjoys widespread application in Bézier curve and surface algorithms, namely, the *subdivision process*.

Notwithstanding its simple elegance, our motivation in deriving Theorem 7 was not purely theoretical. We therefore offer an interpretation which has immediate implications for the floating-point implementation of curve and surface algorithms:

Rule-of-thumb. Let $\mathbf{r}(t)$ be a generic Bézier curve of degree n with control points $\{\mathbf{p}_k\}$ on the parameter interval $t \in [a, b]$. If the $\{\mathbf{p}_k\}$ are specified in floating-point format with a mantissa of a finite number of digits in base β, the greatest allowable "zoom" factor $f = (b-a)/(\overline{b}-\overline{a})$ in the subdivision of $\mathbf{r}(t)$ to a parameter interval $[\overline{a}, \overline{b}] \subset [a, b]$, such that no more than r digits of precision are lost, is roughly

$$(38) \qquad f_{\max} \simeq \frac{\beta^{r/n}}{2\max(u_{\overline{m}}, v_{\overline{m}})},$$

where $u_{\overline{m}}$ and $v_{\overline{m}}$ are the barycentric coordinates of $\overline{m} = \frac{1}{2}(\overline{a}+\overline{b})$ relative to $[a, b]$.

As a simple example, consider the subdivision of cubic ($n=3$) Bézier curves with control points represented in decimal ($\beta = 10$) floating point. If it is desired that no more than $r = 6$ decimal places of accuracy be lost in the subdivision process, equation (38) suggests a maximum zoom factor of between about 50 and 100, depending on the subinterval location.

Of course, in *specific* cases the actual error amplification will often be substantially less than indicated by $\kappa_\infty(\mathbf{M})$, and subdivision could proceed to finer resolutions than suggested by this rule-of-thumb. The actual error can be estimated in particular cases by, for example, a running error analysis such as described in [5]. However, the bound (38) is nevertheless valuable as a conservative a priori indicator of safe subdivision levels for *arbitrary* Bézier curves and surfaces of a given degree.

Another important practical message arises from the exponential growth of $\kappa_\infty(\mathbf{M})$ with the polynomial degree n. High-degree $(n \geq 10)$ curves are gaining increasing popularity in certain design applications for the additional "flexibility" they offer. However, our analysis indicates a potentially severe loss of accuracy in elementary computations with such forms. The desired freedom of shape would be better attained by recourse to low-degree *piecewise polynomial* (i.e., spline) forms.

To conclude, we recall the discussion of §1 concerning subdivision termination criteria. In the context of isolating the real roots of a polynomial $P(t)$ on a given interval $[a, b]$, the appropriate criterion [12] is that the Bernstein coefficients $\bar{c}_0, \ldots, \bar{c}_n$ on each subinterval $[\bar{a}, \bar{b}]$ exhibit precisely zero or one sign change, since the *variation diminishing* property [13] then ensures exactly zero or one root on each of those subintervals. (More stringent conditions, guaranteeing the convergence of Newton's iteration [10] from any initial value $t_0 \in [\bar{a}, \bar{b}]$, are desirable if roots are to be approximated—these may be phrased in terms of sign changes in the first- and second-order *differences* of $\bar{c}_0, \ldots, \bar{c}_n$.)

Unfortunately, the simple characterization of numerical stability given by (30) does not offer sufficiently detailed information on the errors $\delta\bar{c}_0, \ldots, \delta\bar{c}_n$ in the computed coefficients $\bar{c}_0, \ldots, \bar{c}_n$ to ascertain whether they might erroneously indicate satisfaction of the root-isolation criterion. The quantity $\bar{\varepsilon}_p = \|\delta\bar{\mathbf{c}}\|_p / \|\bar{\mathbf{c}}\|_p$ measures an *average* fractional error in the coefficients, whereas the *individual* fractional errors $\delta\bar{c}_i/\bar{c}_i$ determine whether or not the number of sign changes is correct. (In particular, note that $\bar{\varepsilon}_\infty = \max(|\delta\bar{c}_i|)/\max(|\bar{c}_i|)$, where in general the error of greatest magnitude does not occur in the coefficient of greatest magnitude.) Thus, a robust test of the termination criterion demands a more detailed error analysis than given here; we hope to pursue this matter in a subsequent study.

BIBLIOGRAPHY

1. E. Cohen, T. Lyche, and R. Riesenfeld, *Discrete B-splines and subdivision techniques in computer-aided geometric design and computer graphics*, Comput. Graphics Image Processing **14** (1980), 87–111.

2. M. Cotlar and R. Cignoli, *An introduction to functional analysis*, North-Holland, Amsterdam, 1974.

3. P. de Casteljau, *Courbes et surfaces à pôles*, André Citroën Automobiles SA, Paris (unpublished notes), 1963.
4. G. Farin, *Curves and surfaces for computer aided geometric design*, Academic Press, New York, 1988.
5. R. T. Farouki and V. T. Rajan, *On the numerical condition of polynomials in Bernstein form*, Comput. Aided Geom. Design **4** (1987), 191–216.
6. ———, *Algorithms for polynomials in Bernstein form*, Comput. Aided Geom. Design **5** (1988), 1–26.
7. W. Feller, *An introduction to probability theory and its applications.* Vol. I, Wiley, New York, 1950.
8. R. N. Goldman, *Markov chains and computer-aided geometric design: Part* I—*problems and constraints*, ACM Trans. Graphics **3** (1984), 204–222.
9. ———, *Markov chains and computer-aided geometric design: Part* II—*examples and subdivision matrices*, ACM Trans. Graphics **4** (1985), 12–40.
10. P. Henrici, *Elements of numerical analysis*, Wiley, New York, 1964.
11. J. M. Lane and R. F. Riesenfeld, *A theoretical development for the computer generation and display of piecewise polynomial surfaces*, IEEE Trans. Pattern Anal. Mach. Intell. **2** (1980), 35–46.
12. ———, *Bounds on a polynomial*, BIT **21** (1981), 112–117.
13. I. J. Schoenberg, *On variation diminishing approximation methods*, On Numerical Approximation (R. E. Langer, ed.), Univ. of Wisconsin Press, Madison, WI, 1959.
14. E. M. Stein and G. L. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton Univ. Press, Princeton, N.J., 1971.
15. G. W. Stewart, *Introduction to matrix computations*, Academic Press, New York, 1973.
16. J. H. Wilkinson, *Rounding errors in algebraic processes*, Prentice-Hall, Englewood Cliffs, N. J., 1963.
17. ———, *The algebraic eigenvalue problem*, Oxford Univ. Press, 1988.

IBM Research Division, Thomas J. Watson Research Center, P. O. Box 218, Yorktown Heights, New York 10598

E-mail address, R. T. Farouki: farouki@ibm.com

E-mail address, C. A. Neff: caneff@ibm.com