New York Journal of Mathematics

New York J. Math. **13** (2007) 107–115.

# All Bézier curves are attractors of iterated function systems

**Chand T. John**

Abstract. The fields of computer aided geometric design and fractal geometry have evolved independently of each other over the past several decades. However, the existence of so-called *smooth fractals*, i.e., smooth curves or surfaces that have a self-similar nature, is now well-known. Here we describe the self-affine nature of quadratic Bézier curves in detail and discuss how these self-affine properties can be extended to other types of polynomial and rational curves. We also show how these properties can be used to control shape changes in complex fractal shapes by performing simple perturbations to smooth curves.

Contents



## 1. Introduction

In the late 1950s, advancements in hardware technology made it possible to efficiently manufacture curved 3D shapes out of blocks of wood or steel. It soon became apparent that the bottleneck in mass production of curved 3D shapes was the lack of adequate software for designing these shapes. Bézier curves were first introduced in the 1960s independently by two engineers in separate French automotive companies: first by Paul de Casteljau at Citroën, and then by Pierre Bézier at Rénault. De Casteljau never published his results, and so the theory of these curves largely contains Bézier's name. However, the algorithm used to compute points on these

Received January 22, 2001.
*Mathematics Subject Classification.* 53A.
*Key words and phrases.* Bézier, subdivision, IFS, self-affine, self-similar, polynomial, curve.

ISSN 1076-9803/07

curves bears de Casteljau's name. The theory quickly extended to more general Bézier and B-spline surfaces. Computer aided geometric design (CAGD) became an established field of study by 1974. Farin [2] gives a more complete history of Bézier curves and surfaces.

Benoit Mandelbrot [5] popularized fractals (self-similar sets) as tools for generating images of realistic objects such as mountain ranges and for modeling various physical phenomena. In 1981, John Hutchinson [3] introduced the current version of iterated function systems (IFSs) in order to establish a general theory of strictly self-similar sets. IFSs were popularized by Michael Barnsley in his book *Fractals everywhere* [1]. IFSs were shown to be useful for generating images of very complex real objects such as clouds. Although the problem of using IFSs for image compression and other applications proved to be harder than initially expected, there are still many interesting problems and approaches in the ongoing research in the field.

Here we demonstrate that these seemingly unrelated mathematical objects actually have a close relationship. Although this relationship has been known to some of the computer graphics community, applications of this knowledge and extensions to more general classes of curves and surfaces have been relatively unexplored. We show how this relationship between the two types of objects may be useful in modeling and shape control.

## 2. Quadratic Bézier curves

Throughout this paper, $\mathbb{E}^2$ denotes the set of points in the Euclidean plane, while $\mathbb{R}^2$ denotes the set of vectors in the plane. Let $P_0, P_1, \ldots, P_n \in \mathbb{E}^2$. Let $\alpha_0, \alpha_1, \ldots, \alpha_n \in [0,1]$ such that

$$\alpha_0 + \alpha_1 + \cdots + \alpha_n = 1.$$

Then the *barycenter* of the points $\{P_i\}_{i=0}^n$ with weights $\{\alpha_i\}_{i=0}^n$ is the point

$$P = \sum_{i=0}^{n} \alpha_i P_i. \tag{1}$$

Simply put, $P$ is the center of mass ("barycenter") of the points $\{P_i\}_{i=0}^n$ with weights $\{\alpha_i\}_{i=0}^n$. The process of computing a barycenter using Equation (1) is called a *barycentric combination*. Note that although addition and scalar multiplication are not defined over $\mathbb{E}^2$, Equation (1) is still valid when it is written as

$$P = P_0 + \sum_{i=1}^{n} \alpha_i (P_i - P_0),$$

since each $P_i - P_0$ is a vector in $\mathbb{R}^2$, where addition and scalar multiplication are valid operations, and the addition of a point to a vector is also valid operation.

Suppose we are given three distinct, noncollinear points $P_0$, $P_1$, and $P_2$ in the plane. Suppose we are also given a real number $t \in [0,1]$. The *de Casteljau algorithm* proceeds as follows. First compute two barycentric combinations to obtain two intermediate points:

$$P_0^1(t) = (1-t)P_0 + tP_1, \tag{2}$$

$$P_1^1(t) = (1-t)P_1 + tP_2. \tag{3}$$

Then compute a similar barycentric combination over these intermediate points:

$$P_0^2(t) = (1-t)P_0^1(t) + tP_1^1(t) = (1-t)^2 P_0 + 2t(1-t)P_1 + t^2 P_2. \tag{4}$$

The set of points $\{P_0^2(t) : t \in [0,1]\}$ is the *quadratic Bézier curve* (QBC) with *control points* $P_0$, $P_1$, and $P_2$. The triangle $\triangle P_0P_1P_2$ is called the *control polygon* of the curve. For more on the theory of Bézier curves and surfaces, see [2].

## 3. Iterated function systems

An *affine map* is a transformation $w : \mathbb{E}^2 \to \mathbb{E}^2$ such that

$$w(x,y) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}, \tag{5}$$

where $a, b, c, d, e$ and $f$ are real numbers. We may abbreviate Equation (5) with the notation $w(X) = AX + T$, where $A$ is a $2 \times 2$ matrix, $X = [\ x \quad y\ ]^T$ and $T = [\ e \quad f\ ]^T$.

An important fact is that barycentric combinations are invariant with respect to affine maps. That is, if $w$ is an affine map and $P$ is a barycenter defined as in Equation (1), then

$$w(P) = w\left(\sum_{i=0}^{n} \alpha_i P_i\right) = \sum_{i=0}^{n} \alpha_i w(P_i), \tag{6}$$

i.e., $w(P)$ is still the barycenter of the points $\{w(P_i)\}$ with weights $\{\alpha_i\}$. In fact, affine maps are precisely those maps that preserve barycentric combinations. The proof of this fact is straightforward — see page 18 in [2].

An *iterated function system* (IFS) is a set of $N$ affine maps $w_1, \ldots, w_N$. Here we will only focus on IFSs with $N = 2$, so we shall denote an IFS as a pair $\{w_1, w_2\}$ of affine maps. Let $H(\mathbb{E}^2)$ denote the set of nonempty compact subsets of $\mathbb{E}^2$. Associated with each IFS is a function $W : H(\mathbb{E}^2) \to H(\mathbb{E}^2)$ such that $W(K) = w_1(K) \cup w_2(K)$, for every $K \in H(\mathbb{E}^2)$. Let $W^{\circ n}(K)$ denote the repeated application of the map $W$ to the set $K$ a total of $n$ times.

A common restriction is to assume that $w_1$ and $w_2$ are *contractive* maps: that, for each $i = 1, 2$,

$$\|w_i(X) - w_i(Y)\| \leq s_i \cdot \|X - Y\|, \forall X, Y \in \mathbb{E}^2, \tag{7}$$

where $0 \leq s_i < 1$ is the *contractivity factor* of $w_i$. Assuming that $w_1$ and $w_2$ are contractive, then we know from the *contraction mapping theorem* [1] that $w_1$ and $w_2$ respectively have unique fixed points $X_1$ and $X_2$, and furthermore, for any $X \in \mathbb{E}^2$,

$$\lim_{n\to\infty} w_1^{\circ n}(X) = X_1 \text{ and } \lim_{n\to\infty} w_2^{\circ n}(X) = X_2. \tag{8}$$

Convergence for $w_1$ and $w_2$ are with respect to the Euclidean metric. Not only do $w_1$ and $w_2$ push every point toward their own fixed points, but also $W$ maps every $K \in H(\mathbb{E}^2)$ to *its* own unique fixed point $L = \lim_{n\to\infty} W^{\circ n}(K)$, called the *attractor* of the IFS $\{w_1, w_2\}$. Note that we can start with *any* nonempty compact set $K$ and end up with the *same* attractor $L$, for a fixed IFS. Convergence in $H(\mathbb{E}^2)$ is with respect to the Hausdorff metric. Formally, an IFS consisting only of contractive maps is called a *hyperbolic* IFS. The term "IFS" can be used to refer to an arbitrary

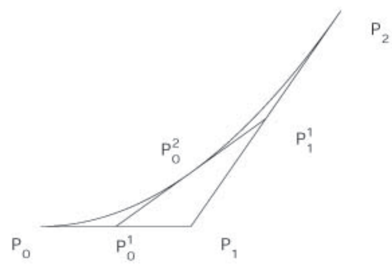


FIGURE 1. The de Casteljau algorithm is applied to a quadratic Bézier curve with control points $P_0 = (0, 0)$, $P_1 = (1/2, 0)$, and $P_2 = (1, 1)$. This curve is the graph of $y = x^2$ for $x \in [0, 1]$.

collection of maps with no condition imposed on the maps. For our purposes, we will always require an IFS to be composed of affine maps, but they need not be contractive unless explicitly stated. We shall see below that it is not necessary for $w_1$ and $w_2$ to be contraction mappings in order for $W$ to converge to the attractor of its IFS, but simply that $w_1$ and $w_2$ must mimick the general behavior of an IFS made up of contraction mappings. See [1] for a thorough treatment of IFSs.

## 4. An IFS with a QBC attractor

We now describe a connection between the two seemingly unrelated mathematical objects introduced above: QBCs and IFSs. Consider the QBC defined by $P_0 = (0, 0), P_1 = (1/2, 0)$, and $P_2 = (1, 1)$. It is easy to verify that the image of the function $P_0^2(t)$ for $t \in [0, 1]$ is the graph of $y = x^2$ for $x \in [0, 1]$.

Now suppose we were to use the de Casteljau algorithm to compute $P_0^2(u)$, where $0 < u < 1$ is an arbitrary real number. We would compute the points $P_0^1(u), P_1^1(u)$, and $P_0^2(u)$. Now define $w_1$ and $w_2$ to be the unique affine transformations satisfying

$$w_1(P_0) = P_0 \qquad w_1(P_1) = P_0^1(u) \qquad w_1(P_2) = P_0^2(u) \tag{9}$$

$$w_2(P_0) = P_0^2(u) \qquad w_2(P_1) = P_1^1(u) \qquad w_2(P_2) = P_2. \tag{10}$$

So $w_1$ maps the original control polygon $\triangle P_0P_1P_2$ to the polygon

$$T_1 = \triangle P_0P_0^1(u)P_0^2(u)$$

and $w_2$ maps the original control polygon to the polygon $T_2 = \triangle P_0^2(u)P_1^1(u)P_2$. Let $S_1$ denote the QBC whose control polygon is $T_1$ and let $S_2$ be the QBC whose control polygon is $T_2$. It is easy to verify algebraically that $S_1$ and $S_2$ are respectively the graphs of $y = x^2$ for $x \in [0, u]$ and $x \in [u, 1]$ respectively. In other words, the maps $w_1$ and $w_2$ subdivide the original curve into two subcurves that intersect in exactly one point: $P_0^2(u) = (u, u^2)$. The functions also map the original control polygon to the control polygons that generate each of the two subcurves.

We can compute $w_1$ and $w_2$ by solving a system of linear equations directly from their definition. This yields

$$w_1 \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u & 0 \\ 0 & u^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{11}$$

and

$$w_2 \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1-u & 0 \\ 2u(1-u) & (1-u)^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u \\ u^2 \end{bmatrix}. \tag{12}$$

Barnsley [1] describes a way to construct an IFS whose attractor is the graph of a function interpolating a set of points in $\mathbb{E}^2$. Formally, given data points $(x_0, y_0), (x_1, y_1), \ldots, (x_N, y_N)$ where $x_0 < x_1 < \cdots < x_N$, for some $N > 1$, define an IFS $\{w_1, \ldots, w_N\}$ satisfying the following conditions:

$$a_n = \frac{x_n - x_{n-1}}{x_N - x_0}, \tag{13}$$

$$e_n = \frac{x_N x_{n-1} - x_0 x_n}{x_N - x_0}, \tag{14}$$

$$c_n = \frac{y_n - y_{n-1} - d_n(y_N - y_0)}{x_N - x_0}, \tag{15}$$

$$f_n = \frac{x_N y_{n-1} - x_0 y_n - d_n(x_N y_0 - x_0 y_N)}{x_N - x_0}, \tag{16}$$

$b_n = 0$, and $0 \le d_n < 1$ for each $n \in \{1, \ldots, N\}$, where the variables are the coefficients of each $w_n$: $w_n(x, y) = (a_n x + b_n y + e_n, c_n x + d_n y + f_n)$. Then two facts from [1] hold:

(1) There is a metric $d$ on $\mathbb{E}^2$ equivalent to the Euclidean metric, such that the IFS is hyperbolic with respect to $d$. There is a unique nonempty compact set $S \in \mathbb{E}^2$ such that

$$S = \bigcup_{n=1}^{N} w_n(S).$$

(2) Moreover, $S$ is the attractor of this IFS, and $S$ is the graph of a continuous function $f : [x_0, x_N] \to \mathbb{R}$ interpolating the original $N + 1$ data points. $f$ is called a *fractal interpolation function*.

We shall set $N = 2$ and have the data points $(0, 0)$, $(u, u^2)$, and $(1, 1)$, and set $d_1 = u^2$ and $d_2 = (1-u)^2$. Note the resulting IFS is $\{w_1, w_2\}$ where $w_1$ and $w_2$ are defined as in Equations (11) and (12). Then if we define $W : H(\mathbb{E}^2) \to H(\mathbb{E}^2)$ such that $W(B) = w_1(B) \cup w_2(B)$ for all $B \in H(\mathbb{E}^2)$, we have from the above facts and the IFS definitions that $\{W^{\circ n}(B)\}$ converges to the QBC above (call it $S$) with respect to the metric $d$, and $S$ is the unique fixed point of $W$. In summary, we have shown how to construct a whole family of hyperbolic IFSs (parameterized by $0 < u < 1$) whose attractor is a particular QBC: the graph of $y = x^2$ for $x \in [0, 1]$.

## 5. All QBCs are attractors of IFSs

Suppose we are given three points $Q_0, Q_1, Q_2 \in \mathbb{E}^2$ that are distinct and noncollinear. Let $T$ be a QBC with control points $Q_0$, $Q_1$, and $Q_2$. Let $0 < u < 1$ be an arbitrary real number. If we let $Q_0^2(t)$ denote the point on $T$ with parameter value $t \in [0, 1]$, then define $T_1 = \{Q_0^2(t) : t \in [0, u]\}$ and $T_2 = \{Q_0^2(t) : t \in [u, 1]\}$. Define an affine map $\omega : \mathbb{E}^2 \to \mathbb{E}^2$ such that $\omega(P_i) = Q_i$ for $i = 0, 1, 2$. Clearly this map is unique and invertible. Moreover, since affine maps preserve barycentric combinations, $\omega(S) = T$, $\omega(S_1) = T_1$, and $\omega(S_2) = T_2$. Let $v_1$ and $v_2$ be the unique

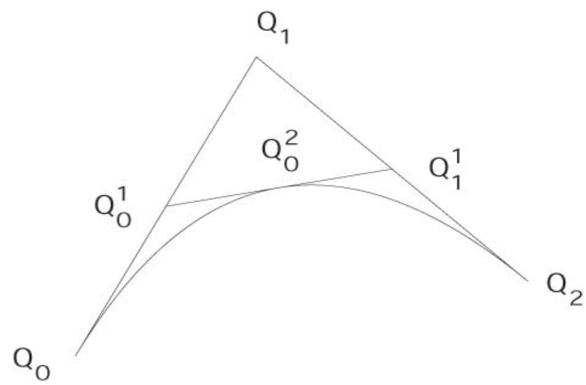


FIGURE 2. The de Casteljau algorithm is applied to a quadratic Bézier curve with control points $Q_0$, $Q_1$, and $Q_2$. The behavior of the affine maps $v_1$ and $v_2$ is analogous to the behavior of $w_1$ and $w_2$ on $y = x^2$.

affine maps mapping $T$ to $T_1$ and $T_2$, respectively. It is easy to see that

$$v_1 = \omega \circ w_1 \circ \omega^{-1}, \tag{17}$$

$$v_2 = \omega \circ w_2 \circ \omega^{-1}. \tag{18}$$

Define $V : H(\mathbb{E}^2) \to H(\mathbb{E}^2)$ such that $V(B) = v_1(B) \cup v_2(B)$ for all $B \in \mathbb{E}^2$. Clearly $V(T) = v_1(T) \cup v_2(T) = T_1 \cup T_2 = T$, so $T$ is a fixed point of $V$. It is easy to see that $V = \omega \circ W \circ \omega^{-1}$, which implies that $V^{\circ n} = \omega \circ W^{\circ n} \circ \omega^{-1}$. Now for any $B \in H(\mathbb{E}^2)$, we know $\omega^{-1}(B) \in H(\mathbb{E}^2)$, so $W^{\circ n}(\omega^{-1}(B)) \to S$ as $n \to \infty$. But since $\omega$ is continuous, $V^{\circ n}(B) = \omega(W^{\circ n}(\omega^{-1}(B))) \to \omega(S) = T$ as $n \to \infty$. Furthermore, if $A_1$ and $A_2$ are both fixed points of $V$, then $V(A_1) = A_1$ and $V(A_2) = A_2$, so $\omega(W(\omega^{-1}(A_1))) = A_1$ and $\omega(W(\omega^{-1}(A_2))) = A_2$, so $W(\omega^{-1}(A_1)) = \omega^{-1}(A_1)$ and $W(\omega^{-1}(A_2)) = \omega^{-1}(A_2)$, but since $W$ has $S$ as its unique fixed point, it follows that $S = \omega^{-1}(A_1) = \omega^{-1}(A_2)$, so $A_1 = A_2 = T$, so $V$ does have a unique fixed point $T$ to which every sequence $\{V^{\circ n}(B)\}$ converges.

Here we have proven that, even if $v_1$ and $v_2$ are not contraction mappings in a conventional sense, they still mimick the behavior of $w_1$ and $w_2$, and therefore the IFS $\{v_1, v_2\}$ still converges to its attractor, the QBC with control points $Q_0$, $Q_1$, and $Q_2$. Thus we have given a constructive proof of the following result.

**Theorem 1.** *A quadratic Bézier curve with distinct noncollinear control points $P_0$, $P_1$, and $P_2$ is the attractor of a family of iterated function systems $\{w_1, w_2\}$, where the family is parameterized by a real number $0 < u < 1$.*

## 6. Controlling fractals with Bézier curves

In computer graphics, there is a need for succinct description of a complex object's geometry and an efficient way to vary the geometry of an object to make movies. Here we show how a small number of QBCs can be used to represent a complex fractal. We also demonstrate two simple ways of continuously varying such a fractal's shape.

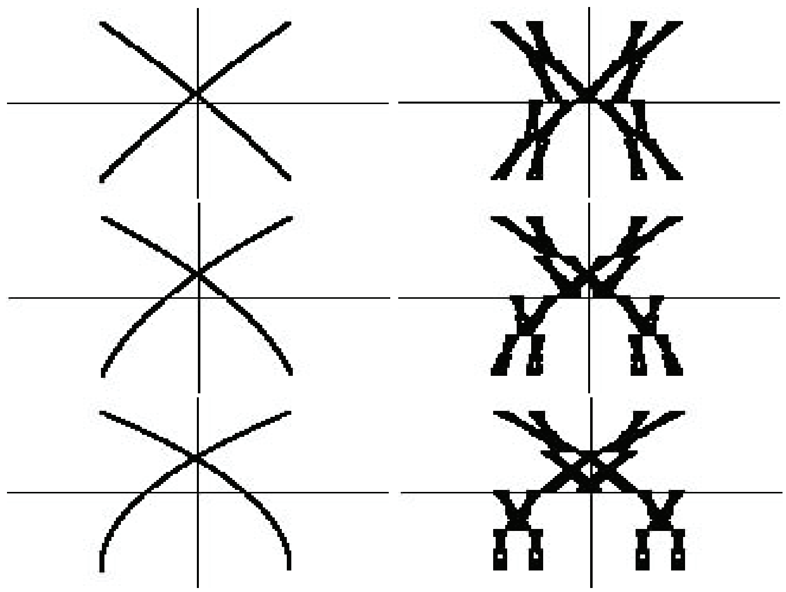

FIGURE 3. Each picture in the left column shows two curves, one with control points $(-5, 5)$, $(p, 0)$, and $(5, -5)$, and another with control points $(-5, -5)$, $(-p, 0)$, $(5, 5)$, as $p$ varies. For both curves, $u = 0.5$. Each picture in the right column shows the fractal that results from combining IFSs for the two curves in the picture next to it. The first row shows the QBCs and corresponding combined IFS's attractor when $p = 1$. The second row shows the case where $p = 3$ and the third row shows the case where $p = 5$.

We have developed a method for finding an IFS whose attractor is a given QBC. Suppose we have several such QBCs $S_1, S_2, \ldots, S_n$. Suppose that we have constructed an IFS $I_i = \{w_{2i-1}, w_{2i}\}$ whose attractor is the curve $S_i$, for $i = 1, 2, \ldots, n$. Then we can combine all the $w_j$ into one IFS $I = \{w_1, w_2, \ldots, w_{2n}\}$ whose attractor is a dusty cloud-like fractal. We can move control points of the curves $S_i$ in a smooth fashion and recalculate the attractor of the aggregate IFS $I$ to produce an animation that shows the original fractal being deformed smoothly. We would also have chosen arbitrary constraints $u_i$ for each IFS $I_i$, and these values can also be varied to smoothly deform the attractor of $I$.

Suppose we have two QBCs, one with control points $(-5, 5)$, $(1, 0)$, $(5, -5)$ and another with control points $(-5, -5)$, $(-1, 0)$, $(5, 5)$. For both curves, let $u = 0.5$. Let the affine maps for the first QBC's generating IFS be $w_1$ and $w_2$, and let the affine maps for the second QBC be $w_3$ and $w_4$. Let $I = \{w_1, w_2, w_3, w_4\}$ denote the combined IFS of the two curves. Now if we continuously move the middle control points of both curves and recompute the IFS $I$ and its attractor, we will see that the attactor's shape also changes continuously, and in a way that bears some geometric resemblance to how the curves' shapes are changing. In fact, Barnsley [1] proves that if the coefficients of an IFS are varied continuously, then the attractor of the IFS also varies its shape continuously in $H(\mathbb{E}^2)$. Since the coefficients of $I$ are continuous functions of the control point coordinates of the two curves, it follows then that continuous changes to the control point coordinates cause continuous changes to the attractor of $I$. See Figure 3 for the variations in the shape of the attractor of $I$ as the middle control points' $x$-coordinates are varied.

Suppose we now fix $p = 1$, allow $u$ to vary for both curves equally, and recompute the corresponding IFSs for each curve, and render the resulting fractal generated by the curves' combined IFS. Figure 4 shows the fractal shapes that can be generated by these varying IFSs. This demonstrates the second way that we can vary fractal shapes generated by QBCs.

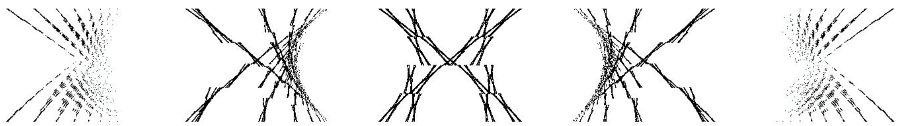

FIGURE 4. Each picture shows the attractor of the combined IFS obtained from two QBCs with control points $(-5, 5)$, $(1, 0)$, $(5, -5)$ and $(-5, -5)$, $(-1, 0)$, $(5, 5)$, as the parameter $u$ is varied for both curves. From left to right, the images show the attractor generated by the IFS when both curves have $u$ equal to 0.1, 0.3, 0.5, 0.7, and 0.9.

## 7. Conclusion and future work

We have shown that all quadratic Bézier curves are attractors of iterated function systems. We demonstrated a potential application of this result to shape representation and animation for computer graphics. We can generalize the self-affineness results beyond quadratic curves to all polynomial Bézier curves using affine maps in higher dimensions. We can parameterize a family of IFSs converging to a single curve by multiple real knots $0 < u_1 < \cdots < u_k < 1$. We can go further to characterize all *rational* Bézier curves as attractors of IFSs that contain more general projective maps instead of affine maps. It is also known that projective maps transform algebraic curves into other algebraic curves. What kinds of algebraic curves (beyond rational curves) can be characterized as attractors of IFSs consisting of projective maps, if any? And, how can these results be extended to polynomial, rational, and algebraic surfaces? These questions are natural extensions of the work presented here. In addition, there is potential for these methods to be applied to geometric modeling.

## References

[1] BARNSLEY, MICHAEL F. Fractals everywhere. Second ed. Revised with the assistance of and with a foreword by Hawley Rising, III. *Academic Press Professional, Boston, MA*, 1993. MR1231795 (94h:58101), Zbl 0784.58002.

[2] FARIN, GERALD. Curves and surfaces for computer aided geometric design. A practical guide. Firth edition. *Academic Press, Boston*, 2002. MR1412572 (97e:65022), Zbl 0919.68120.

[3] HUTCHINSON, JOHN E. Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981) 713–747. MR0625600 (82h:49026), Zbl 0598.28011.

[4] John, Chand T. Visualization by subdivision: Two applications for future graphics platforms. *Workshop on Unique Chips and Systems* (UCAS-1) **1** (2005) 12–19.

[5] Mandelbrot, Benoit B. The fractal geometry of nature. Schriftenreihe für den Referenten. [Series for the Referee]. W. H. Freeman and Company, San Francisco, 1982. MR0665254 (84h:00021), Zbl 0504.28001.

[6] Munkres, James R. Topology: A first course. *Prentice-Hall, Inc., Englewood Cliffs, N.J.*, 1975. MR0464128 (57 #4063), Zbl 0306.54001.

[7] Szeliski, R.; Terzopoulos, D. From splines to fractals, *Computer Graphics* (SIGGRAPH '89), **23** (1989) 51–60.

Computer Science Department, Stanford University, Stanford, CA 94305
ctj@stanford.edu http://www.stanford.edu/~ctj

This paper is available via http://nyjm.albany.edu/j/2007/13-7.html.