Physica D 50 (1991) 478–492
North-Holland

"Missing moment" and perturbative methods for polynomial iterated function systems

E.R. Vrscay and D. Weil

Department of Applied Mathematics, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Received 24 May 1990
Revised manuscript received 11 February 1991
Accepted 14 February 1991
Communicated by J. Guckenheimer

An iterated function system (IFS) over a compact metric space X is defined by a set of contractive maps w_i: X → X, $i = 1, \dots, N$, with associated nonzero probabilities $p_i > 0$, $\Sigma p_i = 1$. The "parallel" action of the maps defines a unique compact invariant attractor set $A \subset X$ which supports an invariant measure μ and which is balanced with respect to the p_i. For linear w_i on $X \subset \mathbb{R}$, the invariance of μ yields a relation between the moments $g_n = \int x^n \, d\mu$ which permits their recursive computation from the initial value $g_0 = 1$. For *nonlinear* w_i, however, the moment relations are incomplete and do not permit a recursive computation. This paper describes two methods of obtaining accurate estimates of the moments when the IFS maps w_i are polynomials: (i) application of the necessary Hausdorff conditions on the g_i to obtain convergent upper and lower bounds and (ii) a perturbation expansion approach. The methods are applied to some model problems.

1. Introduction

The iterated function system (IFS) has become a powerful tool for the construction as well as the analysis of typically fractal sets. The idea of constructing such sets by repeated (parallel) action of contractive maps goes back to Hutchinson [1]. The contraction mapping principle was used to show the existence of an invariant "attractor" set A which supports an invariant measure μ. Quite independently, Barnsley and Demko [2] developed this method from a probabilistic outlook of Markov walks on the attractor. (For a comprehensive treatment of the IFS, the reader is referred to Barnsley's recent textbook [3].)

The invariance properties of the measure μ imply a relation between the power moments $g_n = \int x^n \, d\mu$, $n = 0, 1, \dots$. This relation is complete only when the IFS maps are affine. In this case, the g_n may be computed explicitly in a recursive manner. When the maps are *nonlinear*, however, the relations are not complete and the moments cannot be computed explicitly. In this paper, we outline two algorithms to compute converging approximations to the moments for polynomial IFS. Following the spirit of earlier work by Handy and Bessis [10] on quantum mechanical eigenvalue problems, the first method makes use of the necessity that the g_n satisfy sequences of Hausdorff inequalities. We first isolate a set of indeterminate, or *missing* moments for the system and express all other moments in terms of these quantities. The application of a set of Hausdorff conditions on the moments results in a linear programming problem involving the missing moments. As the number of conditions increases, the upper and lower bounds on the missing moments are observed to converge to common limits. In the second method, the nonlinear IFS is considered as a perturbation of a corresponding linear affine IFS. The moments of the former are expanded as power series in the perturbation parameter ϵ, viz. $g_k = g_k(\epsilon)$. The coefficients of these power series may then be computed in a recursive manner. The series have non-zero radii of convergence, and their partial sums afford accurate estimates of the moments $g_k(\epsilon)$. The procedure is quite analogous to difference equation methods associated with the standard Rayleigh–Schrödinger perturbation theory of quantum mechanics.

The paper is structured as follows. Section 2 outlines the mathematical setting and the basic theory behind IFS, as well as introducing the nonlinear IFS which are to be studied. Section 3 focuses on the moment relations for IFS, and demonstrates the breakdown of recursiveness for nonlinear IFS. The method of missing moments is described in section 3.2 and applied to some model problems. A perturbation series approach to calculate the moments for nonlinear IFS is outlined in section 4, and applied to the model problems of section 3.2 for comparison.

2. Brief review of IFS

In the discussion below, (X, d) denotes a compact metric space with metric d. In applications, X is usually a bounded subset of $\mathbb{R}^D$ $(D = 1, 2)$. Later in this paper, we shall focus on the specific case $\mathrm{X} = [0, 1]$.

Let $\boldsymbol{w} = \{w_1, w_2, \ldots, w_N\}$ denote a set of N continuous contraction maps w_i: $\mathrm{X} \to \mathrm{X}$, so that for some s, $0 \le s < 1$,

$$d(w_i(x), w_i(y)) \le s d(x, y), \quad \forall x, y \in \mathrm{X}, \quad i = 1, 2, \ldots, N. \tag{2.1}$$

From refs. [1–3] there exists a unique, nonempty, compact set A, the *attractor*, which is invariant under the "parallel" action

$$\mathrm{A} = \boldsymbol{w}(\mathrm{A}) = \bigcup_{i=1}^{N} w_i(\mathrm{A}), \tag{2.2}$$

where $w_i(\mathrm{A}) = \{w_i(x), x \in \mathrm{A}\}$. Now let there be associated with the maps w_i a set of non-zero probabilities $\boldsymbol{p} = \{p_1, p_2, \ldots, p_N\}$, $p_i > 0$ and $\sum_i p_i = 1$, whose role will be seen below. The system $\{\mathrm{X}, \boldsymbol{w}, \boldsymbol{p}\}$ will be referred to as a *contractive IFS*.

In order to describe the relevant measures, let $\mathscr{A}$ denote the σ-algebra generated by the set of Borel subsets $\mathscr{B}(\mathrm{X})$, and let $\mathscr{M}$ denote the set of all finite measures with domain $\mathscr{A}$. In particular, we consider $\overline{\mathscr{M}}$, the subset of probability measures $\overline{\mathscr{M}} = \{\mu \in \mathscr{M}: \mu(\mathrm{X}) = 1\}$. From refs. [1–3], there exists a unique measure $\mu \in \overline{\mathscr{M}}$, called the ($\boldsymbol{p}$-balanced) *invariant measure*, which obeys the relation

$$\mu(\mathrm{B}) = (M\mu)(\mathrm{B}) = \sum_{i=1}^{N} p_i \mu(w_i^{-1}(\mathrm{B})), \quad \mathrm{B} \in \mathscr{B}(\mathrm{X}), \tag{2.3}$$

where $w_i^{-1}(\mathrm{B}) \equiv \{x \in \mathrm{X}: w_i(x) \in \mathrm{B}\}$. Moreover, the support of μ is A, i.e. $\mathrm{supp}(\mu) = \mathrm{A}$.

The following relation for integration over A with respect to the invariant measure is a noteworthy consequence of eq. (2.3): for μ-integrable functions f: $\mathrm{X} \to \mathbb{R}$,

$$\int_{\mathrm{A}} f(x)\,\mathrm{d}\mu(x) = \sum_{i=0}^{N} p_i \int_{\mathrm{A}} (f \circ w_i)(x)\,\mathrm{d}\mu(x). \tag{2.4}$$

From a more dynamical viewpoint, the role of the probabilities $\boldsymbol{p}$ in determining the invariant measure can be seen in the context of the random iteration algorithm, or "chaos game" [3], which provides a convenient and fast method of generating "pictures" of A: Pick an $x_0 \in \mathrm{X}$ and define the iteration sequence

$$x_{n+1} = w_{\sigma_n}(x_n), \quad n = 0, 1, 2, \ldots, \tag{2.5}$$

where the index σ_n is chosen randomly and independently from the set of indices $\{1, 2, \ldots, N\}$, with probabilities $P(\sigma_n = i) = p_i$. Almost every orbit $\{x_n\}$ is dense on A. Moreover, Elton [4] has shown that the chaos game is ergodic, in the Birkhoff sense: If we consider the time-averaged distribution of the first $n+1$ points in the orbit of eq. (2.5), i.e.

$$\nu_n = \frac{1}{n+1} \sum_{k=0}^{n} \delta_{x_k}, \tag{2.6}$$

where δ_x denotes a unit mass measure at x, then the ν_n converge weakly to the invariant measure μ. Hence, the chaos game provides a picture not only of the attractor A but also of the measure μ. This follows from the property $\int f\,\mathrm{d}\nu_n \to \int f\,\mathrm{d}\mu$: Let $f = \chi_{\mathrm{B}}$, the characteristic function of the subset $\mathrm{B} \in \mathrm{X}$ represented by a pixel $P(i, j)$ on a computer screen, then the probability of visiting $P(i, j)$ during the chaos game is $\mu(\mathrm{B})$. In this way, a "histogram" representation of the measure may be obtained by counting the number of visits made to each pixel. Examples will be shown below.

In practical applications (in fact, in *most* treatments to date), the IFS maps w_i are taken to be linear transformations. Since our attention here will be restricted to problems on the real line, the most general linear IFS maps assume the form

$$w_i(x) = s_i x + a_i, \quad |s_i| < 1, \quad i = 1, \ldots, N. \tag{2.7}$$

The moment theory associated with affine IFS on $\mathbb{R}$ has been examined, starting with ref. [5] and followed by more complete developments in refs. [6, 7]. In what follows, however, we consider the moment theory for nonlinear, specifically, polynomial IFS maps of the form

$$w_i(x) = \sum_{k=0}^{n_i} c_{ik} x^k, \quad n_i \geq 1, \quad i = 1, \ldots, N, \quad n_{\max} = \max_i (n_i) > 1, \tag{2.8}$$

where the contractivity property of (2.1) is still assumed to hold.

We conclude this section with some examples, first comparing some simple nonlinear IFS on $\mathrm{X} = [0, 1]$, with their "linear analogues": both systems define the same geometrical attractor A, however the measures supported on A are drastically different.

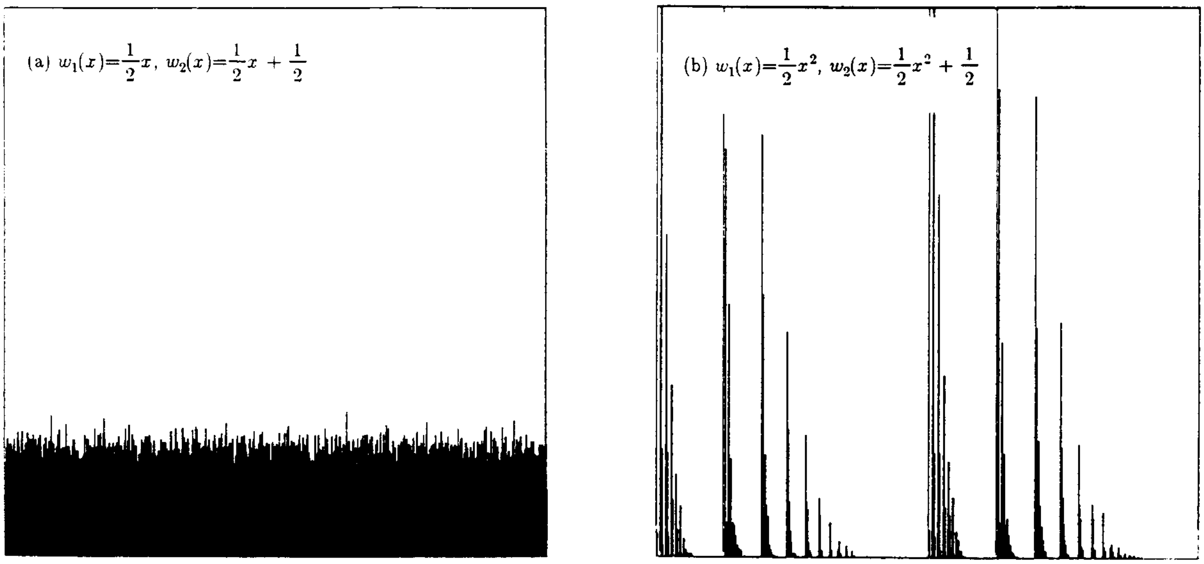


Fig. 1. Histogram approximations to invariant measures on [0, 1] for (a) linear IFS, (b) nonlinear IFS. In both cases $p_1 = p_2 = \frac{1}{2}$.

(1) Linear IFS:

$$w_1(x) = sx, \quad w_2(x) = sx + (1-s), \quad 0 < s \leq \tfrac{1}{2}.$$

(2) Nonlinear IFS:

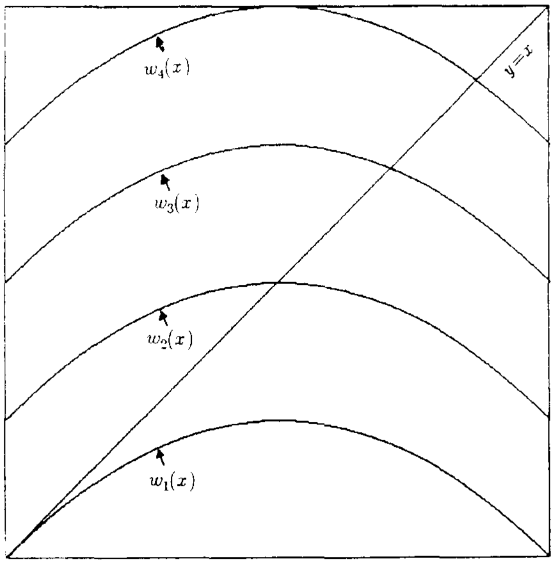
$$w_1(x) = sx^2, \quad w_2(x) = sx^2 + (1-s), \quad 0 < s \leq \tfrac{1}{2}.$$

In both cases, let $p_1 = p_2 = \frac{1}{2}$. Note that the condition $s \leq \frac{1}{2}$ is necessary for strict contractivity of the nonlinear maps on [0, 1]. (For $\frac{1}{2} < s < 1$, the structure of A is altered with the introduction of a new fixed point for w_2 which migrates away from $x = 1$). Both IFS perform the same geometric action (dissection) on X: w: $[0,1] \to [0,s] \cup [1-s,1]$, hence their attractors are identical. Two particular cases are (i) $s = \frac{1}{2}$, A = [0, 1], (ii) $s = \frac{1}{3}$, A is the ternary Cantor set C on [0, 1].

In fig. 1 are shown histogram approximations of the invariant measures for both the linear and nonlinear IFS given above, with $s = \frac{1}{2}$. For the linear case (fig. 1a), μ is uniform Lebesgue measure on [0, 1]. This is not the case for its nonlinear counterpart (fig. 1b). The quadratic IFS is very weakly contractive to the immediate left of the point $x = \frac{1}{2}$, accounting for the low visitation, hence measure μ, in this area (and its forward images under the actions of the two IFS maps). When $s = 1/3$, μ is the uniform Cantor–Lebesgue measure on the ternary Cantor set C (see, for example ref. [8], p. 77) for the linear IFS. The invariant measure for the nonlinear Cantor case also demonstrates the nonuniformity seen in fig. 1b.

The nonlinear IFS maps on $X \subset \mathbb{R}$ need not be monotonic (i.e. homeomorphic) as are their linear counterparts in eq. (2.7). As an example, consider the IFS on [0, 1] composed of the following contractive maps on [0, 1]:

$$w_i(x) = x - x^2 + \tfrac{1}{4}(i-1), \quad p_i = \tfrac{1}{4}, \quad i = 1,2,3,4. \tag{2.9}$$



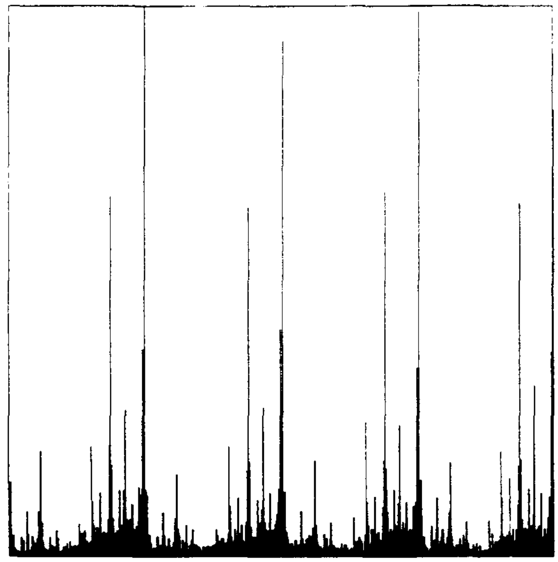


Fig. 2. (a) nonlinear IFS on [0, 1] composed of four "logistic" maps given in eq. (2.15); (b) Resulting invariant measure μ for $p_i = \frac{1}{4}$.

These "logistic"-type maps are sketched in fig. 2a. Each map $w_i(x)$ contracts $A = [0, 1]$ to the subinterval $A_i = [\frac{1}{4}(i-1), \frac{1}{4}i]$. The dynamics on [0, 1] associated with the iteration of each of these maps, hence the entire IFS, is more complicated than that for linear maps. A histogram approximation of the invariant measure μ supported on the attractor of this measure is presented in fig. 2b.

3. Moment relations for IFS

If $\{X, \boldsymbol{w}, \boldsymbol{p}\}$ is a contractive IFS in $\mathbb{R}^D$, with attractor A, then we define the power moments of the associated invariant measure by the (Lebesgue) integrals

$$g_{i_1 i_2 \ldots i_D} = \int_A x_1^{i_1} x_2^{i_2} \ldots x_D^{i_D} \, d\mu . \tag{3.1}$$

For convenience, the measure is assumed to be normalized, i.e. $g_{00\ldots0} = \int_A d\mu = 1$. In usual applications of the IFS method, including the inverse problem, linear affine transformations are used since the geometry associated with such maps is quite simple. It is also well known that for affine IFS, at least in one dimension, the power moments are easily computed recursively, with the use of the invariance property of eq. (2.4), as shown below.

3.1. Linear IFS

First consider the general linear IFS on $\mathbb{R}$ as defined by the maps

$$w_i(x) = s_i x + a_i, \quad |s_i| < 1, \quad i = 1, 2, \ldots, N, \tag{3.2}$$

with associated probabilities p_i. Setting $f(x)=x^n$ in eq. (2.4), we have

$$g_n = \int_A x^n \, d\mu(x) = \sum_{i=1}^{N} p_i \int_A (s_i x + a_i)^n \, d\mu(x), \quad n = 1,2,\ldots . \tag{3.3}$$

Expanding the polynomial, collecting like powers in x and integrating, we obtain the following well-known recursion relation [2–7]

$$\left(1 - \sum_{i=1}^{N} p_i s_i^n\right) g_n = \sum_{j=1}^{n} \binom{n}{j} g_{n-j} \left(\sum_{i=1}^{N} p_i s_i^{n-j} a_i^j \right), \quad n = 1,2,\ldots . \tag{3.4}$$

By the assumptions in our definition of a contractive IFS, the coefficient of g_n on the left cannot vanish. Thus, the moments may be computed explicitly and uniquely by this recursion formula, with the initial value $g_0 = 1$. Conversely, since the attractor A is bounded, the moment problem is determinate, and the infinite sequence of moments g_n, $n = 0,1,2\ldots$ determines a unique probability measure [9]. (Formulas for recursive computation of the moments g_{ij} for the more complicated two-dimensional case have also been derived [10].)

Examples

$X = [0,1]$, $N = 2$, $w_1(x) = sx$, $w_2(x) = sx + (1-s)$, $p_1 = p_2 = \frac{1}{2}$, with $0 \le s < 1$. The first five moments are given by

$$g_0 = 1, \quad g_1 = \tfrac{1}{2}, \quad g_2 = \frac{1}{2(1+s)}, \quad g_3 = \frac{2-s}{4(1+s)}, \quad g_4 = \frac{1+s^2-s^3}{2(1+s)(1+s+s^2+s^3)} .$$

Some special cases:

(1) $s = \frac{1}{2}$: $A = [0,1]$, μ = uniform Lebesgue measure, $g_n = \int_0^1 x^n \, dx = 1/(n+1)$.

(2) $s = 0$: $A = \{0\} \cup \{1\}$, $\mu = \frac{1}{2}(\delta_0 + \delta_1)$, where δ_a denotes unit mass measure at $x = a$, $g_0 = 1$, $g_n = \frac{1}{2}$, $n = 1,2,\ldots$.

(3) $s = \frac{1}{3}$: A = ternary Cantor set on $[0,1]$, μ = (uniform) Cantor–Lebesgue measure,

$$g_0 = 1, \quad g_1 = \tfrac{1}{2}, \quad g_2 = \tfrac{3}{8}, \quad g_3 = \tfrac{5}{16}, \quad g_4 = \tfrac{87}{320}, \quad g_5 = \tfrac{31}{128}, \quad g_6 = \tfrac{10215}{46592} .$$

The partial derivatives of the moments with respect to the IFS parameters s_i, a_i, p_i can also be computed recursively in closed form by differentiating eq. (3.4) implicitly [10, 11]. In principle, derivatives to arbitrary orders can be calculated.

3.2. Polynomial IFS and missing moments

When any or all of the maps constituting a contractive IFS are nonlinear, the recursive property between the power moments generally breaks down. To illustrate, consider the simplest nonlinear analogue of the IFS given in eq. (3.2), and scaled on $X = [0,1]$ (without loss of generality):

$$w_i(x) = s_i x^2 + a_i, \quad |s_i| \le \tfrac{1}{2}, \quad i = 1,2,\ldots,N, \tag{3.5}$$

with associated probabilities p_i. Note that the condition $|s_i| \leq \frac{1}{2}$ guarantees contractivity of the maps on $[0,1]$. From eq. (2.4), with $f(x)=x^n$,

$$g_n = \int_\Lambda x^n \,\mathrm{d}\mu(x) = \sum_{i-1}^{N} p_i \int_\Lambda (s_i x^2 + a_i)^n \,\mathrm{d}\mu(x), \quad n = 1,2,\dots . \tag{3.6}$$

Setting $g_0 = 1$, the first three equations corresponding to $n = 1,2,3$ in eq. (3.6) are

$$\begin{aligned} g_1 &= g_2 \sum p_i s_i + \sum p_i a_i, \\ g_2 &= g_4 \sum p_i s_i^2 + 2 g_2 \sum p_i s_i a_i + \sum p_i a_i^2, \\ g_3 &= g_6 \sum p_i s_i^3 + 3 g_4 \sum p_i s_i^2 a_i + 3 g_2 \sum p_i s_i a_i^2 + \sum p_i a_i^3, \end{aligned} \tag{3.7}$$

where the summations range over $i = 1,\dots,N$. The relations are insufficient to permit a recursive computation of the moments. Each value of n introduces a new set of moments g_k for $n+1 \leq k \leq 2n$. Note, however, that the sequence of odd moments can be considered as independent variables, since all even moments may be written as *linear* functions of the odd ones, i.e.

$$g_{2k} = g_{2k}(g_1, g_3, \dots, g_{[(2k-1)/2]}), \tag{3.8}$$

where $[x] = \mathrm{int}(x)$ denotes the integer part of x. In the spirit of Handy and Bessis [12], the sequence of unknown odd moments will be referred to as *missing moments*. We now proceed to find approximations to these unknown variables, using the following result from the Hausdorff moment problem [9]:

Let g_n, $n = 0,1,2,\dots$ denote an infinite sequence of real numbers. A necessary and sufficient condition that there exist a unique measure μ on $[0,1]$, such that

$$g_n = \int_0^1 x^n \,\mathrm{d}\mu,$$

is that the g_n satisfy the following inequalities:

$$I(m,n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k g_{m+k} \geq 0, \quad m,n = 0,1,2,\dots . \tag{3.9}$$

(The equality holds only when μ consists of point masses at $x = 0$ and/or 1. Without loss of generality, this degenerate case will be ignored.) Trivially, for $n = 0$, $I(m,0) = g_m > 0$, while the next set $I(m,1) > 0$ implies the nonincreasing property $g_m > g_{m+1}$. As shown below, the Hausdorff conditions may now be applied, several at a time, to a finite set of moments to produce bounds on the missing moments (hence bounds to the even moments). We expect that the bounds on the g_{2k-1} improve as the number of Hausdorff conditions is increased.

At any one time, we consider a finite number $M > 0$ of missing moments. Let $\boldsymbol{x}$ denote the M-vector of missing moments, i.e.

$$\boldsymbol{x} = (x_1,\dots,x_M)^{\mathrm{T}} = (g_1,\dots,g_{2M-1})^{\mathrm{T}}. \tag{3.10}$$

For the particular nonlinear IFS problem of eq. (3.5), this vector uniquely defines the moment sequence g_i, $i = 0, 1, \ldots, 2M$ (cf. eq. (3.8)). The only Hausdorff inequalities which employ the missing moments in eq. (3.10) are $I(m, n)$ for which $1 \le m + n \le 2M$, a total of $N_{\max} = (M+1)(2M+1)$. (The trivial case $I(0, 0)$ is ignored.) Some, or all, of these inequalities are then applied to the moments, producing a set of linear inequalities in terms of the missing moments:

$$\mathbf{A}\,x > b, \tag{3.11}$$

where $\mathbf{A}$ is an $N \times M$ matrix, and $N \le N_{\max}$ is the number of inequalities employed. The determination of upper and lower bounds to each missing moment $x_i = g_{2i-1}$ then becomes a linear programming problem:

$$\text{maximize (minimize) } S(x) = x_i, \text{ subject to } \mathbf{A}x > b.$$

These bounds are achieved on vertices (extreme points) of the convex polytope defined by the intersection of the hyperplanes defined by the inequalities in eq. (3.11).

Sample calculations

The special case of the nonlinear IFS of eq. (3.5) will be studied:

$$w_1(x) = sx^2, \quad w_2(x) = sx^2 + (1-s), \quad p_1 = p_2 = \tfrac{1}{2}. \tag{3.12}$$

Case 1. $s = \frac{1}{2}$ (cf. fig. 1b). The first few even moments are given by

$$g_2 = 2g_1 - \tfrac{1}{2}, \quad g_4 = 6g_1 - 2, \quad g_6 = 8g_3 - 12g_1 + \tfrac{13}{4}, \quad g_8 = -16g_3 + 98g_1 - 32. \tag{3.13}$$

The first set of terms $I(m, n)$, $0 \le m \le 3$, $1 \le n \le 4$, in terms of missing moments, is presented in table 1. The symbolic manipulation language MAPLE [13] was used to compute the inequalities shown this table. It proved to be quite useful, since the complexity of these expressions increases rapidly. (Note that the term $I(0, 2) = \frac{1}{2}$ adds no information to the bounding procedure.) A little algebra reveals that the nonincreasing conditions $g_1 > g_2 > g_4$ imply the bounds

$$\tfrac{1}{3} < g_1 < \tfrac{3}{8}. \tag{3.14}$$

For a relatively low number $M > 0$ of missing moments, bounds to the moments in rational or real arithmetic may be computed with the linear programming package provided in MAPLE (Version 4.3).

Table 1
Hausdorff inequalities $I(m, n)$ of eq. (3.9) for $0 \le m \le 3$, $1 \le n \le 4$, in terms of the missing moments g_1 and g_3, as applied to the nonlinear IFS: $w_1(x) = \frac{1}{2}x^2$, $w_2(x) = \frac{1}{2}x^2 + \frac{1}{2}$, $p_1 = p_2 = \frac{1}{2}$.

m	$I(m,n)$ $n = 1$	$n = 2$	$n = 3$	$n = 4$
0	$1 - g_1$	$\frac{1}{2}$	$-\frac{1}{2} + 3g_1 - g_3$	$-4 + 14g_1 - 4g_3$
1	$\frac{1}{2} - g_1$	$1 - 3g_1 + g_3$	$\frac{7}{2} - 11g_1 + 3g_3$	
2	$-\frac{1}{2} + 2g_2 - g_3$	$-\frac{5}{2} + 8g_1 - 2g_3$		
3	$2 - 6g_1 + g_3$			

Table 2
Upper and lower bounds to missing moments $g_1, \ldots, g_{2M+1}$ yielded by application of Hausdorff inequalities $I(m,n), 1 \le m+n \le 2M$. IFS: $w_1(x) = \frac{1}{2}x^2$, $w_2(x) = \frac{1}{2}x^2 + \frac{1}{2}$, $p_1 = p_2 = \frac{1}{2}$. Final column: numerical estimates of the moments from the method of eq. (3.15).

	$M=2$	$M=3$	$M=4$	$M=5$	g_{2k-1}
g_1	0.375000	0.352778	0.346963	0.346046	0.345701
	0.312500	0.340190	0.344832	0.345472	
g_3	0.250000	0.137500	0.121495	0.117626	0.116424
	0.000000	0.099138	0.113084	0.115620	
g_5		0.116667	0.063785	0.051893	0.048834
		0.000000	0.038932	0.046905	
g_7			0.058411	0.032330	0.027473
			0.000000	0.016168	
g_9				0.030487	0.011441
				0.000000	

For larger values of M, where computation in MAPLE becomes very tedious, the linear programming problem was run in FORTRAN, using the IMSL subroutine ZX4LP. The results obtained for $M = 2, 3, 4, 5$ are presented in table 2.

The entries in the final column of table 2 are accurate estimates of the moments obtained by exploiting the following property [14]:

$$T^n f(x) \to \int f \, d\mu, \quad x \in X. \tag{3.15}$$

Here, μ is the invariant measure of the IFS $\{X, \boldsymbol{w}, \boldsymbol{p}\}$. The operator T: C(X) $\to$ C(X), where C(X) denotes the space of continuous functions on X, is given by

$$T(f)(x) = \sum_1^N p_i (f \circ w_i)(x). \tag{3.16}$$

The operator M in eq. (2.3) is the adjoint operator of T [14]. (Using eq. (3.15), Bessis and Demko [15] calculated integrals such as $\int \sqrt{x} \, d\mu$ over an IFS attractor A.) The iterate in eq. (3.15) is given by the nested sum

$$T^n f(x) = \sum_{i_1=1}^{N} \cdots \sum_{i_n=1}^{N} p_{i_1} \cdots p_{i_n} f(w_{i_1} \circ \cdots \circ w_{i_n})(x), \tag{3.17}$$

which amounts to enumerating an N-tree to n generations. For $n = 15$, the estimates $y_n = T^n f(x)$, where $f(x) = x^{2k-1}$, agreed to better than one part in 10^7 in all cases. The accuracy of these estimates was quite independent of the choice of the initial seed x.

Case 2. $s = \frac{1}{3}$, A = ternary Cantor set on [0, 1] with nonuniform measure. Table 3 lists numerical results of the missing moment method applied to this problem.

Table 3
Upper and lower bounds to missing moments $g_1, \ldots, g_{2M-1}$ for IFS: $w_1(x) = \frac{1}{3}x^2$, $w_2(x) = \frac{1}{3}x^2 + \frac{2}{3}$, $p_1 = p_2 = \frac{1}{2}$.

	$M = 2$	$M = 3$	$M = 4$	$M = 5$	g_{2k-1}
g_1	0.444444 0.428571	0.439270 0.436274	0.438105 0.437876	0.438029 0.437986	0.438011
g_3	0.333333 0.142857	0.251863 0.235294	0.246095 0.243628	0.245225 0.244854	0.245069
g_5		0.219448 0.088235	0.172310 0.152670	0.164676 0.162922	0.163943
g_7			0.153100 0.067284	0.125103 0.107772	0.117811
g_9				0.112681 0.056061	0.089197

4. Perturbation theory for polynomial IFS on $\mathbb{R}$

In this section, a perturbation series approach to the calculation of moments of polynomial IFS on $\mathbb{R}$ is outlined. The computation of series coefficients is recursive in nature, providing increasingly accurate estimates of the moments. It is assumed that the contractive nonlinear maps may be written in the form

$$w_i(x) = w_i^{(0)}(x) + \epsilon w_i^{(1)}(x), \quad i = 1, 2, \ldots, N, \tag{4.1}$$

which will be written compactly as $w = w^{(0)} + \epsilon w^{(1)}$. The superscripts (0) and (1) identify, respectively, the "unperturbed" and "perturbed" parts of the problem: in eq. (4.1) these are the linear and nonlinear parts of the $w_i(x)$ in eq. (4.1), which will take the forms

$$w_i^{(0)} = s_i x + a_i, \quad |s_i| \leq 1, \quad i = 1, 2, \ldots, \tag{4.2}$$

$$w_i^{(1)} = \sum_{k=2}^{n_i} c_{ik} x^k, \quad n_i \geq 2. \tag{4.3}$$

Let $g_k^{(0)}$, $k = 0, 1, 2, \ldots$ denote the moments of the *linear* contractive IFS $\{\mathrm{X}, w^{(0)}, p\}$, which can be computed recursively as shown in section 3.1. The moments of the *nonlinear* IFS w in (4.1) will be written as $g_k(\epsilon)$. The perturbative approach now consists of assuming the formal expansions

$$g_k(\epsilon) = g_k^{(0)} + g_k^{(1)}\epsilon + g_k^{(2)}\epsilon^2 + \ldots, \quad k = 1, 2, \ldots. \tag{4.4}$$

As always, the normalization condition $g_0(\epsilon) = g_0^{(0)} = 1$ is assumed. No assumptions are made concerning the convergence of the expansion in (4.4). For the moment, it can be treated as a formal power series which provides an asymptotic expansion to $g_k(\epsilon)$. Some statements regarding actual convergence of the series will be made below.

The invariance relation in eq. (2.4) is again employed, with $f(x) = x^n$, to give

$$g_n(\epsilon) = \sum_{i=1}^{N} p_i \int_{\mathrm{A}} \left[s_i x + a_i + \epsilon w_i^{(1)} \right]^n \mathrm{d}\mu, \quad n = 1, 2, 3, \ldots. \tag{4.5}$$

It will be convenient to expand the expression in square brackets in terms of ϵ:

$$\sum_{j=0}^{\infty} g_n^{(j)}\epsilon^j = \sum_{j=0}^{\infty} \epsilon^j \binom{n}{j} \left(\sum_{i=1}^{N} p_i \int_{A} (s_i x + a_i)^{n-j} \left(c_{i,2}x^2 + \ldots + c_{i,n_i}x^{n_i}\right)^j \mathrm{d}\mu \right). \tag{4.6}$$

One now proceeds in a straightforward fashion: the polynomials on the right are expanded and integrated termwise to produce a linear combination of moments $g_k(\epsilon)$,

$$\sum_{j=0}^{\infty} g_n^{(j)}\epsilon^j = \sum_{j=0}^{\infty} \epsilon^j \sum_{k=0}^{n\times n_i} B_{jk} g_k(\epsilon), \tag{4.7}$$

where the B_{jk} are constants involving the IFS parameters s_i, a_i, p_i as well as the polynomial coefficients c_{ik}. The moments on the right are now expanded in powers of ϵ according to eq. (4.4) and like powers of ϵ^j on both sides equated. The appearance of the extra ϵ on the right side of eq. (4.5) permits a recursive computation of the series coefficients $g_n^{(j)}$, as shown below. (The only algebraic complication lies in the expansion of the term $(w_i^{(1)})^j$ but the net result is always a linear combination of moments.) For example, in the case $n = 1$,

$$g_1(\epsilon) = \sum_{i=1}^{N} p_i s_i g_1(\epsilon) + \sum_{i=1}^{N} p_i a_i + \epsilon \sum_{i=1}^{N} p_i \left[c_{i,2} g_2(\epsilon) + \ldots + c_{i,n_i} g_{n_i}(\epsilon) \right]. \tag{4.8}$$

The zeroth-, first- and second-order terms in this equation yield:

$$g_1^{(0)}\left(1 - \sum_{i=1}^{N} p_i s_i\right) = \sum_{i=1}^{N} p_i a_i, \tag{4.9a}$$

$$g_1^{(1)}\left(1 - \sum_{i=1}^{N} p_i s_i\right) = \sum_{i=1}^{N} p_i \left(c_{i,2} g_2^{(0)} + \ldots + c_{i,n_i} g_{n_i}^{(0)}\right), \tag{4.9b}$$

$$g_1^{(2)}\left(1 - \sum_{i=1}^{N} p_i s_i\right) = \sum_{i=1}^{N} p_i \left(c_{i,2} g_2^{(1)} + \ldots + c_{i,n_i} g_{n_i}^{(1)}\right). \tag{4.9c}$$

Eq. (4.9a) is simply the result for the linear "unperturbed" IFS, cf. eq. (3.4). Eq. (4.9b) shows how a knowledge of the $g_k^{(0)}$, $k = 2, \ldots, n_i$ permits the computation of $g_1^{(1)}$. This procedure is now continued for the computation of first-order corrections $g_k^{(1)}$: it is easy to see that to compute $g_2^{(1)}$, we need to know $g_k^{(0)}$, $k = 2, \ldots, 2n_i$, etc. Eq. (4.9c) shows that a knowledge of first-order corrections is necessary to compute second-order coefficients.

The net result of this perturbation procedure may be summarized as follows: First arrange the perturbation coefficients $g_k^{(l)}$ in an array as shown in table 4. The kth row of this array is composed of the perturbation series coefficients of $g_k(\epsilon)$. From normalization, all elements in the first row are zero, except $g_0^{(0)} = 1$. The elements of the first column $g_k^{(0)}$ are simply the moments of the linear IFS, which are recursively computable. The array is then computed in a columnwise manner from the top, noting that in general, the computation of $g_k^{(l)}$ requires a knowledge of $g_k^{(i)}$, $i = 1, \ldots, l-1$ as well as at least $g_{k-1}^{(i)}$, $i = 1, \ldots, l + n_{\max}$, where $n_{\max} = \max_i(n_i)$. The effects of the degree of the maximum nonlinear term $n_{\max}$ are now seen: only a triangular part of the array may be calculated, since a knowledge of the coefficients $g_1^{(l)}$ for $l = 1, \ldots, L$ requires a knowledge of $g_k^{(0)}$, $k = 1, \ldots, Ln_{\max}$.

Table 4
Array of series coefficients $g_k^{(n)}$ for the moments $g_k(\epsilon)$ of a perturbed IFS. The kth row corresponds to the expansion of $g_k(\epsilon)$. The $n = 0$ column consists of the moments $g_k^{(0)}$ of the linear unperturbed IFS.

	$n=0$	$n=1$	$n=2$	
$k=0$	1	0	0	...
$k=1$	$g_1^{(0)}$	$g_1^{(1)}$	$g_1^{(2)}$	...
$k=2$	$g_2^{(0)}$	$g_2^{(1)}$	$g_2^{(2)}$	...
	$\vdots$	$\vdots$	$\vdots$	

Without loss of generality, we illustrate this method for the special case of quadratic IFS maps (in order to minimize the complications with algebra):

$$w_i(x) = s_i x + a_i + \epsilon c_i x^2, \quad i = 1, 2, \dots, N. \tag{4.10}$$

Expanding eq. (4.6), the general recursion relation for the series coefficients becomes

$$\begin{aligned} g_n^{(l)}\left(1 - \sum_{i=1}^{N} p_i s_i^n\right) &= \sum_{k=0}^{n-1} \binom{n}{k} g_k^{(l)} \left(\sum_{k=1}^{N} p_i s_i^k a_i^{n-k} \right) + n \sum_{k=0}^{n-1} \binom{n-1}{k} g_{k+2}^{(l-1)} c_i \left(\sum_{i=1}^{N} p_i s_i^k a_i^{n-1-k} \right) \\ &\quad + \tfrac{1}{2} n(n-1) \sum_{k=0}^{n-2} \binom{n-2}{k} g_{k+4}^{(l-2)} c_i^2 \left(\sum_{i=1}^{N} p_i s_i^k a_i^{n-2-k} \right) + \dots + g_{2n}^{(l-n)}, \end{aligned} \tag{4.11}$$

where "negative orders" vanish, i.e. $g_k^{(m)} = 0$ for $m < 0$. The triangular nature of the computation, which proceeds columnwise, follows from these relations.

Before proceeding with some examples, let us mention that the procedure outlined above is quite similar to the recursive method obtained when traditional quantum mechanical Rayleigh–Schrödinger perturbation theory is combined with the hypervirial and Hellmann–Feynman theorems to produce a "perturbation theory without wavefunctions". A discussion of the quantum mechanical method, with reference to the array of series coefficients and its recursive computation, appears in ref. [16].

Sample calculations

Example 1. The following problem will illustrate the method as well as provide some insight into the convergence properties of the expansions. Consider the quadratic IFS

$$w_1(x) = \tfrac{1}{2}x + \epsilon x^2, \quad w_2(x) = \tfrac{1}{2}x + \tfrac{1}{2} + \epsilon x^2, \quad p_1 = p_2 = \tfrac{1}{2}. \tag{4.12}$$

The unperturbed case, with moments $g_n^{(0)} = 1/(n+1)$, was discussed in section 3.1. Using (4.11), the expansions for the first two moments, to fourth order, are given by

$$\begin{aligned} g_1(\epsilon) &= \tfrac{1}{2} + \tfrac{2}{3}\epsilon + \tfrac{4}{3}\epsilon^2 + \tfrac{128}{35}\epsilon^3 + \tfrac{379154}{33075}\epsilon^4 + \dots, \\ g_2(\epsilon) &= \tfrac{1}{3} + \tfrac{2}{3}\epsilon + \tfrac{64}{35}\epsilon^2 + \tfrac{189577}{33075}\epsilon^3 + \tfrac{429088564}{21531825}\epsilon^4 + \dots. \end{aligned} \tag{4.13}$$

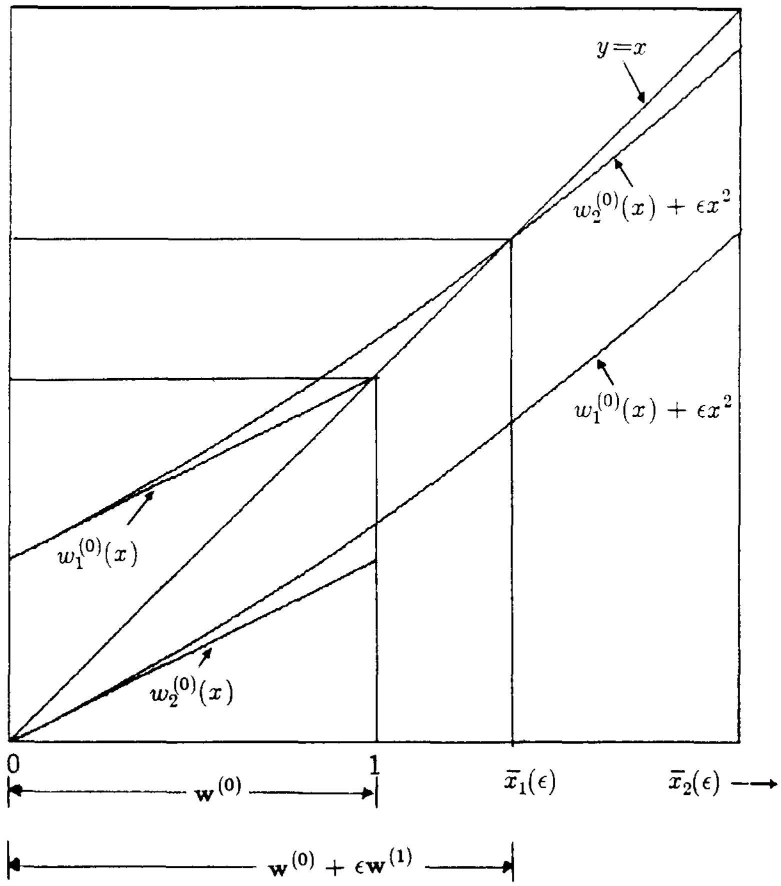


Fig. 3. Plot of IFS maps of eq. (4.12), both for unperturbed case ($\epsilon = 0$), and perturbed, nonlinear case ($\epsilon = 0.1$). For the latter, the second fixed point, $\bar{x}_2(\epsilon)$, of $w_2(x)$ lies to the right of the interval shown.

Some numerical "ratio tests" were then performed to estimate the radii of convergence of these expansions. The series for $g_k(\epsilon)$ were computed to large order (ϵ^{50}) for $k = 1,2,3$. The sequences $g_k^{(l)}/g_k^{(l+1)}$ were then extrapolated to the limit $l \to \infty$ to estimate the radii of convergence R. (The extrapolations were performed using the Thiele–Padé continued fraction method [17, p. 105].) In all cases, an estimate of $R = \frac{1}{8}$ was obtained.

In order to understand this result, consider the unperturbed and perturbed IFS, as shown in fig. 3. The interval $X = [0, \bar{x}_1]$ defining the IFS is determined by the lesser fixed point $\bar{x}_1$ of $w_2(x)$, given by

$$\bar{x}_1(\epsilon) = (1/4\epsilon)\left[1 - (1 - 8\epsilon)^{1/2}\right] = 1 + 2\epsilon + 8\epsilon^2 + 40\epsilon^3 + \mathscr{O}(\epsilon^4). \tag{4.14}$$

As ϵ increases from 0, the two fixed points $\bar{x}_{1,2}$ approach each other and merge at $\epsilon = \frac{1}{8}$, for which the map $w_2(x)$ is still contractive. For $\epsilon > \frac{1}{8}$, the fixed points are complex. (Also note that an IFS is well defined for $-\frac{1}{2} \leq \epsilon \leq 0$.)

Example 2. We now return to the two-map quadratic IFS of the previous section, eq. (3.12), in order to compare the results afforded by both missing moment and perturbative methods.

Table 5
Partial sums of perturbation expansions for $g_1(\epsilon)$ and $g_2(\epsilon)$ in eqs. (4.12) and (4.13) for the two nonlinear IFS problems.

	$\frac{1}{2}x^2, \frac{1}{2}x^2 + \frac{1}{2}$		$\frac{1}{3}x^2, \frac{1}{3}x^2 + \frac{2}{3}$	
n	$S_n(g_1(\epsilon))$	$S_n(g_3(\epsilon))$	$S_n(g_1(\epsilon))$	$S_n(g_3(\epsilon))$
1	0.3125000000	0.0859375000	0.4074074074	0.1975308642
2	0.3281250000	0.0976562500	0.4238683128	0.2194787380
3	0.3359375000	0.1044921875	0.4311842707	0.2308591170
4	0.3398437500	0.1085205078	0.4344358075	0.2368202679
5	0.3420410156	0.1110839844	0.4360615760	0.2401922321
10	0.3451421261	0.1153960824	0.4378578411	0.2445815008
15	0.3455695377	0.1161504805	0.4379922740	0.2450008659
20	0.3456620093	0.1163349283	0.4380082217	0.2450572021
25	0.3456877270	0.1163904775	0.4380106631	0.2450665514
30	0.3456961953	0.1164098500	0.4380111066	0.2450683512
35	0.3456993517	0.1164173943	0.4380111979	0.2450687386
40	0.3457006462	0.1164205981	0.4380112186	0.2450688295
45	0.3457012194	0.1164220574	0.4380112237	0.2450688525
50	0.3457014896	0.1164227616	0.4380112250	0.2450688586

Case 1. $s = \frac{1}{2}$, a perturbation of the linear IFS $w_1^{(0)}(x) = 0$, $w_2^{(0)}(x) = \frac{1}{2}$, with $c_1 = c_2 = 1$, and $\epsilon = \frac{1}{2}$. The attractor $A = \{0, 1/2\}$, with invariant measure $\mu = \frac{1}{2}(\delta_0 + \delta_{1/2})$. Hence the moments of this measure are given by $g_0^{(0)} = 1$, $g_n^{(0)} = (\frac{1}{2})^{n+1}$. Using eq. (4.11), the expansions for the first three moments, to fifth order, are given by

$$\begin{aligned} g_1(\epsilon) &= \tfrac{1}{4} + \tfrac{1}{8}\epsilon + \tfrac{1}{16}\epsilon^2 + \tfrac{1}{16}\epsilon^3 + \tfrac{1}{16}\epsilon^4 + \tfrac{9}{128}\epsilon^5 + \dots, \\ g_2(\epsilon) &= \tfrac{1}{8} + \tfrac{1}{8}\epsilon + \tfrac{1}{16}\epsilon^2 + \tfrac{1}{16}\epsilon^3 + \tfrac{9}{128}\epsilon^4 + \tfrac{21}{256}\epsilon^5 + \dots, \\ g_3(\epsilon) &= \tfrac{1}{16} + \tfrac{3}{64}\epsilon + \tfrac{3}{64}\epsilon^2 + \tfrac{7}{128}\epsilon^3 + \tfrac{33}{512}\epsilon^4 + \tfrac{21}{256}\epsilon^5 + \dots. \end{aligned} \tag{4.15}$$

The expansions for $g_1(\epsilon)$ and $g_3(\epsilon)$ have been computed to large order (ϵ^{50}), and some partial sums are shown in table 5. In both cases, the partial sums are observed to converge to limits consistent with the numerical results in table 3 obtained from both (i) the missing moment method and (ii) the property in (3.15).

In all cases, the numerical ratio test outlined above yielded $R = \frac{1}{2}$. This is consistent with the fact that $\epsilon = \frac{1}{2}$ is a crossing point for the two fixed points of $w_2(x) = \frac{1}{2} + \epsilon x^2$. It is the maximum value for which $w_2(x)$ is contractive on the interval $[0, \bar{x}_1(\epsilon)]$, where $\bar{x}_1(\epsilon)$ is the lesser fixed point of $w_2(x)$.

Case 2. $s = \frac{1}{3}$, a perturbation of the linear IFS $w_1^{(0)}(x) = 0$, $w_2^{(0)}(x) = \frac{2}{3}$, with $c_1 = c_2 = 1$, and $\epsilon = \frac{1}{3}$. The attractor $A = \{0, 2/3\}$, with invariant measure $\mu = \frac{1}{2}(\delta_0 + \delta_{2/3})$. Hence the moments of this measure are given by $g_0^{(0)} = 1$, $g_n^{(0)} = \frac{1}{2}(2/3)^n$. Expansions for the first three moments, to fifth order, are given by

$$\begin{aligned} g_1(\epsilon) &= \tfrac{1}{3} + \tfrac{2}{9}\epsilon + \tfrac{4}{27}\epsilon^2 + \tfrac{16}{81}\epsilon^3 + \tfrac{64}{243}\epsilon^4 + \tfrac{96}{243}\epsilon^5 + \dots, \\ g_2(\epsilon) &= \tfrac{2}{9} + \tfrac{4}{27}\epsilon + \tfrac{16}{243}\epsilon^2 + \tfrac{64}{243}\epsilon^3 + \tfrac{352}{729}\epsilon^4 + \tfrac{448}{729}\epsilon^5 + \dots, \\ g_3(\epsilon) &= \tfrac{4}{27} + \tfrac{4}{27}\epsilon + \tfrac{16}{81}\epsilon^2 + \tfrac{224}{729}\epsilon^3 + \tfrac{1280}{2187}\epsilon^4 + \tfrac{1792}{2187}\epsilon^5 + \dots. \end{aligned} \tag{4.16}$$

The partial sums for the expansions of $g_1(\epsilon)$ and $g_3(\epsilon)$ are also shown in table 5. Convergence to limits consistent with the numerical results of table 3 is also observed.

Numerical extrapolations yield an estimate of $R = \frac{3}{8}$ for the radii of convergence of the above expansions. This is the value of ϵ for which a crossing of fixed points occurs: equivalently, it is the maximum real value of ϵ for which the map $w_2(x) = \epsilon x^2 + \frac{2}{3}$ is contractive on the interval $[0, \bar{x}_1(\epsilon)]$, where $\bar{x}_1(\epsilon)$ is the lesser fixed point of $w_2(x)$.

The applicability of the perturbation method has been shown in the problems studied above. On the basis of the numerical results, we conjecture that R, the radius of convergence of the perturbation expansions, is the distance from $\epsilon = 0$ to the nearest "singularity", that is, the smallest absolute value of ϵ beyond which the perturbed IFS ceases to be contractive, or at which a crossing of fixed points occurs. Of course, the quadratic maps studied here are still rather simple in form. It would be interesting to study more complicated polynomial cases where the crossing of fixed points is less trivial. A deeper question concerns the analyticity properties of the moment functions $g_k(\epsilon)$, as well as the invariant measure μ itself. An ultimate goal is to apply such ideas to the study of some more standard nonlinear dynamical systems on $\mathbb{R}$.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada, both in the form of an Operating Grant (ERV) as well as an Undergraduate Research Assistantship (DW), which are gratefully acknowledged.

References

[1] J. Hutchinson, Fractals and self-similarity, Indiana Univ. J. Math. 30 (1981) 713–747.
[2] M.F. Barnsley and S. Demko, Iterated function systems and the global construction of fractals, Proc. R. Soc. London A 399 (1985) 243–275.
[3] M.F. Barnsley, Fractals Everywhere (Academic Press, New York, 1988).
[4] J. Elton, An ergodic theorem for iterated maps, Ergod. Th. Dynam. Sys. 7 (1987) 481–488.
[5] P. Diaconis and M. Shahshahani, Products of random matrices and computer image generation, in: Random Matrices and Their Applications, Contemp. Math. 50 (Am. Math. Soc., Providence, RI, 1986).
[6] J. Elton and Z. Yan, Approximation of measures by Markov processes and homogeneous affine iterated function systems, Constr. Approx. 5 (1989) 69–87.
[7] C.R. Handy and G. Mantica, Inverse problems in fractal construction: moment method solution, Physica D43 (1990) 17–36.
[8] R.B. Ash, Measure, Integration and Functional Analysis (Academic Press, New York, 1972).
[9] N.I. Akhiezer, The Classical Moment Problem (Hafner, New York, 1965).
[10] E.R. Vrscay and C.J. Roehrig, Iterated function systems and the inverse problem of fractal construction using moments, in: Computers and Mathematics, eds. E. Kaltofen and S.M. Watt (Springer, Berlin, 1989) pp. 250–259. [There is a misprint in eq. (3.10) of this reference: the two moments $g_{i_1 i_2}$ and $g_{j_1 j_2}$ should be replaced by the single moment $g_{i_1+j_1, i_2+j_2}$.]
[11] E.R. Vrscay, Iterated function systems: theory, applications and the inverse problem, in: NATO Advanced Study Institute Lectures, 28th Session, Fractal Geometry and Analysis, Université de Montréal, July 3–21, 1989, to appear.
[12] C.R. Handy and D. Bessis, Rapidly convergent upper and lower bounds for the Schrödinger equation ground state energy, Phys. Rev. Lett. 55 (1985) 931.
[13] B.W. Char, K.O. Geddes, G.H. Gonnet, M.B. Monagan and S.M. Watt, MAPLE Reference Manual, 5th Ed. (Watcom, Waterloo, ON, Canada, 1988).
[14] M.F. Barnsley, S.G. Demko, J. Elton and J.S. Geronimo, Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities, Ann. Inst. H. Poincaré 24 (1988) 367–394.
[15] D. Bessis and S. Demko, Stable recovery of fractal measures by polynomial sampling, Physica D 47 (1991) 427–438.
[16] E.R. Vrscay, Nonlinear "self-interaction" hamiltonians of the form $H^{(0)} + \lambda\langle r^p \rangle r^q$ and their Rayleigh–Schrödinger perturbation expansions, J. Math. Phys. 29 (1988) 901–911.
[17] G.A. Baker, Jr., Essentials of Padé Approximants (Academic Press, New York, 1975).