

The Calculus of Fractal Interpolation Functions

MICHAEL F. BARNSELY AND ANDREW N. HARRINGTON*

*Iterated Systems, Inc., 5550 Peachtree Parkway,
Norcross, Georgia 30092, U.S.A.*

Communicated by E. W. Cheney

Received September 13, 1985

The calculus of deterministic fractal functions is introduced. Fractal interpolation functions can be explicitly indefinitely integrated any number of times, yielding a hierarchy of successively smoother interpolation functions which generalize splines and which, just as in the case for the original fractal functions, are attractors for iterated function systems. The fractal dimension for a class of fractal interpolation functions is explicitly computed. © 1989 Academic Press, Inc.

1. INTRODUCTION

In [B] Barnsley introduced some real-valued interpolation functions, defined on a compact interval in \mathbb{R} , which appear well suited for approximating naturally occurring functions which display some sort of self-similarity under magnification. The functions are analogous to splines and polynomial interpolations in that their graphs are set to go through a finite number of prescribed points. They differ from classical interpolants in that they obey a functional relation related to self-similarity on smaller scales:

Let $x_0 < x_1 < \dots < x_N$. Let L_j be the affine map satisfying $L_j(x_0) = x_{j-1}$, $L_j(x_N) = x_j$, $j = 1, 2, \dots, N$. Let $y_0, y_1, \dots, y_N \in \mathbb{R}$. Let $-1 < \alpha_j < 1$, $j = 1, 2, \dots, N$. Let $X = [x_0, x_N]$. Let $F_j: X \times \mathbb{R} \rightarrow \mathbb{R}$ satisfy

$$|F_j(x, y_1) - F_j(x, y_2)| \leq \alpha_j |y_1 - y_2|, \quad x \in X, y_1, y_2 \in \mathbb{R}, \quad (1.0)$$

$$F_j(x_0, y_0) = y_{j-1}, \quad F_j(x_N, y_N) = y_j, \quad j = 1, 2, \dots, N. \quad (1.1)$$

* Supported in part by NSF Grant DMS 8401921.

The fractal interpolation function, abbreviated FIF, associated with $\{(L_j(x), F_j(x, y))\}_{j=1}^N$ is the unique function $f: X \rightarrow \mathbb{R}$ satisfying

$$f(L_j(x)) = F_j(x, f(x)), \quad j = 1, 2, \dots, N, \quad x \in X. \tag{1.2}$$

The function f is continuous because of the continuity conditions (1.1) at the joints. On each subinterval $[x_{j-1}, x_j]$ f interpolates between y_{j-1} and y_j , and the graph is also related to the function over the whole interval.

The graph of f is the attractor for the iterated function system on $X \times \mathbb{R}$, $\{(L_j(x), F_j(x, y))\}_{j=1}^N$ (see [BD] and the brief review below).

Barnsley’s main original examples of FIF theory were derived from affine functions F_j , and they typically were fractals—having a noninteger Hausdorff–Besicovitch dimension [M] or fractal dimension [CFMT].

In this paper we deal with FIFs where

$$F_j(x, y) = \alpha_j y + q_j(x). \tag{1.3}$$

We show that if the q_j ’s are linear, and the x_j ’s are equally spaced, and the $|\alpha_j| = \alpha$, $j = 1, 2, \dots, N$, with $1/N < \alpha < 1$, then whenever the interpolation points do not lie on a straight line, the fractal dimension of the graph of f is

$$\lim_{\delta \rightarrow 0} \frac{\log(\mathcal{N}_\delta(f))}{\log(1/\delta)} = 2 + \log_N \alpha,$$

where $\mathcal{N}_\delta(f)$ is the minimum number of $\delta \times \delta$ squares needed to cover the graph of f .

We note that the integrals of FIFs satisfying (1.3) are also FIFs. (The differentiable examples cannot actually be fractals. We retain the name fractal interpolation function because of the flavor of the scaling in the definition (1.2) and because some derivative of these functions is typically fractal.) We show that we can make C^n interpolations with FIFs associated with $n + 1$ degree polynomial q_j ’s. There are more problems specifying end conditions for these interpolants than for classical splines.

The casino functions of Dubins and Savage [DS] are a special case of FIFs derived from just two maps where the q_j ’s are constant. These and related FIFs are cumulative distribution functions for probability measures. These measures are simple cases of p -balanced measures [BD].

An outline of this paper follows. In Section 2 facts about iterated function systems and FIFs are summarized. Sections 3 and 4 concern integrals of FIFs and the inverse problem of interpolating with a C^n FIF. In Section 5 we refer to casino functions and generalizations. In Section 6 we introduce binary operations on iterated function systems for FIFs. We use this formalism in Section 7 to produce a sum representation of FIFs. We

calculate the fractal dimension of some FIFs in Section 8. Finally, in Section 9, we mention a generalization and some directions for further exploration.

2. A SPECIAL CASE OF ITERATED FUNCTION SYSTEMS

FIFs are the attractors of a special class of iterated function systems. The general definition involves a complete metric space K . Let H be the set of all nonempty compact subsets of K . Let $w_n: K \rightarrow K$ be continuous, $n = 1, 2, \dots, N$. Then $\{K, w_n: n = 1, 2, \dots, N\}$ is an iterated function system (IFS).

For any $S \in H$, let $W(S) = \bigcup_n w_n(S)$. $G \in H$ is an *attractor* for the IFS if

$$G = W(G).$$

If K is compact, then any IFS has an attractor (not necessarily unique). Starting with any $S \in H$ we can form $W(S) = W^{\circ 1}(S)$, $W^{\circ m}(S) = W \circ W^{\circ m-1}(S)$. Let G be the set of accumulation points of $\{W^{\circ m}(S)\}_{m=1}^{\infty}$.

An IFS is *hyperbolic* if

$$\frac{d(w_n(x), w_n(y))}{d(x, y)} \leq \alpha < 1 \quad \text{for all } x, y \in K, n = 1, 2, \dots, N.$$

In this case the IFS has a unique attractor,

$$G = \lim_{m \rightarrow \infty} W^{\circ m}(S).$$

A FIF is associated with a hyperbolic IFS where $K = [x_0, x_N] \times \mathbb{R}$ and $w_j(x, y) = (L_j(x), F_j(x, y))$. In this paper we consider only cases where

$$F_j(x, y) = \alpha_j y + q_j(x), \quad |\alpha_j| < 1. \quad (2.1)$$

Barnesley [B] has shown that the moments of such FIFs can be recursively calculated.

3. INTEGRAL OF FIFs AND DIFFERENTIABLE FIFs

Suppose f is a FIF satisfying (2.1). We show that its integral is also a FIF.

THEOREM 1. *If f is the FIF associated with $\{(L_j(x), F_j(x, y))\}_{j=1}^N$ where $F_j(x, y) = \alpha_j y + q_j(x)$, and*

$$\hat{f}(x) = \hat{y}_0 + \int_{x_0}^x f(t) dt, \quad (3.1)$$

then \hat{f} is the FIF associated with $\{(L_j(x), \hat{F}_j(x, y))\}_{j=1}^N$ where for $j = 1, 2, \dots, N$

$$\hat{F}_j = a_j \alpha_j y + \hat{q}_j(x),$$

$$a_j = \frac{x_j - x_{j-1}}{x_N - x_0},$$

$$\hat{q}_j(x) = \hat{y}_{j-1} - a_j \alpha_j \hat{y}_0 + a_j \int_{x_0}^x q_j,$$

$$\hat{y}_j = \hat{f}(x_j) = \hat{y}_0 + \sum_{n=1}^j a_n \left[\alpha_n (\hat{y}_N - \hat{y}_0) + \int_{x_0}^{x_N} q_n \right], \quad (3.2)$$

$$\hat{y}_N = \hat{y}_0 + \frac{\sum_{n=1}^N a_n \int_{x_0}^{x_N} q_n}{1 - \sum_{n=1}^N a_n \alpha_n}. \quad (3.3)$$

Proof.

$$\begin{aligned} \hat{f}(L_j(x)) &= \left(\hat{y}_0 + \int_{x_0}^{x_{j-1}} f \right) + \int_{x_{j-1}}^{L_j(x)} f \\ &= \hat{y}_{j-1} + a_j \int_{x_0}^x f(L_j(t)) dt. \end{aligned}$$

Now use the functional equation

$$f(L_j(x)) = \alpha_j f(x) + q_j(x);$$

$$\begin{aligned} \hat{f}(L_j(x)) &= \hat{y}_{j-1} + a_j \int_{x_0}^x (\alpha_j f(t) + q_j(t)) dt \\ &= \hat{y}_{j-1} + a_j \alpha_j (\hat{f}(x) - \hat{y}_0) + a_j \int_{x_0}^x q_j. \end{aligned}$$

Hence \hat{f} is the attractor of the stated IFS. The continuity conditions at the joints y_j must be satisfied since the antiderivative \hat{f} is continuous. Now let $x = x_N$ above:

$$\hat{y}_j - \hat{y}_{j-1} = f(L_j(x_N)) - \hat{y}_{j-1} = a_j \left[\alpha_j (\hat{y}_N - \hat{y}_0) + \int_{x_0}^{x_N} q_j \right].$$

Next (3.2) follows from $\hat{y}_j = \hat{y}_0 + \sum_{n=1}^j (\hat{y}_n - \hat{y}_{n-1})$. Substituting $j = N$ into (3.2) yields (3.3). ■

COROLLARY. Use the notation of the theorem associated with f . $\hat{f}' = f$ if and only if \hat{f} is the FIF associated with $\{(L_j(x), \hat{F}_j(x, y))\}_{j=1}^N$ where

$$\hat{F}_j(x, y) = \hat{\alpha}_j y + \hat{q}_j(x)$$

$$\frac{\hat{\alpha}_j}{a_j} = \alpha_j, \quad \hat{q}_j' = a_j q_j, \quad j = 1, 2, \dots, N.$$

Proof. The if part is immediate from the theorem. For the converse suppose $\hat{f}(0) = \hat{y}_0$ is given:

$$\begin{aligned} \hat{y}_j - \hat{y}_{j-1} &= \hat{F}_j(x_N, \hat{y}_N) - \hat{F}_j(x_0, \hat{y}_0) \\ &= \hat{\alpha}_j(\hat{y}_N - \hat{y}_0) + \hat{q}_j(x_N) - \hat{q}_j(x_0) \\ &= a_j \alpha_j (\hat{y}_N - \hat{y}_0) + a_j \int_{x_0}^{x_N} q_j. \end{aligned}$$

This is the same information that we used to uniquely determine $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N$ in the theorem. The equations $\hat{F}_j(0, \hat{y}_0) = \hat{y}_N$ then uniquely determine the constant terms in $\hat{q}_1, \hat{q}_2, \dots, \hat{q}_N$. ■

EXAMPLE. Let f be the FIF passing through $(0, 0)$, $(\frac{1}{2}, 1)$, $(\frac{3}{4}, -1)$, $(1, 0)$ with

$$\begin{aligned} L_1(x) &= \frac{1}{2}x, & F_1(x, y) &= \frac{3}{4}y + x \\ L_2(x) &= \frac{1}{2} + \frac{1}{4}x, & F_2(x, y) &= \frac{3}{4}y + 1 - 2x \\ L_3(x) &= \frac{3}{4} + \frac{1}{4}x, & F_3(x, y) &= \frac{3}{4}y - 1 + x. \end{aligned}$$

Figure 1 shows the graph of f . Figure 2 shows the graph of $\hat{f}(x) = \int_0^x f$, which passes through $(0, 0)$, $(\frac{1}{2}, \frac{7}{16})$, $(\frac{3}{4}, \frac{17}{32})$, $(1, \frac{1}{2})$ and is generated by

$$\begin{aligned} \hat{F}_1(x, y) &= \frac{3}{8}y + \frac{1}{4}x^2 \\ \hat{F}_2(x, y) &= \frac{3}{16}y + \frac{7}{16} + \frac{x}{4} - \frac{x^2}{4} \\ \hat{F}_3(x, y) &= \frac{3}{16}y + \frac{17}{32} - \frac{1}{4}x + \frac{x^2}{8}. \end{aligned}$$

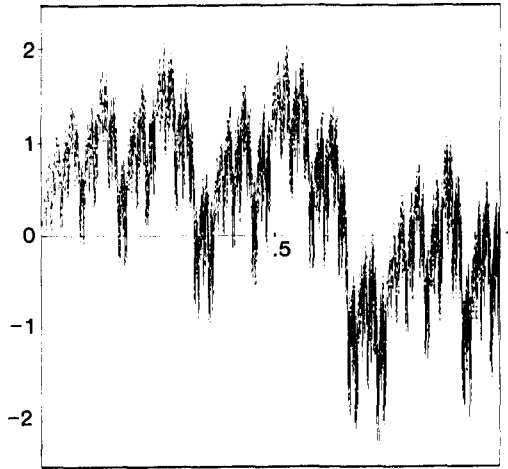


FIGURE 1

The simple method used to generate the graphs of f and \hat{f} on a microcomputer is described in [B].

If we iterate the corollary we get conditions satisfied by an n th antiderivative.

THEOREM 2. *Let $x_0 < x_1 < x_2 < \dots < x_N$. $L_j(x)$ is the affine function satisfying $L_j(x_0) = x_{j-1}$, $L_j(x_N) = x_j$, $j = 1, 2, \dots, N$. Let $a_j = L'_j = (x_j - x_{j-1}) / (x_N - x_0)$,*

$$F_j(x, y) = \alpha_j y + q_j(x), \quad j = 1, 2, \dots, N.$$

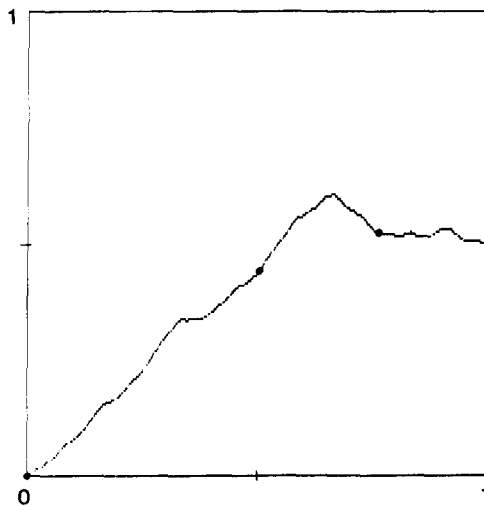


FIGURE 2

Suppose for some integer $n \geq 0$, $|\alpha_j| < a_j^n$, $q_j \in C^n[x_0, x_N]$, $j = 1, 2, \dots, N$. Let

$$F_{jk}(x, y) = \frac{\alpha_j y + q_j^{(k)}(x)}{a_j^k} \quad (\text{where } q_j^{(0)} = q_j)$$

$$y_{0,k} = \frac{q_1^{(k)}(x_0)}{a_1^k - \alpha_1}, \quad y_{N,k} = \frac{q_N^{(k)}(x_N)}{a_N^k - \alpha_N}, \quad k = 1, 2, \dots, n. \quad (3.4)$$

If $F_{j-1,k}(x_N, y_{N,k}) = F_{jk}(x_0, y_{0,k})$, $j = 2, 3, \dots, N$ and $k = 1, 2, \dots, n$, then $\{(L_j(x), F_j(x, y))\}_{j=1}^N$ determines a FIF $f \in C^n[x_0, x_N]$, and $f^{(k)}$ is the FIF determined by $\{(L_j(x), F_{jk}(x, y))\}_{j=1}^N$, $k = 1, 2, \dots, n$.

Proof. Equations (3.4) are just

$$F_{1k}(x_0, y_{0k}) = y_{0k}, \quad F_{Nk}(x_N, y_{Nk}) = y_{Nk}$$

solved for y_{0k} and y_{Nk} , so the hypotheses directly imply that $\{(L_j(x), F_{jk}(x, y))\}_{j=1}^N$ is the FIF for some continuous function g_k , $k = 0, 1, 2, \dots, n$. The corollary to Theorem 1 shows $g'_{k-1} = g_k$, $k = 1, 2, \dots, n$. ■

4. INTERPOLATING WITH C^n FIFs

In practice one is less likely to want to recognize a C^n FIF than to construct an interpolation through specified points (x_j, y_j) .

The simplest C^n FIFs are constructed from $n+1$ degree polynomials. With all α_j 's set to 0, the result is a classical spline, but the general case provides some new algebraic twists.

THEOREM 3. *Given are a positive integer n , interpolation points (x_j, y_j) $j = 0, 1, 2, \dots, N$, horizontal contraction factors $a_j = (x_j - x_{j-1})/(x_N - x_0)$, and vertical contraction factors α_j with the bound $|\alpha_j| < a_j^n$, $j = 1, 2, \dots, N$. Suppose f is a C^n FIF associated with these parameters where the q_j 's are polynomials of degree at most $n+1$. Let \mathbf{d}_j be the n dimensional vector given by $(\mathbf{d}_j)_k = f^{(k)}(x_j)$, $k = 1, 2, \dots, n$. Let $\mathbf{y} = (y_0, y_1, \dots, y_N)$. There are $n \times n$ matrices C_N and D_N and an $n \times (N+1)$ matrix E_N (computed in the proof) depending only on $a_j, \alpha_j, j = 1, 2, \dots, N$, such that*

$$C_N \mathbf{d}_0 + (D_N - I) \mathbf{d}_N + E_N \mathbf{y} = \mathbf{0}, \quad (4.1)$$

where I is the identity matrix. Either C_N or $D_N - I$ or both may be singular. They are both nonsingular if, for given a_j 's, the α_j 's are close enough to 0.

Remark. If the a_j 's and α_j 's were to be given so C_N or $D_N - I$ is singular,

then, unlike with spline interpolations, there will sometimes not be a solution for given \mathbf{y} and derivatives at one endpoint.

Proof of Theorem 3. If we let $f^{(0)} = f$ and $q_j^{(0)} = q_j$, then the formulas in Theorem 2 for the IFSs associated with f and its derivatives imply that the joint conditions to be satisfied are

$$f^{(k)}(x_j) = \frac{\alpha_j f^{(k)}(x_N) + q_j^{(k)}(x_N)}{a_j^k}, \quad \begin{array}{l} j = 1, 2, \dots, N, \\ k = 0, 1, 2, \dots, n \end{array} \quad (4.2)$$

$$f^{(k)}(x_{j-1}) = \frac{\alpha_j f^{(k)}(x_0) + q_j^{(k)}(x_0)}{a_j^k}, \quad \begin{array}{l} j = 0, 1, 2, \dots, N \\ k = 0, 1, 2, \dots, n. \end{array} \quad (4.3)$$

Let $q_j(x) = \sum_{k=0}^{n+1} q_{kj}((x - x_0)/(x_N - x_0))^k$, $j = 1, 2, \dots, N$. Define the vectors \mathbf{S}_j by

$$(\mathbf{S}_j)_k = \frac{q_j^{(k)}(x_0)}{a_j^k} = \frac{k! q_{kj}}{(x_j - x_{j-1})^k}, \quad \begin{array}{l} k = 1, 2, \dots, n, \\ j = 1, 2, \dots, N. \end{array} \quad (4.4)$$

Let β_j be the diagonal $n \times n$ matrix with $(\beta_j)_{kk} = q_j/a_j^k$, $k = 1, 2, \dots, n$, $j = 1, 2, \dots, N$. We can rewrite (4.3) when $k > 0$

$$\mathbf{d}_{j-1} = \beta_j \mathbf{d}_0 + \mathbf{S}_j, \quad j = 1, 2, \dots, N. \quad (4.5)$$

For $k = 0$ we obtain

$$y_{j-1} = \alpha_j y_0 + q_{0j}$$

or

$$q_{0j} = y_{j-1} - \alpha_j y_0, \quad j = 1, 2, \dots, N. \quad (4.6)$$

From (4.2) with $k = 0$ we find

$$y_j = \alpha_j y_N + \sum_{i=0}^{n+1} q_{ij}, \quad j = 1, 2, \dots, N. \quad (4.7)$$

We can use (4.6) and (4.7) to solve for the highest order coefficient in the q_j 's:

$$q_{n+1,j} = y_j - \alpha_j y_N - (y_{j-1} - \alpha_j y_0) - \sum_{i=1}^n q_{ij}. \quad (4.8)$$

Eliminate q_{0j} and $q_{n+1,j}$ from the expression for q_j :

$$q_j(x) = \sum_{i=1}^n q_{ij} \left[\left(\frac{x-x_0}{x_N-x_0} \right)^i - \left(\frac{x-x_0}{x_N-x_0} \right)^{n+1} \right] + (y_{j-1} - q_j y_0) \\ + [(y_j - y_{j-1}) - \alpha_j(y_N - y_0)] \left(\frac{x-x_0}{x_N-x_0} \right)^{n+1}. \quad (4.9)$$

Divide by a_j^k and differentiate using the notation $n_r = n!/(n-r)!$:

$$\frac{q_j^{(k)}(x_N)}{a_j^k} = \frac{1}{(x_j - x_{j-1})^k} \left\{ \sum_{i=k}^n q_{ij} i_k - \sum_{i=1}^n (n+1)_k q_{ij} \right. \\ \left. + (n+1)_k [(y_j - y_{j-1}) - q_j(y_N - y_0)] \right\}, \quad \begin{matrix} j=1, 2, \dots, N, \\ k=1, 2, \dots, n. \end{matrix} \quad (4.10)$$

If we replace q_{ij} by $(\mathbf{S}_j)_i((x_j - x_{j-1})^i/i!)$ and let \mathbf{t}_j be the vector with

$$(\mathbf{t}_j)_k = \frac{q_j^{(k)}(x_N)}{a_j^k}, \quad k=1, 2, \dots, n,$$

then (4.10) can be written as matrix equations

$$\mathbf{t}_j = T_j \mathbf{S}_j + P_j \mathbf{y}, \quad j=1, 2, \dots, N,$$

where

$$(T_j)_{ki} = \begin{cases} \frac{i_k - (n+1)}{i!} (x_j - x_{j-1})^{i-k}, & 1 \leq k \leq i \leq n \\ -\frac{(n+1)_k}{i!} (x_j - x_{j-1})^{i-k}, & 1 \leq i < k \leq n \end{cases}$$

and the $n \times (N+1)$ matrix P_j is defined by

$$(P_j \mathbf{y})_k = \frac{(n+1)_k}{(x_j - x_{j-1})^k} [(y_j - y_{j-1}) - \alpha_j(y_N - y_0)], \quad k=1, 2, \dots, N.$$

Now we can rewrite (4.2) when $k > 0$:

$$\mathbf{d}_j = \beta_j \mathbf{d}_N + T_j \mathbf{S}_j + P_j \mathbf{y}, \quad j=1, 2, \dots, N \quad (4.11)$$

Substitute for \mathbf{S}_j using (4.5)

$$\mathbf{d}_j = \beta_j \mathbf{d}_N + T_j(\mathbf{d}_{j-1} - \beta_j \mathbf{d}_0) + P_j \mathbf{y}. \quad (4.12)$$

We wish to define matrices C_j , D_j , and E_j so

$$\mathbf{d}_j = C_j \mathbf{d}_0 + D_j \mathbf{d}_N + E_j \mathbf{y}. \tag{4.13}$$

Starting from $\mathbf{d}_0 = I\mathbf{d}_0 + 0\mathbf{d}_N + 0\mathbf{y}$ we have

$$C_0 = I, \quad D_0 = 0, \quad E_0 = 0.$$

Then we use (4.12) to recursively calculate

$$\begin{aligned} C_j &= T_j(C_{j-1} - \beta_j) \\ D_j &= T_j D_{j-1} + \beta_j, \quad j = 1, 2, \dots, N \\ E_j &= T_j E_{j-1} + P_j. \end{aligned} \tag{4.14}$$

Taking $j = N$ in (4.13) is equivalent to (4.1) of the theorem.

Once \mathbf{y} , \mathbf{d}_0 , and \mathbf{d}_N are specified consistent with (4.1) then we can always calculate the q_j 's determining f . Use (4.14) and (4.13) to obtain \mathbf{d}_j 's, and (4.5) for \mathbf{S}_j 's, and (4.4) and (4.9) for $q_j(x)$.

If all the α_j 's are 0 then we reduce to the case of classical splines, and C_N and $D_N - I$ are invertible. Since these matrices depend continuously on the α_j 's, they are invertible for small enough α_j 's. ■

Example 1, below, shows that C_N and $D_N - I$ can be singular.

EXAMPLE 1. We take the simplest case, where $n = 1$. All the matrices but E_j and P_j reduce to numbers:

$$\begin{aligned} T_j &= -1, \quad \beta_j = \frac{\alpha_j}{a_j} \\ C_N &= (-1)^N + \beta_N - \beta_{N-1} + \dots + (-1)^{N-1} \beta_1 \\ D_N - I &= -1 + \beta_N - \beta_{N-1} + \dots + (-1)^{N-1} \beta_1. \end{aligned}$$

If each β_j is the same (there is a consistent ratio of horizontal and vertical scalings), then C_N and $D_N - I$ will not be singular. In other cases with $N \geq 3$ they may be.

Take $N = 3$. Both C_N and $D_N - I$ are 0 if

$$0 = -1 + \frac{\alpha_3}{a_3} - \frac{\alpha_2}{a_2} + \frac{\alpha_1}{a_1}.$$

This would be true if each $a_j = \frac{1}{3}$, $\alpha_1 = \alpha_3 = \frac{2}{9}$, $\alpha_2 = \frac{1}{9}$. In this extreme case there is a C^1 FIF f only if $y_0 - 3y_1 + 3y_2 - y_3 = 0$, to make $E_2 \mathbf{y} = 0$.

With N any even number ≥ 4 , the two conditions for singularity would not be the same, but each is attainable.

EXAMPLE 2. Interpolate through $(0, 0)$, $(\frac{1}{2}, \frac{1}{3})$, $(1, 0)$ with $\alpha_1 = \alpha_2 = \frac{1}{3}$ and make a C^2 FIF:

$$T_1 = T_2 = \begin{bmatrix} -2 & -\frac{1}{4} \\ -12 & -2 \end{bmatrix}, \quad \beta_1 = \beta_2 = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{4}{3} \end{bmatrix},$$

$$P_1 \mathbf{y} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} = -P_2 \mathbf{y}$$

$$C_1 = \begin{bmatrix} -\frac{6}{5} & -\frac{1}{20} \\ -\frac{36}{5} & -\frac{2}{3} \end{bmatrix}, \quad D_1 = \begin{bmatrix} \frac{2}{5} & 0 \\ 0 & \frac{4}{5} \end{bmatrix}, \quad E_1 \mathbf{y} = \begin{bmatrix} 2 \\ 8 \end{bmatrix},$$

$$C_2 = \frac{1}{5} \begin{bmatrix} 25 & 2 \\ 168 & 15 \end{bmatrix}, \quad D_2 = -\frac{1}{5} \begin{bmatrix} 7 & 1 \\ 24 & 9 \end{bmatrix}, \quad E_2 \mathbf{y} = \begin{bmatrix} -8 \\ -48 \end{bmatrix}.$$

If we choose to make $\mathbf{d}_0 = 0$ then

$$\mathbf{d}_2 = -D_2^{-1} E_2 \mathbf{y} = -\frac{40}{13} \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$

$$\mathbf{d}_1 = \frac{2}{13} \begin{bmatrix} 5 \\ -44 \end{bmatrix}$$

$$\mathbf{S}_1 = 0, \quad q_{11} = q_{21} = 0$$

$$\mathbf{S}_2 = \mathbf{d}_1, \quad q_{12} = \frac{5}{13}, \quad q_{22} = -\frac{11}{13}$$

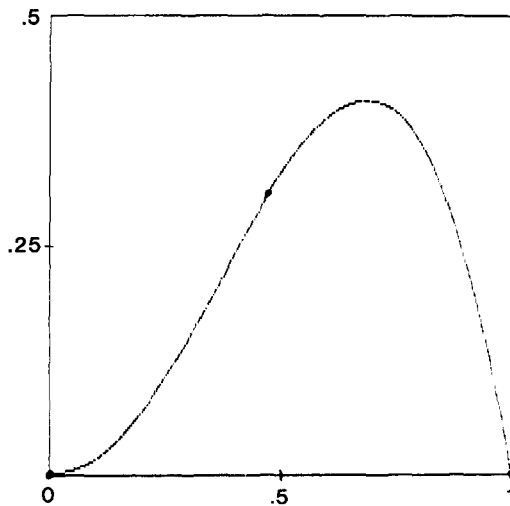


FIGURE 3

$$F_1(x, y) = \frac{1}{5}y + \frac{x^3}{3}$$

$$F_2(x, y) = \frac{1}{5}y + \frac{1}{3} + \frac{5}{13}x - \frac{11}{13}x^2 + \frac{5}{39}x^3$$

$$L_1(x) = \frac{1}{2}x, \quad L_2(x) = \frac{1}{2} + \frac{1}{2}x.$$

Figure 3 is the computer generated graph of f .

5. CUMULATIVE DISTRIBUTION FUNCTIONS

A special case of FIFs has been discussed by Dubins and Savage [DS, Chaps. 5 and 6]. The graph of f goes through $(0, 0)$, (x_1, y_1) , $(1, 1)$ with $0 < y_1 < 1$. $\alpha_1 = y_1$ and $\alpha_2 = 1 - y_1$ are chosen so $q_1(x)$ and $q_2(x)$ are constants, 0 and y_1 . $w_j(x, y) = (L_j(x), F_j(y))$ maps the unit square onto the rectangular regions I and II in Fig. 4 for $j = 1, 2$. The resulting FIF is a continuous, increasing function, and hence a cumulative distribution function for some measure.

The idea can be generalized to have N intervals and $0 = y_0 < y_1 < \dots < y_N = 1$. In this case $w_j(x, y)$ maps the unit square onto the rectangle with diagonally opposite corners (x_{j-1}, y_{j-1}) and (x_j, y_j) . The resulting FIF is again a cumulative distribution function. The associated probability measure is the p -balanced measure [BD] arising from the IFS on $[0, 1]$, $\{L_j(x)\}_{j=1}^N$, with the probability $\alpha_j = y_j - y_{j-1}$ associated with the map L_j .

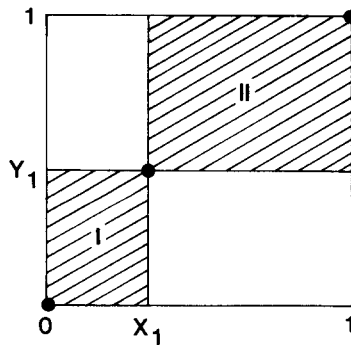


FIGURE 4

6. BINARY OPERATIONS AND SHAPE COMBINATIONS

We discuss some algebraic properties of the iterated function systems that generate FIFs. We fix the space $K = [x_0, x_N] \times \mathbb{R}$ and use a shorthand notation for an IFS:

$$W = \{w_j\}_{j=1}^N, \quad \text{where } w_j(x, y) = (L_j(x), F_j(x, y)), F_j(x, y) = \alpha_j y + q_j(x).$$

We use the same symbol for the operator on compact sets $W(S) = \bigcup_{j=1}^N w_j(S)$. We use U , V , and Z to denote other IFSs for FIFs on the same space K . Rather than introducing more new characters for all the subsidiary functions and parameters associated with U , V , and Z , we will subscript. For example, Z consists of N_Z functions, $Z = \{Z_j\}_{j=1}^{N_Z}$, W consists of $N = N_W$ functions,

Of course W has the same attractor as $W \circ W$, which also can be expressed as

$$W \circ W = \{w_j \circ w_k\}_{j,k=1}^N.$$

If a fractal function exhibits one pattern first on a larger scale and a different one on a smaller scale, associated with W and Z , respectively, then we might want to consider a hybrid

$$W \circ Z = \{W_j \circ Z_k\}_{j=1, k=1}^{N_W, N_Z}$$

which is again an IFS but its attractor is a FIF only if the initial and final interpolation points for W match those for Z .

We wish to modify the composition operation so the result is always an IFS for a FIF. First we introduce the normalization of W , denoted \tilde{W} , defined to shift y coordinates so $\tilde{y}_0 = \tilde{y}_N = 0$.

The coordinate change is

$$B(x, y) = (x, y + b(x)),$$

where $b(x)$ is the affine function satisfying $b(x_0) = y_0$, $b(x_N) = y_N$. Then $\tilde{W} = \{\tilde{w}_j\}$ where $\tilde{w}_j = B^{-1}w_jB$.

We can now define the shifted composition $W \odot Z$ to end up with the coordinate system of W :

$$W \odot Z = \{B_W \circ \tilde{W}_j \circ \tilde{Z}_k \circ B_W^{-1}\} = \{W_j \circ B_W \circ B_Z^{-1} \circ Z_k \circ B_Z \circ B_W\}.$$

This shifted composition is clearly associative, so we could define a semigroup.

There is an obvious additional operation for W and Z , if they have the same x_j 's and α_j 's. If

$$\begin{aligned} W_j(x, y) &= (L_j(x), \alpha_j y + q_{W_j}(x)) \\ Z_j(x, y) &= (L_j(x), \alpha_j y + q_{Z_j}(x)), \quad j = 1, 2, \dots, N \end{aligned}$$

then the sum of the FIFs associated with these IFSSs will be the FIF associated with $U = W \oplus Z$, defined

$$U_j(x, y) = (L_j(x), \alpha_j y + q_{W_j}(x) + q_{Z_j}(x)).$$

We can use a lower triangular matrix of functions to represent W_j and B

$$W_j \simeq \begin{bmatrix} L_j & 0 \\ q_j & \alpha_j \cdot \mathcal{I} \end{bmatrix}, \quad B \simeq \begin{bmatrix} \mathcal{I} & 0 \\ b & \mathcal{I} \end{bmatrix},$$

where \mathcal{I} is the identity function.

Matrix multiplication then involves the composition of component functions. We can make a consistent definition of \oplus for lower triangular matrices of functions with the same main diagonal,

$$\begin{bmatrix} b & 0 \\ c & l \end{bmatrix} \oplus \begin{bmatrix} b & 0 \\ d & l \end{bmatrix} = \begin{bmatrix} b & 0 \\ c+d & l \end{bmatrix}.$$

LEMMA.

$$\prod_{j=1}^m \left\{ \begin{bmatrix} b_j & 0 \\ c_j & l_j \end{bmatrix} \oplus \begin{bmatrix} b_j & 0 \\ d_j & l_j \end{bmatrix} \right\} = \prod_{j=1}^m \begin{bmatrix} b_j & 0 \\ c_j & l_j \end{bmatrix} \oplus \prod_{j=1}^m \begin{bmatrix} b_j & 0 \\ d_j & l_j \end{bmatrix}$$

if b_j, c_j, d_j, l_j are functions $\mathbb{R} \rightarrow \mathbb{R}$ with l_j strictly linear.

Proof. First the \oplus operation on the R.H.S. makes sense since the main diagonal of the product of lower triangular matrices is independent of the off-diagonal elements. One may easily show by induction that the off-diagonal element of a product depends linearly on the off-diagonal elements of the matrices in the product. That is reflected in the \oplus operations in the lemma. ■

There is a similar statement relating \odot and \oplus for the IFSSs we have been discussing.

THEOREM 4. *If $W \oplus Z$ and $U \oplus V$ are well defined by the definition above, then*

$$(W \oplus Z) \odot (U \oplus V) = (W \odot U) \oplus (Z \odot V).$$

Proof. If $W = \tilde{W}$ and $Z = \tilde{Z}$, then the theorem is a direct translation of the previous lemma with $m = 2$. In the general case, however, there must be a change of coordinate matrices included at either end. Since

$$B_{W \oplus Z} = B_W \oplus B_Z, \quad B_{W \oplus Z}^{-1} = B_W^{-1} \oplus B_Z^{-1}$$

the theorem follows from the lemma with $m = 4$. ■

Given W we can define 0_W by $W \oplus 0_W = W$, thus 0_W is the function system with x_j 's and α_j 's matching W , but with all q_j 's identically to zero.

Theorem 4 is easily equivalent to the collection of identities

1. $W \odot Z = W \odot 0_Z \oplus 0_W \odot Z$,
2. $0_U \odot (W \oplus Z) = (0_U \odot W) \oplus (0_U \odot Z)$,
3. $(W \oplus Z) \odot 0_U = (W \odot 0_U) \oplus (Z \odot 0_U)$.

Identity 1 says that any shifted composition can be written simply as a sum. In the next section we give a geometric interpretation to identity 1 and use it to analyze FIFs.

7. SUM REPRESENTATIONS OF FIFs

Recall from Section 2 that if we are given an IFS $W = \{w_j\}_{j=1}^N$, and if S is a nonempty compact set, then $W^{\circ n}(S)_{n \rightarrow \infty}$ approaches the attractor for W . We suppose W is associated with a FIF, so the attractor is the graph of a function f with $f(x_0) = y_0$, $f(x_N) = y_N$. If g is any continuous function satisfying these end conditions, then there is another function satisfying these end conditions so W of the graph of g is the graph of h . We use the notation

$$W \odot g = h.$$

In a manner analogous to the way we defined shifted compositions for IFSs and FIFs, we can extend the definition of $W \odot g$ to allow any $g \in C[x_0, x_N]$.

Let b_g be the affine function matching the values of g at x_0 and x_N . Define $\tilde{g} = g - b_g$ so \tilde{g} equals 0 at the endpoints. If f is the FIF for W then $b_W = b_f$. We define

$$W \odot g = \tilde{W} \odot \tilde{g} + b_g = \tilde{W} \odot (g - b_g) + b_W = W \odot (g - b_g + b_W).$$

We obtain results like the ones in the last section

$$\begin{aligned} (W \odot Z) \odot g &= W \odot (Z \odot g) \\ (W \oplus Z) \odot (g + h) &= W \odot g + W \odot h \\ (W \oplus Z) \odot g &= W \odot g + Z \odot 0. \end{aligned}$$

If $f_0 \in C[x_0, x_N]$, $f_j = W \odot f_{j-1}$, $j = 1, 2, \dots$, then the FIF for W , $f = \lim f_j$. Since the limit is independent of f_0 , we can take f_0 to be just b_W (or equivalently 0).

Then $f_1 = W \odot f_0$ interpolates between the points (x_j, y_j) .

Using identity 1 from the end of the last section we find

$$\begin{aligned} f_2 &= W \odot W \odot b_W = [(W \odot 0_W) \oplus (0_W \odot W)] \odot b_W \\ &= W \odot (0_W \odot b_W) + 0_W \odot (W \odot b_W) \\ &= W \odot b_W + 0_W \odot f_1 \\ &= f_1 + 0_W \odot f_1. \end{aligned} \tag{7.1}$$

The graph of $0_W \odot g$ has the graph of \tilde{g} compressed sideways into each interval $[x_{j-1}, x_j]$ and scaled vertically by α_j , $j = 1, 2, \dots, N$. Let $S_1 = f_1$, $S_j = 0_W \odot S_{j-1}$, $j = 2, 3, \dots$.

The manipulations in (7.1) can be extended with more shifted compositions with W to obtain

$$f_k = \sum_{n=1}^k S_n.$$

Hence

$$f = \sum_{n=1}^{\infty} S_n. \tag{7.2}$$

We use the representation in the next section.

8. THE FRACTAL DIMENSION OF SOME FIFS

For continuous real valued functions f on $[x_0, x_N]$ let

$\mathcal{N}_\delta(f)$ = the number of $\delta \times \delta$ square regions
required to cover the graph of f .

THEOREM 5. *Let f be the FIF interpolating between $(x_0, y_0), \dots, (x_N, y_N)$*

with affine q_j 's, $x_j - x_{j-1}$ constant, and $|\alpha| = \alpha$, $j = 1, 2, \dots, N$, $1/N < \alpha < 1$. If the interpolation points do not lie on a straight line then

$$\lim_{\delta \rightarrow 0} \frac{\log \mathcal{N}_\delta(f)}{\log 1/\delta} = 2 + \log_N^\alpha.$$

Remark. This limit is called the fractal dimension of the graph of f [CFMT]. We use the notation $D(f)$ for it.

This is a fairly easy number of approximate with a computer, unlike the Hausdorff-Besicovitch dimension. If we are looking for a simple mathematical model for a given fractal function, we would specify that a sequence of points match and use an approximation of the fractal dimension to choose α .

For the proof of Theorem 5 we use two simple lemmas.

LEMMA 1. $\mathcal{N}_\delta(f) < (1/\delta)((x_N - x_0) + \int_{x_0}^{x_N} |df(x)|) + 1$.

Proof. $\mathcal{N}_\delta(f) - 1 < 1/\delta \cdot \text{length of graph} < 1/\delta((x_N - x_0) + \int_{x_0}^{x_N} |df|)$ by the triangle inequality. ■

LEMMA 2. $\mathcal{N}_\delta(\sum_j f_j) \leq 2 \sum_j \mathcal{N}_\delta(f_j)$.

Proof. (1) If we require that the x coordinates of each box go exactly from one multiple of δ to the next, then we claim $\sum_j \mathcal{N}_\delta(f_j) \geq \mathcal{N}_\delta(\sum_j f_j)$. Because of our extra assumption we can treat each x -interval $[j\delta, (j+1)\delta]$ separately. Consider just one such interval. The height of the image interval of $\sum f_j$ is less than the sum of the height for each f_j separately. The claim follows from the fact that the number of boxes in a column required to cover an image of height d is exactly the smallest integer $\geq d/\delta$.

(1) If we do not assume the above restriction on x -coordinates, then we can cover any unrestricted box with two restricted ones, and double the bound. ■

Proof of Theorem 5. We can make an affine change of coordinates if necessary to assume $y_0 = y_N = 0$. Let

$$\Delta_j = y_j - y_{j-1}, \quad \Delta = \max |A_j|, \quad V = \sum |A_j|.$$

(1) First we show $D(f) \leq 2 + \log_N^\alpha$. We use the notation of the last section for f_k and S_k . We find

$$\int_{x_0}^{x_N} |dS_j| = (N\alpha)^{j-1} V,$$

$$\int_{x_0}^{x_N} |df_n| \leq V \cdot \frac{(N\alpha)^n - 1}{N\alpha - 1}.$$

Let

$$t_n = f - f_n = \sum_{j=n+1}^{\infty} S_j.$$

$$\|t_n\|_{\infty} \leq \sum_{j=n+1}^{\infty} \|S_j\|_{\infty} \leq \sum_{j=n+1}^{\infty} \alpha^j \frac{V}{2} = \frac{\alpha^{n+1}(V/2)}{1-\alpha}.$$

Hence the graph of t_n is contained in a rectangle of length $(x_N - x_0)$ and height $(\alpha^{n+1}V)/(1-\alpha)$, so $N_{\delta}(t_n) \leq K\alpha^n/\delta^2$ for some constant K .

We may apply Lemmas 1 and 2 and choose large enough constants K_j so

$$\mathcal{N}_{\delta}(f) = \mathcal{N}_{\delta}(t_n + f_n) \leq \frac{K_1 \alpha^n}{\delta^2} + \frac{K_2 (N\alpha)^n}{\delta} + \frac{K_3}{\delta}.$$

We make $n = k \ln \delta$, with k chosen so the exponent on δ in the first two terms is the same. Since

$$\alpha^n = \delta^{k \ln \alpha}, \quad (N\alpha)^n = \delta^{k \ln(N\alpha)}$$

we obtain

$$k \ln \alpha - 2 = k \ln(N\alpha) - 1$$

$$k = -1/\ln N$$

$$\mathcal{N}_{\delta}(f) \leq K_4 \left(\frac{1}{\delta}\right)^{2 + \ln \alpha / \ln N} + K_3 \left(\frac{1}{\delta}\right)$$

$$D(f) \leq 2 + \frac{\ln \alpha}{\ln N} = 2 + \log_N^{\alpha}.$$

(2) Now we obtain the corresponding lower bound for $D(f)$. Use I as a multi-index (i_1, i_2, \dots, i_n) , $|I| = n$, $L_I = L_{i_1} \circ L_{i_2} \circ \dots \circ L_{i_n}$. On $L_I[x_0, x_N]$, S_1, \dots, S_n and f_n are affine functions, $L_I[x_0, x_N]$ is $1/N^{n-1}$ of $L_{i_1}[x_0, x_N]$. On $L_{i_1}[x_0, x_N]$ S_1 varies by Δ_{i_1} . On $L_I[x_0, x_N]$, S_1 varies by $1/N^{n-1} \Delta_{i_1}$. On $L_{i_1} \circ L_{i_2}[x_0, x_N]$ S_2 varies by $\alpha_{i_1} \Delta_{i_2}$. On $L_I[x_0, x_N]$ S_2 varies by $1/N^{n-2} \alpha_{i_1} \Delta_{i_2} \dots$.

For the general case define σ_{Ij} as the sign making

$$\sigma_{Ij}^{j-1} = \sigma_{i_1} \sigma_{i_2} \dots \sigma_{i_{j-1}}.$$

Then on $L_I[x_0, x_N]$ S_j varies by $(\sigma_{Ij} \alpha^{j-1} / N^{n-j}) \Delta_{i_j}$, $j = 1, 2, \dots, n$.

The total jump of f_n on $L_I[x_0, x_N]$ is

$$\frac{1}{N^{n-1}} \sum_{j=1}^n \sigma_{Ij} (N\alpha)^{j-1} \Delta_{i_j}.$$

Since f has the same terminal values as f_n on $L_I[x_0, x_N]$, the size of its range on this interval has at least as large a magnitude. Hence the total number of squares used to cover the graph of f on this interval is at least

$$\frac{1}{\delta} \left| \frac{1}{N^{n-1}} \sum_{j=1}^n \sigma_{I_j}(N\alpha)^{j-1} \Delta_{I_j} \right|.$$

If we take $\delta \leq 1/N^n$, then

$$\mathcal{N}_\delta(f) \geq \frac{1}{2\delta} \sum_{|I|=n} \frac{1}{N^{n-1}} \left| \sum_{j=1}^n \sigma_{I_j}(N\alpha)^{j-1} \Delta_{I_j} \right|,$$

where the $\frac{1}{2}$ is included because a square may be doubled counted if it covers parts of two adjacent intervals. To bound some of the absolute values of sums in the last inequality we note

$$\left| \sum_{j=1}^n \sigma_{I_j}(N\alpha)^{j-1} \Delta_{I_j} \right| \leq \frac{(N\alpha)^n - 1}{N\alpha - 1} \Delta.$$

Since $N\alpha > 1$ there is an m such that for all n large enough

$$\left| \sum_{j=1}^{n-m} \sigma_{I_j}(N\alpha)^{j-1} \Delta_{I_j} \right| \leq \left| \sum_{j=n-m+1}^n (-N\alpha)^j \Delta \right|.$$

Suppose $\max |\Delta_j|$ is attained for $j=l$. Then by considering only all multi-indices whose last m entries are l

$$\mathcal{N}_\delta(f) \geq \frac{1}{2\delta} \sum_{|I|=n-m} \frac{1}{N^{n-1}} \left| \sum_{j=1}^{n-m} \sigma_{I_j}(N\alpha)^{j-1} \Delta_{I_j} + \sum_{j=n-m+1}^n (N\alpha_l)^{j-1} \Delta_l \right|,$$

where the sum for all I with $|I|=n-m$ can be replaced by the number of terms, N^{n-m} , times the average term. Since the sign of the part inside the absolute value is the same for all I and since the average of the D_j 's is zero, we find that the first summation inside the absolute values drops out of the average, so

$$\mathcal{N}_\delta(f) \geq \frac{1}{2\delta} N^{1-m} \left| \sum_{j=n-m+1}^n (N\alpha_l)^j \Delta \right| \geq \frac{K_5}{\delta} (N\alpha)^n, \quad \delta \leq N^{-n}.$$

Use $\delta = N^{-n}$ so $n = -\log_N \delta$ and $\alpha^n = \delta^{-\log_N \alpha}$ to obtain

$$\mathcal{N}_\delta(f) \geq \frac{K_5}{\delta^2 + \log_N \alpha}$$

$$D(f) \geq 2 + \log_N \alpha. \quad \blacksquare$$

9. OTHER DIRECTIONS AND GENERALIZATIONS

We make one immediate generalization concerning the coordinate change matrices B , and then suggest other possible generalizations and related investigations.

We assumed $b(x)$ was an affine function, but any continuous function satisfying

$$b(x_0) = y_0, \quad b(x_N) = y_N \tag{9.1}$$

would work, except for the formulas with the \oplus operation to work we need the added restriction

$$B_{W+Z} = B_W \oplus B_Z. \tag{9.2}$$

For a C^n FIF, the sum representation still works, but the summands are not generally C^n . This can be remedied by requiring

$$b_W^{(k)}(x_0) = f^{(k)}(x_0), \quad b_W^{(k)}(x_N) = f^{(k)}(x_N), \quad k = 1, 2, \dots, n. \tag{9.3}$$

If b_W is chosen as the polynomial of degree at most $2n + 1$ satisfying both (9.1) and (9.3), then (9.2) will also hold.

Now we suggest some other avenues.

Much more could be said about the new classes of smooth interpolation functions. For what sort of endpoint conditions, in what circumstances is there a unique solution? Is there a minimum energy interpretation as with cubic splines?

We allowed the scaling in the y direction to include a reflection (if $\alpha_j < 0$). What if we allow the horizontal scalings to also reverse direction? For some j , $L_j(x_0) = x_j$ and $L_j(x_N) = x_{j-1}$ (x_j and x_{j-1} are switched). There should be reasonable generalizations of Section 3–8. What happens to moment theory [B]?

Another generalization would be to allow x_0, x_1, \dots, x_N not to be a monotonic sequence. The attractor would be a continuous curve Γ (but not the graph of a function) passing in order through (x_j, y_j) , $j = 0, 1, \dots, N$. We believe $\int_{\Gamma} y \, dx$ should still be calculable. It should be expressible in terms of signed measure on the attracting curve (negative where dx is negative). Can moments be calculated? Does the original formalism need much change?

REFERENCES

- [B] M. F. BARNESLEY, Fractal functions and interpolation, *Constr. Approx.* **2**, 303–329 (1986).
- [BD] M. F. BARNESLEY AND S. G. DEMKO, Iterated function systems and the global construction of fractals, *Proc. Roy. Soc. London Ser. A* **399**, 243–274 (1985).
- [CFMT] P. CONSTANTIN, C. FOIAS, O. P. MANLEY, AND R. TEMAN, Determining modes and fractal dimension of turbulent flows, *J. Fluid Mech.* **150**, 427–440 (1985).
- [DS] L. DUBINS AND L. SAVAGE, “How to Gamble if You Must,” McGraw-Hill, New York, 1965.
- [M] D. MANDELBROT, “The Fractal Geometry of Nature,” Freeman, San Francisco, 1982.