A FRACTAL VALUED RANDOM ITERATION ALGORITHM AND FRACTAL HIERARCHY

MICHAEL BARNSLEY* and JOHN HUTCHINSON†

Department of Mathematics
Australian National University, Canberra, Australia
*Mbarnsley@aol.com
†john.hutchinson@anu.edu.au

ÖRJAN STENFLO
Department of Mathematics, Stockholm University
SE-10691 Stockholm, Sweden
stenflo@math.su.se

Received May 26, 2004
Accepted October 8, 2004

Abstract

We describe new families of random fractals, referred to as “V-variable”, which are intermediate between the notions of deterministic and of standard random fractals. The parameter V describes the degree of “variability”: at each magnification level any V-variable fractals has at most V key “forms” or “shapes”. V-variable random fractals have the surprising property that they can be computed using a forward process. More precisely, a version of the usual Random Iteration Algorithm, operating on sets (or measures) rather than points, can be used to sample each family. To present this theory, we review relevant results on fractals (and fractal measures), both deterministic and random. Then our new results are obtained by constructing an iterated function system (a super IFS) from a collection of standard IFSs together with a corresponding set of probabilities. The attractor of the super IFS is called a superfractal; it is a collection of V-variable random fractals (sets or measures) together with an associated probability distribution on this collection. When the underlying space is for example $\mathbb{R}^2$, and the transformations are computationally straightforward (such as affine transformations), the superfractal can be sampled by means of the algorithm, which is highly efficient in terms of
1. INTRODUCTION AND NOTATION

1.1. Fractals and Random Fractals

A theory of deterministic fractal sets and measures, using a “backward” algorithm, was developed in Hutchinson.\(^1\) A different approach using a “forward” algorithm was developed in Barnsley and Demko.\(^2\)

Falckner,\(^3\) Graf\(^4\) and Mauldin and Williams\(^5\) randomized each step in the backward construction algorithm to obtain random fractal sets. Arbeiter\(^6\) introduced and studied random fractal measures; see also Olsen.\(^7\) Hutchinson and Richter\(^8\) introduced new probabilistic techniques which allowed one to consider more general classes of random fractals. For further material see Zähle,\(^9\) Patzschke and Zähle,\(^1\) and the references in all of these.

This paper begins with a review of material on deterministic and random fractals generated by IFSs, and then introduces the class of \(V\)-variable fractals which in a sense provides a link between deterministic and “standard” random fractals.

Deterministic fractal sets and measures are defined as the attractors of certain iterated function systems (IFSs), as reviewed in Sec. 2. Approximations in practical situations quite easily can be computed using the associated random iteration algorithm. Random fractals are typically harder to compute because one has to first calculate lots of fine random detail at low levels, then one level at a time, build up the higher levels.

In this paper we restrict the class of random fractals to ones that we call random \(V\)-variable fractals. Superfractals are sets of \(V\)-variable fractals. They can be defined using a new type of IFS, in fact a “super” IFS made of a finite number \(N\) of IFSs, and there is available a novel random iteration algorithm; each iteration produces new sets, lying increasingly close to \(V\)-variable fractals belonging to the superfractal, and moving ergodically around the superfractal.

Superfractals appear to be a new class of geometrical object, their elements lying somewhere between fractals generated by IFSs with finitely many maps, which correspond to \(V = N = 1\), and realizations of the most generic class of random fractals, where the local structure around each of two distinct points are independent, corresponding to \(V = \infty\). They seem to allow geometric modeling of some natural objects, examples including realistic-looking leaves, clouds and textures; and good approximations can be computed fast in elementary computer graphics examples. They are fascinating to watch, one after another, on a computer screen, diverse, yet ordered enough to suggest coherent natural phenomena and potential applications.

Areas of potential applications include computer graphics and rapid simulation of trajectories of stochastic processes. The forward algorithm also enables rapid computation of good approximations to random (including “fully” random) processes, where previously there was no available efficient algorithm.

1.2. An Example

Here we give an illustration of an application of the theory in this paper. By means of this example we introduce informally \(V\)-variable fractals and superfraactals. We also explain why we think these objects are of special interest and deserve attention.

We start with two pairs of contractive affine transformations, \(\{f_1, f_2\}\) and \(\{f_3, f_4\}\), where \(f_n: \square \rightarrow \square\) with \(\square = [0, 1] \times [0, 1] \subset \mathbb{R}^2\). We use two pairs of screens, where each screen corresponds to a copy of \(\square\) and represents for example a computer monitor. We designate one pair of screens to be the Input Screens, denoted by \((\square_1, \square_2)\). The other pair of screens is designated to be the Output Screens, denoted by \((\square_3, \square_4)\).

Initialize by placing an image on each of the Input Screens, as illustrated in Fig. 2, and clearing both of the Output Screens. We construct an image on each of the two Output Screens as follows.

(i) Pick randomly one of the pairs of functions \(\{f_1, f_2\}\) or \(\{f_3, f_4\}\), say \(\{f_1, f_2\}\). Apply \(f_1^m\) to one of the images on \(\square_1\) or \(\square_2\), selected randomly, to make an image on \(\square_3\). Then apply \(f_2^n\) to one of the images on \(\square_1\) or \(\square_2\), also selected randomly, and overlay the resulting image \(I\) on the image now...
already on $\square$. (For example, if black-and-white images are used, simply take the union of the black region of $I$ with the black region on $\square$, and put the result back onto $\square$.)

(ii) Again pick randomly one of the pairs of functions $(f_1^0, f_1^0)$ or $(f_2^0, f_2^0)$, say $(f_1^0, f_2^0)$. Apply $f_1^0$ to one of the images on $\square_1$, or $\square_2$, selected randomly, to make an image on $\square_2$. Also apply $f_2^0$ to one of the images on $\square_1$, or $\square_2$, also selected randomly, and overlay the resulting image on the image now already on $\square$.

(iii) Switch Input and Output, clear the new Output Screens, and repeat steps (i), and (ii).

(iv) Repeat step (iii) many times, to allow the system to settle into its “stationary state”.

What kinds of images do we see on the successive pairs of screens, and what are they like in the “stationary state”? What does the theory developed in this paper tell us about such situations?

As a specific example, let us choose

$$f_1(x, y) = \left(\frac{1}{2}, \frac{3}{8} y + \frac{5}{16}, \frac{1}{2} x + \frac{3}{8} y + \frac{3}{16}\right).$$

$$f_2(x, y) = \left(\frac{1}{2} x + \frac{3}{8} y + \frac{1}{16}, \frac{1}{2} y + \frac{3}{8} y + \frac{3}{16}\right).$$

We describe how these transformations act on the triangle $ABC$ in the diamond $ABCD$, where $A = (\frac{1}{2}, \frac{1}{2}), B = (\frac{1}{2}, \frac{1}{2}), C = (\frac{1}{2}, \frac{1}{2})$, and $D = (\frac{1}{2}, \frac{1}{2})$. Let $B_1 = (\frac{3}{8}, \frac{1}{2}), B_2 = (\frac{9}{16}, \frac{1}{2}), B_3 = (\frac{3}{8}, \frac{1}{2})$, and $B_4 = (\frac{3}{8}, \frac{1}{2})$. See Fig. 1. Then we have

$$f_1(A) = A, \quad f_1(B) = B_1, \quad f_1(C) = B;$$

$$f_2(A) = B, \quad f_2(B) = B_1, \quad f_2(C) = C;$$

$$f_2(A) = A, \quad f_2(B) = B_3, \quad f_2(C) = D;$$

$$f_2(A) = D, \quad f_2(B) = B_4, \quad f_2(C) = C.$$

In Fig. 2, we show an initial pair of images, two jumping fish, one on each of the two screens $\square_1$ and $\square_2$. In Figs. 3 to 9, we show the start of the sequence of pairs of images obtained in a particular trial, for the first seven iterations. Then in Figs. 10 to 12, we show three successive pairs of computed screens, obtained after more than 20 iterations. These latter images are typical of those obtained after 20 or more iterations, very diverse, but always representing continuous “random” paths in $\mathbb{R}^2$; they correspond to the “stationary state”, at the resolution of the images. More precisely, with probability one the empirically obtained distribution on such images over a long experimental run corresponds to the stationary state distribution.

Notice how the two images in Fig. 11 consist of the union of shrunken copies of the images in Fig. 10, while the curves in Fig. 12 are made from two shrunken copies of the curves in Fig. 11.

This example illustrates some typical features of the theory in this paper. (i) New images are generated, one per iteration per screen. (ii) After sufficient iterations for the system to have settled into its “stationary state”, each image looks like a finite resolution rendering of a fractal set that typically changes from one iteration to the next; each fractal belongs to the same family, in the present case a family of continuous curves. (iii) In fact, it follows from the theory that the pictures in this example correspond to curves with this property; for any $\epsilon > 0$ the curve is the union of “little” curves, ones such that the distance apart of any two points is no more than $\epsilon$, each of which is an affine transformation of one of at most two continuous closed paths in $\mathbb{R}^2$. (iv) We will show that the successive images, or rather the abstract objects they represent, eventually all lie arbitrarily close to an object
called a superfractal. The superfractal is the attractor of a superIFS which induces a natural invariant probability measure on the superfractal. The images produced by the algorithm are distributed according to this measure. (v) The images produced in the "stationary state" are independent of the starting images. For example, if the initial images in the example had been of a dot or a line instead of a fish, and the same sequence of random choices had been made, then the images produced in Figs. 10 to 12 would have been the same at the printed resolution.

One similarly obtains $V$-variable fractals and their properties using $V$, rather than two, screens and otherwise proceeding similarly. In (iii) each of the sets of diameter at most $\epsilon$ is an affine
transformation of at most $V$ sets in $\mathbb{R}^2$, where these sets again depend upon $\epsilon$ and the particular image.

This example and the features just mentioned suggest that superfractals are of interest because they provide a natural mathematical bridge between deterministic and random fractals, and because they may lead to practical applications in digital imaging, special effects, computer graphics, as well as in the many other areas where fractal geometric modeling is applied.

1.3. The Structure of This Paper

The main contents of this paper, while conceptually not very difficult, involves potentially elaborate notation because we deal with iterated function systems (IFSs) made of IFSs, and probability measures on spaces of probability measures. So a material part of our effort has been towards a simplified notation. Thus, below, we set out some notation and conventions that we use throughout.
Fig. 8 The two images after six iterations.

Fig. 9 The two images after seven iterations.

Fig. 10 Images on the two screens $\square_1$ and $\square_2$ after a certain number $L > 20$ of iterations. Such pictures are typical of the “stationary state” at the printed resolution.

The core machinery that we use is basic IFS theory, as described in Hutchinson and Barnsley and Demko. So in Sec. 2, we review relevant parts of this theory, using notation and organization that extends to and simplifies later material. To keep the structural ideas clear, we restrict attention to IFSs with strictly contractive transformations and constant probabilities. Of particular relevance to this paper, we explain what is meant by the random iteration algorithm. We illustrate the theorems with simple applications to two-dimensional computer graphics, both to help with understanding and to draw attention to some issues related to discretization that apply a fortiori in computations of $V$-variable fractals.

We begin Sec. 3 with the definition of a superIFS, namely an IFS made of IFSs. We then introduce associated trees, in particular labeled trees, the space of code trees $\Omega$, and construction trees; then we review standard random
In Sec. 4, we study a special class of code trees, called \( V \)-variable trees, where \( V \) is an integer. What are these trees like? At each level they have at most \( V \) distinct subtrees! In fact these trees are described with the aid of an IFS \( \{ \Omega^V, \eta^a, P^a, a \in A \} \) where \( A \) is a finite index set, \( P^a \)'s are probabilities, and each \( \eta^a \) is a contraction mapping from \( \Omega^V \) to itself. The IFS enables one to put a measure attractor on the set of \( V \)-variable trees, such that they can be sampled by means of the random iteration algorithm. We describe the mappings \( \eta^a \) and compositions of them using certain finite doubly labeled trees. This, in turn, enables us to establish the convergence, as \( V \to \infty \), of the probability measure on the set of \( V \)-variable trees, associated with the IFS and the random iteration algorithm, to a corresponding natural probability distribution on the space \( \Omega \).

In Sec. 5, the discussion of trees in Sec. 4 is recapitulated twice over: the same basic IFS theory is applied in two successively more elaborate settings, yielding the formal concepts of \( V \)-variable fractals and superfractals. More specifically, in Sec. 5.1, the superIFS is used to define an IFS of functions that map \( V \)-tuples of compact sets into \( V \)-tuples of compact sets; the attractor of this IFS is a set of \( V \)-tuples of compact sets; these compact sets are named \( V \)-variable fractals and the set of these \( V \)-variable fractals is named a superfractal. We show that these \( V \)-variable fractals can be sampled by means of the random iteration algorithm, adapted to the present setting; that they are distributed according to a certain stationary measure on the superfractal; and that this measure converges to a corresponding measure on the set of “fully” random fractals as \( V \to \infty \), in an appropriate metric. We also provide a continuous mapping from the set of \( V \)-variable trees to the set of \( V \)-variable fractals, and characterize the \( V \)-variable fractals in terms of a property that we name “\( V \)-variability”. Section 5.2 follows the same lines as in Sec. 5.1, except that here the superIFS is used to define an IFS that maps \( V \)-tuples of measures to \( V \)-tuples of measures; this leads to the definition and properties of \( V \)-variable fractal measures. In Sec. 5.3, we describe how to compute the fractal dimensions of \( V \)-variable fractals in certain cases.
and compare them, in a case involving Sierpinski triangles, with the fractal dimensions of deterministic fractals, “fully” random fractals, and “homogeneous” random fractals that correspond to \( V = 1 \) and are a special case of a type of random fractal investigated by Barnsley and Demko.\(^{12-15}\)

In Sec. 6, we describe some potential applications of the theory including new types of space-filling curves for digital imaging, geometric modeling and texture rendering in digital content creation, and random fractal interpolation for computer aided design systems. In Sec. 7, we discuss generalizations and extensions of the theory, areas of ongoing research, and connections to the work of others.

1.4. Some Notation

We use notation and terminology consistent with Barnsley and Demko.\(^2\)

Throughout we reserve the symbols \( M, N \) and \( V \) for positive integers. We will use the variables \( m \in \{1, 2, \ldots, M\} \) and \( v \in \{1, 2, \ldots, V\} \).

Throughout we use an underlying metric space \( (X, d_X) \) which is assumed to be compact unless otherwise stated. We write \( \mathcal{X}^{\mathbb{V}} \) to denote the compact metric space

\[
\mathcal{X} = \mathcal{X} \times \mathcal{X} \times \cdots \times \mathcal{X},
\]

with metric

\[
d(x, y) = \max \{d(x_v, y_v) : v = 1, 2, \ldots, V\},
\]

where \( x = (x_1, x_2, \ldots, x_V) \) and \( y = (y_1, y_2, \ldots, y_V) \).

In some applications, to computer graphics, for example, \( (X, d_X) \) is a bounded region in \( \mathbb{R}^2 \) with the Euclidean metric, in which case we will usually be concerned with affine or projective maps.

Let \( S = S(X) \) denote the set of all subsets of \( X \), and let \( C \subseteq S \). We extend the definition of a function \( f : X \to T \) to \( S \to \mathcal{F} \) by

\[
f(C) = \{ f(x) : x \in C\}.
\]

Let \( \mathcal{H} = \mathcal{H}(X) \) denote the set of non-empty compact subsets of \( X \). Then if \( f : X \to \mathcal{H} \) is continuous we have \( f : H \to \mathcal{H} \). We use \( d_H \) to denote the Hausdorff metric on \( H \) implied by the metric \( d_X \) on \( X \). This is defined as follows. Let \( A \) and \( B \) be two subsets in \( H \), define the distance from \( A \) to \( B \) to be

\[
D(A, B) = \max \{\min \{d_X(x, y) : y \in B\} : x \in A\},
\]

and define the Hausdorff metric by

\[
d_H(A, B) = \max \{D(A, B), D(B, A)\}.
\]

Then \( (\mathcal{H}, d_H) \) is a compact metric space. We will write \( (\mathcal{H}^V, d_{H^V}) \) to denote the \( V \)-dimensional product space constructed from \( (\mathcal{H}, d_H) \) just as \( (\mathcal{X}^V, d_X^V) \) is constructed from \( (X, d_X) \). When we refer to continuous, Lipschitz, or strictly contractive functions acting on \( \mathcal{H}^V \) we assume that the underlying metric is \( d_H^V \).

We will in a number of places start from a function acting on a space, and extend its definition to make it act on other spaces, while leaving the symbol unchanged as above.

Let \( \mathcal{B} = \mathcal{B}(X) \) denote the set of Borel subsets of \( X \). Let \( \mathbb{P} = \mathbb{P}(X) \) denote the set of normalized Borel measures on \( X \). In some applications to computer imaging one chooses \( X = [0, 1] \times [0, 1] \subset \mathbb{R}^2 \) and identifies a black and white image with a member of \( \mathbb{H}(X) \). Greyscale images are identified with members of \( \mathbb{P}(X) \). Probability measures on images are identified with \( \mathbb{P}(\mathbb{H}(X)) \) or \( \mathbb{P}(\mathbb{P}(X)) \).

Let \( d_{\mathbb{P}(X)} \) denote the Monge Kantorovitch metric on \( \mathbb{P}(X) \). This is defined as follows. Let \( \mu \) and \( \nu \) be any pair of measures in \( \mathbb{P} \). Then

\[
d_{\mathbb{P}}(\mu, \nu) = \sup \left\{ \int_X f d\mu - \int_X f d\nu : f : X \to [0, \infty], \int_X f d\mu = \int_X f d\nu \right\}.
\]

Then \( (\mathbb{P}, d_{\mathbb{P}}) \) is a compact metric space. The distance function \( d_{\mathbb{P}} \) metrizes the topology of weak convergence of probability measures on \( X \).\(^{16}\) We define the push-forward map \( f : \mathbb{P}(X) \to \mathbb{P}(X) \) by

\[
f(\mu) = \mu \circ f^{-1} \quad \forall \mu \in \mathbb{P}(X).
\]

Again here we have extended the domain of action of the function \( f : X \to X \).

We will use such spaces as \( \mathbb{P}(\mathbb{H}^V) \) and \( \mathbb{P}(\mathbb{P}(X)^V) \) (or \( \mathbb{H}(\mathbb{H}^V) \) and \( \mathbb{H}(\mathbb{P}^V) \) depending on the context). These spaces may at first seem somewhat Baroque, but as we shall see, they are very natural. In each case we assume that the metric of a space is deduced from the space from which it is built, as above, down to the metric on the lowest space \( X \), and often we drop the subscript on the metric without ambiguity. So, for example, we will write

\[
d(A, B) = d_H(\mathbb{P}(X)^V)(A, B) \quad \forall A, B \in \mathbb{P}(\mathbb{P}(X)^V).
\]

We also use the following common symbols:

\[
N = \{1, 2, \ldots\}, \quad Z = \{\ldots, -2, -1, 0, 1, 2, \ldots\}, \quad \text{and} \quad Z^+ = \{0, 1, 2, \ldots\}.
\]
When $S$ is a set, $|S|$ denotes the number of elements of $S$.

2. ITERATED FUNCTION SYSTEMS

2.1. Definitions and Basic Results

In this section we review relevant information about IFSs. To clarify the essential ideas we consider the case where all mappings are contractive, but indicate in Sec. 5 how these ideas can be generalized. The machinery and ideas introduced here are applied repeatedly later on in more elaborate settings.

Let

$$F = \{X; f_1, f_2, \ldots, f_M; p_1, p_2, \ldots, p_M\} \quad (2.1)$$

denote an IFS with probabilities. The functions $f_m: X \to X$ are contraction mappings with fixed Lipschitz constant $0 \leq l < 1$; that is

$$d(f_m(x), f_m(y)) \leq l \cdot d(x, y) \quad \forall x, y \in X, \quad \forall m \in \{1, 2, \ldots, M\}.$$\hspace{1cm} (2.2)

The $p_m$'s are probabilities, with

$$\sum_{m=1}^M p_m = 1, \quad p_m \geq 0 \quad \forall m.$$\hspace{1cm} (2.3)

We define mappings $F: \mathbb{H}(X) \to \mathbb{H}(X)$ and $F: \mathbb{P}(X) \to \mathbb{P}(X)$ by

$$F(K) = \bigcup_{m=1}^M f_m(K) \quad \forall K \in \mathbb{H},$$

and

$$F(\mu) = \sum_{m=1}^M p_m f_m(\mu) \quad \forall \mu \in \mathbb{P}.$$

In the latter case note that the weighted sum of probability measures is again a probability measure.

**Theorem 1.** The mappings $F: \mathbb{H}(X) \to \mathbb{H}(X)$ and $F: \mathbb{P}(X) \to \mathbb{P}(X)$ are both contractions with factor $0 \leq l < 1$. That is,

$$d(F(K), F(L)) \leq l \cdot d(K, L) \quad \forall K, L \in \mathbb{H}(X),$$

and

$$d(F(\mu), F(\nu)) \leq l \cdot d(\mu, \nu) \quad \forall \mu, \nu \in \mathbb{P}(X).$$

As a consequence, there exists a unique non-empty compact set $A \in \mathbb{H}(X)$ such that

$$F(A) = A,$$

and a unique measure $\mu \in \mathbb{P}(X)$ such that

$$F(\mu) = \mu.$$\hspace{1cm} (2.4)

The support of $\mu$ is contained in or equal to $A$, with equality when all of the probabilities $p_m$ are strictly positive.

**Definition 1.** The set $A$ in Theorem 1 is called the set attractor of the IFS $F$, and the measure $\mu$ is called the measure attractor of $F$.

We will use the term attractor of an IFS to mean either the set attractor or the measure attractor. We will also refer informally to the set attractor of an IFS as a fractal and to its measure attractor as a fractal measure, and to either as a fractal. Furthermore, we say that the set attractor of an IFS is a deterministic fractal. This is in distinction to random fractals, and in particular to $V$-variable random fractals which are the main goal of this paper.

There are two main types of algorithms for the practical computation of attractors of IFS that we term deterministic algorithms and random iteration algorithms, also known as backward and forward algorithms. These terms should not be confused with the type of fractal that is computed by means of the algorithm. Both deterministic and random iteration algorithms may be used to compute deterministic fractals, and as we discuss later, a similar remark applies to our $V$-variable fractals.

Deterministic algorithms are based on the following:

**Corollary 1.** Let $A_0 \in \mathbb{H}(X)$, or $\mu_0 \in \mathbb{P}(X)$, and define recursively

$$A_k = F(A_{k-1}), \quad \text{or} \quad \mu_k = F(\mu_{k-1}), \quad \forall k \in \mathbb{N},$$

respectively, then

$$\lim_{k \to \infty} A_k = A, \quad \text{or} \quad \lim_{k \to \infty} \mu_k = \mu, \quad (2.5)$$

respectively. The rate of convergence is geometrical; for example,

$$d(A_k, A) \leq l^k \cdot d(A_0, A) \quad \forall k \in \mathbb{N}.$$\hspace{1cm} (2.6)

In practical applications to two-dimensional computer graphics, the transformations and the spaces upon which they act must be discretized. The precise behavior of computed sequences of approximations to an attractor of an IFS depends on the details of the implementation and is generally quite complicated; for example, the discrete IFS may have multiple attractors, see Peruggia.

Chap. 4. The following example gives the flavor of such applications.

Fractal Valued Random Iteration Algorithm and Fractal Hierarchy
Example 1. In Fig. 13, we illustrate a practical deterministic algorithm, based on the first formula in Eq. (2.2) starting from a simple IFS on the unit square \( \square \subset \mathbb{R}^2 \). The IFS is \( F = \{ \square, f_1, f_2, f_3 \mathbb{R}^{0.36, 0.28, 0.36} \} \) where the transformations are defined in Eqs. (1.1) to (1.3). The successive images, from left to right, from top to bottom, represent \( A_0, A_2, A_5, A_7, A_20 \) and \( A_21 \). In the last two images, the sequence appears to have converged at the printed resolution to representations of the set attractor. Note however that the initial image is partitioned into two subsets corresponding to the colors red and green. Each successive computed image is made of pixels belonging to a discrete model for \( \square \) and consists of red and green pixels. Each pixel corresponds to a set of points in \( \mathbb{R}^2 \). But for the purposes of computation only one point corresponding to each pixel is used. When both a point in a red pixel and a point in a green pixel belonging to say \( A_n \), are mapped under \( F \) to points in the same pixel in \( A_{n+1} \), a choice has to be made about which color, red or green, to make the new pixel of \( A_{n+1} \). Here we have chosen to make the new pixel of \( A_{n+1} \) the same color as that of the pixel containing the last point in \( A_n \), encountered in the course of running the computer program, to be mapped to the new pixel. The result is that, although the sequence of pictures converge to the set attractor of the IFS, the colors themselves do not settle down, as illustrated in Fig. 15.

We call this “the texture effect”, and comment on it in Example 3. In printed versions of the figures representing \( A_20 \) and \( A_21 \), the red and green pixels are somewhat blended.

The following theorem is the mathematical justification and description of the random iteration algorithm. It follows from Birkhoff’s ergodic theorem and our assumption of contractive maps. A more general version of it is proved in Elton.

**Theorem 2.** Specify a starting point \( x_1 \in \mathcal{X} \). Define a random orbit of the IFS to be \( \{ x_l \}_{l=1}^{\infty} \) where \( x_{l+1} = f_m(x_l) \) with probability \( p_m \). Then for almost all random orbits \( \{ x_l \}_{l=1}^{\infty} \) we have:

\[
\mu(B) = \lim_{l \to \infty} \frac{|B \cap \{x_1, x_2, \ldots, x_l\}|}{l} \tag{2.3}
\]

for all \( B \in \mathcal{B}(\mathcal{X}) \) such that \( \mu(\partial B) = 0 \), where \( \partial B \) denotes the boundary of \( B \).

**Remark 1.** This is equivalent by standard arguments to the following: for any \( x_1 \in \mathcal{X} \) and almost all random orbits the sequence of point measures \( \{ \delta x_1 + \delta x_2 + \cdots + \delta x_l \} \) converges in the weak sense to \( \mu \), see, for example, Billingsley, pp. 11 and 12. (Weak convergence of probability measures is the same as convergence in the Monge Kantorovitch metric, see Dudley, pp. 310 and 311.)

The random iteration algorithm can be applied to the computation of two-dimensional computer graphics. It has benefits compared to deterministic iteration of low memory requirements, high accuracy — as the iterated point can be kept at much higher precision than the resolution of the
Fractal Valued Random Iteration Algorithm and Fractal Hierarchy 121

The following theorem expresses the ergodicity of the IFS \( F \). The proof depends centrally on the uniqueness of the measure attractor. A variant of this theorem, weaker in the constraints on the IFS but stronger in the conditions on the set \( B \), and stated in the language of stochastic processes, is given by Elton.\(^{19}\) We prefer the present version for its simple statement and direct measure theoretic proof.

**Theorem 3.** Suppose that \( \mu \) is the unique measure attractor for the IFS \( F \). Suppose \( B \in \mathcal{B}(\mathbb{X}) \) is such that \( f_m(B) \subseteq B \quad \forall m \in \{1, 2, \ldots, M\} \). Then \( \mu(B) = 0 \) or 1.

**Proof.** Let us define the measure \( \mu|B \) (\( \mu \) restricted by \( B \)) by \( (\mu|B)(E) = \mu(B \cap E) \). The main point of the proof is to show that \( \mu|B \) is invariant under the IFS \( F \). (A similar result applies to \( \mu|B^C \) where \( B^C \) denotes the complement of \( B \).)

If \( E \subseteq B^C \), for any \( m \), since \( f_m(B) \subseteq B \),

\[
f_m(\mu|B)(E) = \mu(B \cap f_m^{-1}(E)) = \mu(B) = 0.
\]

Moreover,

\[
\mu(B) = f_m(\mu|B)(X) = f_m(\mu|B)(B). \tag{2.4}
\]

It follows that

\[
\mu(B) = \sum_{m=1}^{M} p_m f_m(\mu|B)(B) = \sum_{m=1}^{M} p_m f_m(\mu|B^C)(B)
\]

\[
= \mu(B) + \sum_{m=1}^{M} p_m f_m(\mu|B^C)(B) \quad \text{[from (2.4)].}
\]

Hence

\[
\sum_{m=1}^{M} p_m f_m(\mu|B^C)(B) = 0. \tag{2.5}
\]

Hence for any measurable set \( E \subseteq X \)

\[
(\mu|B)(E) = \mu(B \cap E) = \sum_{m=1}^{M} p_m f_m(\mu|B \cap E)
\]

\[
= \sum_{m=1}^{M} p_m f_m(\mu|B)(B \cap E) + \sum_{m=1}^{M} p_m f_m(\mu|B^C)(B \cap E)
\]

\[
= \sum_{m=1}^{M} p_m f_m(\mu|B)(E) + 0 \quad \text{[using (2.5)].}
\]
Thus $\mu(B)$ is invariant and so is either the zero measure or for some constant $c \geq 1$ we have $c\mu(B) = \mu$ (by uniqueness) $= \mu(B) + \mu(B^C)$. This implies $\mu(B^C) = 0$ and in particular $\mu(B^C) = 0$ and $\mu(B) = 1$.

**Example 3.** Figure 15 shows close-ups on the two images at the bottom left in Fig. 13, see Example 1. At each iteration it is observed that the pattern of red and green pixels changes with each iteration, and the distribution of red and green pixels changes with each iteration. A similar texture effect is often observed in other implementations and in applications of $V$-variable fractals to computer graphics. Theorem 3 provides a simple model explanation for this effect as follows. Assume that the red pixels and the green pixels both correspond to sets of points of positive measure, both invariant under $F$. Then we have a contradiction to the corollary above. So neither the red nor the green set can be invariant under $F$. Hence, either one of the sets disappears — which occurs in some other examples — or the pixels must jump around. A similar argument applied to powers of $F$ shows that the way the pixels jump around cannot be periodic, and hence must be "random". (A more careful explanation involves numerical and statistical analysis of the specific computation.)

### 2.2. Fractal Dimensions

In the literature there are many different definitions of a theoretical quantity called the "fractal dimension" of a subset of $\mathbb{X}$. A mathematically convenient definition of the fractal dimension of a set $S \subset \mathbb{X}$ is the Hausdorff dimension. This is always well-defined. Its numerical value often but not always coincides with the values provided by other definitions, when they apply.

Fractal dimensions are useful for a number of reasons. They can be used to compare and classify sets and measures and they have some natural invariance properties. For example, the Hausdorff dimension of a set $S$ is invariant under any bi-Lipschitz transformation; that is, if $f : \mathbb{X} \to \mathbb{X}$ is such that there are constants $c_1$ and $c_2$ in $(0, \infty)$ with $c_1 \cdot d(x, y) \leq d(f(x), f(y)) \leq c_2 \cdot d(x, y) \forall x, y \in \mathbb{X}$ then the Hausdorff dimension of $S$ is the same as that of $f(S)$. Fractal dimensions are useful in fractal image modeling; for example, empirical fractal dimensions of the boundaries of clouds can be used as constraints in computer graphics programs for simulating clouds. Also, as we will see below, the specific value of the Hausdorff dimension of the set attractor $A$ of an IFS can yield the probabilities for most efficient computation of $A$ using the random iteration algorithm. For these same reasons fractal dimensions are an important concept for $V$-variable fractals and superfractals.

The following two definitions are discussed in Falconer, pp. 25 et seq.

**Definition 2.** Let $S \subset \mathbb{X}$, $\delta > 0$, and $0 \leq s < \infty$. Let

$$H_s^\delta(S) = \inf \left\{ \sum_{i=1}^\infty \left| U_i \right|^s \left| \left\{ U_i \right\} \text{ is a } \delta \text{-cover of } S \right. \right\},$$

where $\left| U_i \right|^s$ denotes the $s$th power of the diameter of the set $U_i$, and where a $\delta$-cover of $S$ is a covering of $S$ by subsets of $\mathbb{X}$ of diameter less than $\delta$. Then the $s$-dimensional Hausdorff measure of the set $S$ is defined to be

$$H^s(S) = \lim_{\delta \to 0} H^\delta_s(S).$$

The $s$-dimensional Hausdorff measure is a Borel measure but is not normally even $\sigma$-finite.
Definition 3. The Hausdorff dimension of the set \( S \subset X \) is defined to be
\[
\dim_H S = \inf \{ s | H^s(S) = 0 \}.
\]

The following quantity is often called the fractal dimension of the set \( S \). It can be approximated in practical applications, by estimating the slope of the graph of the logarithm of the number of “boxes” of side length \( \delta \) that intersect \( S \), versus the logarithm of \( \delta \).

Definition 4. The box-counting dimension of the set \( S \subset X \) is defined to be
\[
\dim_B S = \lim_{\delta \to 0} \frac{\log N_\delta(S)}{\log(1/\delta)},
\]
if and only if this limit exists, where \( N_\delta(S) \) is the smallest number of sets of diameter \( \delta > 0 \) that can cover \( S \).

In order to provide a precise calculation of box-counting and Hausdorff dimension of the attractor of an IFS we need the following condition.

Definition 5. The IFS \( F \) is said to obey the open set condition if there exists a non-empty open set \( O \) such that
\[
F(O) = \bigcup_{m=1}^M f_m(O) \subset O,
\]
and
\[
f_m(O) \cap f_\ell(O) = \emptyset \quad \text{if} \ m \neq \ell.
\]

The following theorem provides the Hausdorff dimension of the attractor of an IFS in some special cases.

Theorem 4. Let the IFS \( F \) consist of similitudes, that is \( f_m(x) = s_m O_m x + t_m \), where \( O_m \in \mathbb{R}^K \), \( s_m \in (0,1) \), and \( t_m \in \mathbb{R}^K \). Also let \( F \) obey the open set condition, and let \( A \) denote the set attractor of \( F \). Then
\[
\dim_H A = \dim_B A = D
\]
where \( D \) is the unique solution of
\[
\sum_{m=1}^M s_m^D = 1. \tag{2.6}
\]
Moreover,
\[
0 < H^D(A) < \infty.
\]

Proof. This theorem, in essence, was first proved by Moran in 1946.\textsuperscript{22} A full proof is given in Falconer,\textsuperscript{21} p. 118.

A good choice for the probabilities, which ensures that the points obtained from the random iteration algorithm are distributed uniformly around the set attractor in case the open set condition applies, is \( p_m = s_m^D \). Note that the choice of \( D \) in Eq. 2.6 is the unique value which makes \( (p_1,p_2,\ldots,p_M) \) into a probability vector.

2.3. Code Space

A good way of looking at an IFS as in (2.1) is in terms of the associated code space \( \Sigma = \{1,2,\ldots,M\}^\infty \). Members of \( \Sigma \) are infinite sequences from the alphabet \( \{1,2,\ldots,M\} \) and indexed by \( N \). We equip \( \Sigma \) with the metric \( d_\Sigma \) defined for \( \omega \neq \xi \) by
\[
d_\Sigma(\omega,\xi) = \frac{1}{M^k}
\]
where \( k \) is the index of the first symbol at which \( \omega \) and \( \xi \) differ. Then \( (\Sigma,d_\Sigma) \) is a compact metric space.

Theorem 5. Let \( A \) denote the set attractor of the IFS \( F \). Then there exists a continuous onto mapping \( F : \Sigma \to A \), defined for all \( \sigma_1 \sigma_2 \sigma_3 \ldots \in \Sigma \) by
\[
F(\sigma_1 \sigma_2 \sigma_3 \ldots) = \lim_{k \to \infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \cdots \circ f_{\sigma_k}(x).
\]
The limit is independent of \( x \in X \) and the convergence is uniform in \( x \).

Proof. This result is contained in Hutchinson\textsuperscript{1} Theorem 3.1(3).

Definition 6. The point \( \sigma_1 \sigma_2 \sigma_3 \ldots \in \Sigma \) is called an address of the point \( F(\sigma_1 \sigma_2 \sigma_3 \ldots) \in A \).

Note that \( F : \Sigma \to A \) is not in general one-to-one.

The following theorem characterizes the measure attractor of the IFS \( F \) as the push-forward, under \( F : \Sigma \to A \), of an elementary measure \( \rho \in P(\Sigma) \), the measure attractor of a fundamental IFS on \( \Sigma \).

Theorem 6. For each \( m \in \{1,2,\ldots,M\} \) define the shift operator \( s_m : \Sigma \to \Sigma \) by
\[
s_m(\sigma_1 \sigma_2 \sigma_3 \ldots) = m \sigma_1 \sigma_2 \sigma_3 \ldots
\]
\( \forall \sigma_1 \sigma_2 \sigma_3 \ldots \in \Sigma \). Then \( s_m \) is a contraction mapping with contractivity factor \( \frac{1}{M} \). Consequently \( S := \{\Sigma; s_1,s_2,\ldots,s_M; p_1,p_2,\ldots,p_M\} \)
is an IFS. Its set attractor is \( \Sigma \). Its measure attractor is the unique measure \( \pi \in \mathbb{P}(\Sigma) \) such that
\[
\pi\{\omega_1\omega_2\omega_3 \ldots \in \Sigma | \omega_1 = \sigma_1, \omega_2 = \sigma_2, \ldots, \omega_k = \sigma_k \} = p_{\sigma_1}p_{\sigma_2} \cdots p_{\sigma_k}
\]
\( \forall k \in \mathbb{N}, \forall \sigma_1, \sigma_2, \ldots, \sigma_k \in \{1, 2, \ldots, M\} \).

If \( \mu \) is the measure attractor of the IFS \( F \), with \( F: \Sigma \to A \) defined as in Theorem 5, then
\[ \mu = F(\pi). \]

**Proof.** This result is Hutchinson\(^1\) Theorem 4.4(3) and (4).

We call \( S \) the shift IFS on code space. It has been well studied in the context of information theory and dynamical systems, see for example Billingsley\(^2\) and results can often be lifted to the IFS \( F \). For example, when the IFS is non-overlapping, the entropy (see Barnsley et al.\(^3\) for the definition) of the stationary stochastic process associated with \( F \) is the same as that associated with the corresponding shift IFS, namely:
\[ \sum p_\sigma \log p_\sigma. \]

---

### 3. TREES OF ITERATED FUNCTION SYSTEMS AND RANDOM FRACTALS

#### 3.1. SuperIFSs

Let \( \{X, d_X\} \) be a compact metric space, and let \( M \) and \( N \) be positive integers. For \( n \in \{1, 2, \ldots, N\} \) let \( F^n \) denote the IFS
\[ F^n = \{X; f^n_1, f^n_2, \ldots, f^n_M; p^n_1, p^n_2, \ldots, p^n_M\} \]
where each \( f^n_\sigma: X \to X \) is a Lipschitz function with Lipschitz constant \( 0 \leq l < 1 \) and the \( p^n_\sigma \)'s are probabilities with
\[ \sum_{m=1}^M p^n_\sigma = 1, \quad p^n_\sigma \geq 0 \quad \forall m, n. \]

Let
\[ F = \{X; F^1, F^2, \ldots, F^N; P_1, P_2, \ldots, P_N\}, \]
where the \( P_\sigma \)'s are probabilities with
\[ \sum_{n=1}^N P_\sigma = 1, P_\sigma \geq 0 \quad \forall n \in \{1, 2, \ldots, N\}. \]

With \( P_\sigma > 0 \forall n \in \{1, 2, \ldots, N\} \) and \( \sum_{n=1}^N P_\sigma = 1 \).

As we will see in later sections, given any positive integer \( V \) we can use the set of IFSs \( F \) to construct a single IFS acting on \( \mathbb{H}(X)^V \). In such cases we call \( F \) a superIFS. Optionally, we will drop the specific reference to the probabilities.

#### 3.2. Trees

We associate various collections of trees with \( F \) and the parameters \( M \) and \( N \).

Let \( T \) denote the \((M\text{-fold}) \) set of finite sequences from \( \{1, 2, \ldots, M\} \), including the empty sequence \( \emptyset \). Then \( T \) is called a tree and the sequences are called the nodes of the tree. For \( i = i_1i_2 \ldots i_k \in T \) let \( |i| = k \). The number \( k \) is called the level of the node \( \sigma \). The bottom node \( \emptyset \) is at level zero. If \( j = j_1j_2 \ldots j_k \in T \) then \( ij \) is the concatenated sequence \( i_1j_1i_2j_2 \ldots j_k \).

We define a level-\( k \) (\( M\text{-fold} \)) tree, or a tree of height \( k \), \( T_k \) to be the set of nodes of \( T \) of level less than or equal to \( k \).

A labeled tree is a function whose domain is a tree or a level-\( k \) tree. A limb of a tree is either an ordered pair of nodes of the form \( (i, im) \) where \( i \in T \) and \( m \in \{1, 2, \ldots, M\} \), or the pair of nodes \( (\emptyset, \emptyset) \), which is also called the trunk. In representations of labeled trees, as in Figs. 16 and 19, limbs are represented by line segments and we attach the labels either to the nodes where line segments meet or to the line segments themselves, possibly to both the nodes and limbs when a labeled tree is multivalued. For a two-valued labeled tree \( \tau \) we will write
\[ \tau(i) = (\tau(\text{node } i), \tau(\text{limb } i)) \quad \text{for } i \in T, \]
to denote the two components.

A code tree is a labeled tree whose range is \( \{1, 2, \ldots, N\} \). We write
\[ \Omega = \{\tau | \tau: T \to \{1, 2, \ldots, N\}\} \]
for the set of all infinite code trees.

We define the subtree \( \tilde{\tau}: T \to \{1, 2, \ldots, N\} \) of a node \( i = i_1i_2 \ldots i_k \in T \) corresponding to a node \( \tau = \tau_1 \ldots \tau_k \). For \( j \in T \), by
\[ \tilde{\tau}(j) = \tau(ij) \quad \forall j \in T. \]

In this case we say that \( \tilde{\tau} \) is a subtree of \( \tau \) at level \( k \). (One can think of a subtree as a branch of a tree.)

Suppose \( \tau \) and \( \sigma \) are labeled trees with \( |\tau| \leq |\sigma| \), where we allow \( |\tau| = \infty \). We say that \( \sigma \) extends \( \tau \), and \( \tau \) is an initial segment of \( \sigma \), if \( \sigma \) and \( \tau \) agree on their common domain, namely at nodes up to and including those at level \( |\tau| \). We write
\[ \tau < \sigma. \]

If \( \tau \) is a level-\( k \) code tree, the corresponding cylinder set is defined by
\[ [\tau] = [\tau]_\Omega := \{\sigma \in \Omega : \tau < \sigma\}. \]
Fig. 16  Pictorial representation of a level-4 two-fold tree labeled by the sequences corresponding to its nodes. The labels on the fourth level are shown for every second node. The line segments between the nodes and the line segment below the bottom node are referred to as limbs. The bottom limb is also called the trunk.

We define a metric on \( \Omega \) by, for \( \omega \neq \kappa \),
\[
d_\Omega(\omega, \kappa) = \frac{1}{M^k}
\]
if \( k \) is the least integer such that \( \omega(i) \neq \kappa(i) \) for some \( i \in T \) with \( |i| = k \). Then \( (\Omega, d_\Omega) \) is a compact metric space. Furthermore,
\[
diam(\Omega) = 1
\]
and
\[
diam(\tau) = \frac{1}{M^{k+1}} \quad (3.2)
\]
whenever \( \tau \) is a level-\( k \) code tree.

The probabilities \( P_n \) associated with \( F \) in Eq. (3.1) induce a natural probability distribution \( \rho \in P(\Omega) \) on \( \Omega \). It is defined on cylinder sets \([\tau]\) by
\[
\rho([\tau]) = \prod_{1 \leq |i| \leq |\tau|} P_{\tau(i)} \quad (3.3)
\]
That is, the random variables \( \tau(i) \), with nodal values in \( \{1, 2, \ldots, N\} \), are chosen i.i.d. via \( \Pr(\tau(i) = n) = P_n \). Then \( \rho \) is extended in the usual way to the \( \sigma \)-algebra \( B(\Omega) \) generated by the cylinder sets. Thus we are able to speak of the set of random code trees \( \Omega \) with probability distribution \( \rho \), and of selecting trees \( \sigma \in \Omega \) according to \( \rho \).

A construction tree for \( F \) is a code tree wherein the symbols 1, 2, \ldots, and \( N \) are replaced by the respective IFSs \( F^1, F^2, \ldots, \) and \( F^N \). A construction tree consists of nodes and limbs, where each node is labeled by one of the IFSs belonging to \( F \). We will associate the \( M \) limbs that lie above and meet at a node with the constituent functions of the IFS of that node; taken in order.

We use the notation \( F(\Omega) \) for the set of construction trees for \( F \). For \( \sigma \in \Omega \) we write \( F(\sigma) \) to denote the corresponding construction tree. We will use the same notation \( F(\sigma) \) to denote the random fractal set associated with the construction tree \( F(\sigma) \), as described in the next section.

3.3. Random Fractals

In this section we describe the canonical random fractal sets and measures associated with \( F \) in (3.1).

Let \( F \) be given as in (3.1), let \( k \in \mathbb{N} \), and define
\[
F_k : \Omega \times \mathbb{H}(X) \to \mathbb{H}(X),
\]
by
\[
F_k(\sigma)(K) = \bigcup_{\{i \in T : |i| = k\}} f_\sigma^{\tau(i)} \circ f_\sigma^{\tau(i_1)} \circ \cdots \circ f_\sigma^{\tau(i_{k-1})}(K) \quad (3.4)
\]
\( \forall \sigma \in \Omega \) and \( K \in \mathbb{H}(X) \). (The set \( F_k(\sigma)(K) \) is obtained by taking the union of the compositions of the functions occurring on the branches of the construction tree \( F(\sigma) \) starting at the bottom and working up to the \( k \)th level, acting upon \( K \).)
In a similar way, with measures in place of sets, and unions of sets replaced by sums of measures weighted by probabilities, we define
\[
\tilde{F}_k: \Omega \times \mathcal{P}(X) \to \mathcal{P}(X)
\]
by
\[
\tilde{F}_k(\sigma)(c) = \sum_{i \in \mathcal{I}} \left( P_{i_1}^{\sigma(i_1)} \cdot P_{i_2}^{\sigma(i_2)} \cdots P_{i_k}^{\sigma(i_k)} \right) \times \tilde{f}_{i_1}^\sigma \circ \tilde{f}_{i_2}^\sigma \circ \cdots \circ \tilde{f}_{i_k}^\sigma (c)
\]
for all \( \sigma \in \Omega \) and \( c \in \mathcal{P}(X) \). Note that the \( \tilde{F}_k(\sigma)(c) \) all have unit mass because the \( p_n^m \) sum (over \( m \)) to unity.

**Theorem 7.** Let sequences of functions \( \{ F_k \} \) and \( \{ \tilde{F}_k \} \) be defined as above. Then both the limits
\[
F(\sigma) = \lim_{k \to \infty} F_k(\sigma)(K), \quad \text{and} \quad \tilde{F}(\sigma) = \lim_{k \to \infty} \tilde{F}_k(\sigma)(c),
\]
exist, are independent of \( K \) and \( c \), and the convergence (in the Hausdorff and Monge Kantorovitch metrics, respectively) is uniform in \( \sigma \), \( K \) and \( c \). The resulting functions
\[
F: \Omega \to \mathcal{P}(X) \quad \text{and} \quad \tilde{F}: \Omega \to \mathcal{P}(X)
\]
are continuous.

**Proof.** Make repeated use of the fact that, for fixed \( \sigma \in \Omega \), both mappings are compositions of contraction mappings of contractivity \( l \), by Theorem 1.

Let
\[
\mathcal{H} = \{ F(\sigma) \in \mathcal{H}(X) | \sigma \in \Omega \}, \quad \text{and} \quad \mathcal{H} = \{ \tilde{F}(\sigma) \in \mathcal{P}(X) | \sigma \in \Omega \}.
\]
Similarly let
\[
\mathcal{P} = F(\rho) = \rho \circ F^{-1} \in \mathcal{P}(X), \quad \text{and} \quad \mathcal{P} = \tilde{F}(\rho) = \rho \circ \tilde{F}^{-1} \in \mathcal{P}(X).
\]

**Definition 7.** The sets \( \mathcal{H} \) and \( \mathcal{\tilde{H}} \) are called the sets of fractal sets and fractal measures, respectively, associated with \( F \). These random fractal sets and measures are said to be distributed according to \( \mathcal{P} \) and \( \mathcal{\tilde{P}} \), respectively.

Random fractal sets and measures are hard to compute. There does not appear to be a general simple forwards (random iteration) algorithm for practical computation of approximations to them in two dimensions with affine maps, for example.

The reason for this difficulty lies with the inconvenient manner in which the shift operator acts on trees \( \sigma \in \Omega \) relative to the expressions (3.4) and (3.5).

**Definition 8.** Both the set of IFSs \( \{ F^n : n = 1, 2, \ldots, N \} \) and the superIFS \( F \) are said to obey the (uniform) open set condition if there exists a non-empty open set \( O \) such that for each \( n \in \{ 1, 2, \ldots, N \} \)
\[
F^n(O) = \bigcup_{m=1}^{M} f^m_n(O) \subseteq O,
\]
and
\[
\bigcap_{m=1}^{M} f^m_n(O) = \emptyset \quad \forall m, l \in \{ 1, 2, \ldots, M \} \quad \text{with} \quad m \neq l.
\]

For the rest of this section we restrict attention to \( (X, d) \) where \( X \subseteq \mathbb{R}^K \) and \( d \) is the Euclidean metric. The following theorem gives a specific value for the Hausdorff dimension for almost all of the random fractal sets in the case of “non-overlapping” similitudes.\(^{\text{3-5}}\)

**Theorem 8.** Let the set of IFSs \( \{ F^n : n = 1, 2, \ldots, N \} \) consist of similitudes, i.e. \( f^m_n(x) = s^m_n O^m_n x + t^m_n \) where \( O^m_n \) is an orthonormal transformation on \( \mathbb{R}^K \), \( s^m_n \in (0, 1) \), and \( t^m_n \in \mathbb{R}^K \), for all \( n \in \{ 1, 2, \ldots, N \} \) and \( m \in \{ 1, 2, \ldots, M \} \). Also let \( \{ F^n : n = 1, 2, \ldots, N \} \) obey the uniform open set condition. Then for \( \mathcal{P} \) almost all \( A \in \mathcal{P} \)
\[
\dim_H A = \dim_{\mathcal{P}} A = D
\]
where \( D \) is the unique solution of
\[
\sum_{n=1}^{N} P_n \sum_{m=1}^{M} (s^m_n)^D = 1.
\]

**Proof.** This is an application of Falconer’s Theorem 15.2, p. 230. \( \square \)

### 4. CONTRACTION MAPPINGS ON CODE TREES AND THE SPACE \( \Omega_V \)

#### 4.1. Construction and Properties of \( \Omega_V \)

Let \( V \in \mathbb{N} \). This parameter will describe the variability of the trees and fractals that we are going to introduce. Let
\[
\Omega^V = \Omega \times \Omega \times \cdots \times \Omega \quad \text{v times}
\]
We refer to an element of $\Omega^V$ as a grove. In this section we describe a certain IFS on $\Omega^V$, and discuss its set attractor $\Omega_1$: its points are $(V$-tuples of) code trees that we will call $V$-groves. We will find it convenient to label the trunk of each tree in a grove by its component index, from the set $\{1, 2, \ldots, V\}$.

One reason that we are interested in $\Omega$: is that, as we shall see later, the set of trees that occur in its components, called $V$-trees, provides the appropriate code space for a $V$-variable superfractal.

Next we describe mappings from $\Omega^V$ to $\Omega^V$ that comprise the IFS. The mappings are denoted by $\eta^a: \Omega^V \rightarrow \Omega^V$ for $a \in A$ where

$$A := \{(1, 2, \ldots, N) \times \{1, 2, \ldots, V\}^M\}^V.$$

A typical index $a \in A$ will be denoted $a = (a_1, a_2, \ldots, a_1)$

$$a_v = (n_v; v_{v,1}, v_{v,2}, \ldots, v_{v,M})$$

where $n_v \in \{1, 2, \ldots, N\}$ and $v_{v,m} \in \{1, 2, \ldots, V\}$ for $m \in \{1, 2, \ldots, M\}$.

Specifically, algebraically, the mapping $\eta^a$ is defined in Eq. (4.8) below. But it is very useful to represent the indices and the mappings with trees. See Fig. 17. Each map $\eta^a$ in Eq. (4.8) and each index $a \in A$ may be represented by a $V$-tuple of labeled level-1 trees that we call (level-1) function trees. Each function tree has a trunk, a node and $M$ limbs. There is one function tree for each component of the mapping. Its trunk is labeled by the index of the component $v \in \{1, 2, \ldots, V\}$ to which it corresponds. The node of each function tree is labeled by the IFS number $n_v$ (shown circled) of the corresponding component of the mapping. The $n$th limb of the $v$th tree is labeled by the number $v_{v,m} \in \{1, 2, \ldots, V\}$ for $m \in \{1, 2, \ldots, M\}$.

We will use the same notation $\eta^a$ to denote both a $V$-tuple of function trees and the unique mapping $\eta^a: \Omega^V \rightarrow \Omega^V$ to which it bijectively corresponds. We will use the notation $a$ to denote both a $V$-tuple of function trees and the unique index $a \in A$ to which it bijectively corresponds.

Now we can describe the action of $\eta^a$ on a grove $\omega \in \Omega^V$. We illustrate with an example where $V = 3$, $N = 5$ and $M = 2$. In Fig. 18, an arbitrary grove $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega^3$ is represented by a triple of colored tree pictures, one blue, one orange and one magenta, with trunks labeled one, two and three, respectively. The top left of Fig. 18 shows the map $\eta^a$, and the index $a \in A$, where

$$\eta^a(\omega_1, \omega_2, \omega_3) = (\xi_1(\omega_1, \omega_2), \xi_2(\omega_2, \omega_3), \xi_4(\omega_2, \omega_3)), \quad (4.4)$$

and

$$a = ((1, 1, 2), (5, 3, 2), (4, 3, 1)).$$

represented by function trees. The functions $\{\xi_n: n = 1, 2, \ldots, 5\}$ are defined in Eq. (4.7) below. The result of the action of $\eta^a$ on $\omega$ is represented, in the bottom part of Fig. 18, by a grove whose lowest nodes are labeled by the IFS numbers 1, 5 and 4, respectively, and whose subtrees at level zero consist of trees from $\omega$ located according to the limb labels on the function trees. (Limb labels of the top left expression, the function trees of $\eta^a$, are matched to trunk labels in the top right expression, the components of $\omega$.) In general, the result of the action of $\eta^a$ in Fig. 17 on a grove $\omega \in \Omega^V$ (represented by $V$ trees with trunks labeled from 1 to $N$) is obtained by matching the limbs of the function trees to the trunks of the $V$-trees, in a similar manner.

![Fig. 17](image_url)
We are also going to need a set of probabilities \( \{ p^a \}_{a \in A} \), with

\[
\sum_{a \in A} p^a = 1, \quad p^a \geq 0 \quad \forall a \in A.
\]  

These probabilities may be more or less complicated. Some of our results are specifically restricted to the case

\[
p^a = \frac{P_n \cdot P_2 \cdots P_n}{V(MV)},
\]  

which uses only the set of probabilities \( \{ P_1, P_2, \ldots, P_n \} \) belonging to the superIFS (3.1). This case corresponds to labeling all nodes and limbs in a function tree independently with probabilities such that limbs are labeled according to the uniform distribution on \([1, 2, \ldots, V]\), and nodes are labeled \( j \) with probability \( P_j \).

For each \( n \in \{1, 2, \ldots, N\} \) define the \( n \)th shift mapping \( \zeta_n: \Omega^V \rightarrow \Omega^V \)

\[
(\zeta_n(\omega)(i)) = n \quad \text{and} \quad (\zeta_n(\omega)(m)) = \omega_n(i)
\]  

\forall i \in \mathbb{F}, \, m, \, n \in \{1, 2, \ldots, M\},

for \( \omega = (\omega_1, \omega_2, \ldots, \omega_M) \in \Omega^M \). That is, the mapping \( \zeta_n \) creates a code tree with its bottom node labeled \( n \) attached directly to the \( M \) trees \( \omega_1, \omega_2, \ldots, \omega_M \).

**Theorem 9.** For each \( a \in A \) define \( \eta^a: \Omega^V \rightarrow \Omega^V \) by

\[
\eta^a(\omega_1, \omega_2, \ldots, \omega_V) := \left( \zeta_{v_1}(\omega_{v_1,1}, \omega_{v_1,2}, \ldots, \omega_{v_1,M}), \ldots, \zeta_{v_1}(\omega_{v_1,1}, \omega_{v_1,2}, \ldots, \omega_{v_1,M}) \right)
\]  

Then \( \eta^a: \Omega^V \rightarrow \Omega^V \) is a contraction mapping with Lipshitz constant \( \frac{1}{M} \).

**Proof.** Let \( a \in A \). Let \( (\omega_1, \omega_2, \ldots, \omega_V) \) and \( (\tilde{\omega}_1, \tilde{\omega}_2, \ldots, \tilde{\omega}_V) \) be any pair of points in \( \Omega^V \). Then

\[
d_H(\eta^a(\omega_1, \omega_2, \ldots, \omega_V), \eta^a(\tilde{\omega}_1, \tilde{\omega}_2, \ldots, \tilde{\omega}_V)) = \max_{i \in \{1, 2, \ldots, V\}} d_H(\zeta_{v_i}(\omega_{v_i,1}, \omega_{v_i,2}, \ldots, \omega_{v_i,M}), \zeta_{v_i}(\tilde{\omega}_{v_i,1}, \tilde{\omega}_{v_i,2}, \ldots, \tilde{\omega}_{v_i,M}))
\]  

\[
\leq \frac{1}{M} \cdot \max_{i \in \{1, 2, \ldots, V\}} d_H(\omega_{v_i,1}, \omega_{v_i,2}, \ldots, \omega_{v_i,M}), \tilde{\omega}_{v_i,1}, \tilde{\omega}_{v_i,2}, \ldots, \tilde{\omega}_{v_i,M})
\]  

\[
\leq \frac{1}{M} \cdot \max_{i \in \{1, 2, \ldots, V\}} d_H(\omega_v, \tilde{\omega}_v) = \frac{1}{M} \cdot d_{\Omega^V}((\omega_1, \omega_2, \ldots, \omega_V), (\tilde{\omega}_1, \tilde{\omega}_2, \ldots, \tilde{\omega}_V)).
\]

It follows that we can define an IFS \( \Phi \) of strictly contractive maps by

\[
\Phi = \{ \Omega^V: \eta^a, p^a, a \in A \}.
\]

Let the set attractor and the measure attractor of the IFS \( \Phi \) be denoted by \( \Omega_\Phi \) and \( \mu_\Phi \), respectively. Clearly, \( \Omega_\Phi \in \mathbb{H}(\Omega^V) \) while \( \mu_\Phi \in \mathcal{P}(\Omega^V) \). We call \( \Omega_\Phi \) the space of \( V \)-groves. The elements of \( \Omega_\Phi \) are certain \( V \)-tuples of \( M \)-fold code trees on an alphabet of \( N \) symbols, which we characterize in Theorem 11. But we think of them as special groves of special trees, namely \( V \)-trees.

For all \( v \in \{1, 2, \ldots, V\} \), let us define \( \Omega_{\Phi v} \subset \Omega_\Phi \) to be the set of \( v \)th components of groves in \( \Omega_\Phi \). Also let \( \mu_\Phi \in \mathcal{P}(\Omega) \) denote the marginal probability measure defined by

\[
\mu_\Phi(B) := \mu_\Phi(B, \Omega, \Omega, \ldots, \Omega) \quad \forall B \in \mathcal{B}(\Omega).
\]
Theorem 10. For all \( v \in \{1, 2, \ldots, V\} \) we have:
\[
\Omega_{\|v} = \Omega_{\|1} := \{\text{set of all } V\text{-trees}\}.
\] (4.11)
When the probabilities \( \{P^a|a \in A\} \) obey Eq. (4.6), then starting at an initial grove, the random distribution of trees \( \omega \in \Omega \) that occur in the \( v \)th components of groves produced by the random iteration algorithm corresponding to the IFS \( \Phi \), after \( n \) iteration steps, converge weakly to \( \rho_V \) independently of \( v \) almost always, as \( n \to \infty \).

Proof. Let \( \Xi : \Omega^V \to \Omega^V \) denote any map that permutes the coordinates. Then any IFS \( \Phi = \{\Omega^V, \rho^a, P^a|a \in A\} \) is invariant under \( \Xi \), that is \( \Xi \Omega^V = \Omega^V \). It follows that \( \Xi \Omega_{\|v} = \Omega_{\|v} \) and \( \Xi \rho_V = \rho_V \). It follows that Eq. (4.11) holds, and also that, in the obvious notation, for any \((B_1, B_2, \ldots, B_V) \in (\Xi(\Omega))^V\) we have
\[
\rho_V(B_1, B_2, \ldots, B_V) = \rho_V(B_2, B_3, \ldots, B_{V+1}).
\] (4.12)
In particular
\[
\rho_V(B) = \rho_V(B, \Omega, \Omega, \ldots, \Omega)
\]
for all \( B \in (\Xi(\Omega))^V \). Theorem 9 tells us that we can apply the random iteration algorithm (Theorem 2) to the IFS \( \Phi \). This yields sequences of measures, denoted by \( \{
\rho^l_V : l = 1, 2, 3, \ldots\} \), that converge weakly to \( \rho_V \) almost always. In particular \( \rho^l_V(B, \Omega, \Omega, \ldots, \Omega) \) converges to \( \rho_V(B) \) almost always.

Let \( L \) denote a set of fixed length vectors of labeled trees. We will say that \( L \) and its elements have the property of \( V \)-variability, or that \( L \) and its elements are \( V \)-variable, if and only if, for all \( \omega \in \Omega \), the number of distinct subtrees of all components of \( \omega \) at level \( k \) is at most \( V \), for each level \( k \) of the trees.

Theorem 11. Let \( \omega \in \Omega^V \). Then \( \omega \in \Omega_{\|v} \) if and only if \( \omega \) contains at most \( V \) distinct subtrees at any level \( i \) (i.e., \( \omega \) is \( V \)-variable). Also, if \( \sigma \in \Omega \), then \( \sigma \) is a \( V \)-tree if and only if it is \( V \)-variable.

Proof. Let \( S = \{\omega \in \Omega^V|(|\text{number of components of } \omega \text{ at level } k|) \leq V, \forall k \in \mathbb{N}\}. \)

Then \( S \) is closed. Let \( \{s_n \in S\} \) be a sequence converging to \( s^* \). Suppose that \( s^* \notin S \). Then, at some level \( k \in \mathbb{N} \), \( s^* \) has more than \( V \) subtrees. There exists \( l \in \mathbb{N} \) so that each distinct pair of these subtrees of \( s^* \) first disagrees at some level less than \( l \). Now choose \( n \) so large that \( s_n \) agrees with \( s^* \) through level \( (k + l) \) (i.e., \( d_{K^\infty}(s_n, s^*) < \varepsilon_{K^\infty}\)). Then \( s_n \) has more than \( V \) distinct subtrees at level \( k \), a contradiction. So \( s^* \in S \).

Also \( S \) is non-empty: let \( \sigma \in \Omega \) be defined by \( \sigma(i) = 1 \) for all \( i \in T \). Then \((\sigma, \sigma, \ldots, \sigma) \in S \).

Also \( S \) is invariant under the IFS \( \Phi \); clearly any \( s \in S \) can be written \( s = \eta^i(s) \) for some \( i \in A \) and \( \eta \in \Omega \). Also, if \( s \in S \) then \( \eta^i(s) \in S \). So \( S = \bigcup \{\eta^i(S)|a \in A\} \).

Hence, by uniqueness, \( S \) must be the set attractor of \( \Phi \). That is, \( S = \Omega_{\|V} \). This proves the first claim in the theorem.

It now follows that if \( \sigma \in \Omega \) is a \( V \)-tree then it contains at most \( V \) distinct subtrees at level \( k \), for each \( k \in \mathbb{N} \). Conversely, it also follows that if \( \sigma \in \Omega \) has the latter property, then \((\sigma, \sigma, \ldots, \sigma) \in \Omega_{\|V} \), and so \( \sigma \in \Omega_{\|1} \).

Theorem 12. For all
\[
d_{\Xi(\Omega)}(\Omega_{\|1}, \Omega) < \frac{1}{V},
\]
which implies
\[
\lim_{V \to \infty} \Omega_{\|1} = \Omega,
\]
where the convergence is in the Hausdorff metric.

Let the probabilities \( \{P^a|a \in A\} \) obey Eq. (4.6). Then
\[
d_{\Xi(\Omega)}(\rho_V, \rho) \leq \frac{1.4}{V} \left(\frac{M}{V}\right)^{t} \tag{4.13}
\]
which implies
\[
\lim_{V \to \infty} \rho_V = \rho,
\]
where \( \rho \) is the stationary measure on trees introduced in Sec. 3.2, and convergence is in the Monge-Kantorovich metric.

Proof. To prove the first part, let \( M^{k+l} > V \geq M^k \). Let \( r \) be any level-k code tree. Then \( r \) is clearly \( V \)-variable and it can be extended to an infinite \( V \)-variable code tree. It follows that \( |r \cap \Omega_{\|1}| \neq \emptyset \). The collection of such cylinder sets \( |r \) forms a disjoint partition of \( \Omega \) by subsets of diameter \( \frac{1}{M^{k+l}} \), see Eq. (3.2), from which it follows that
\[
d_{\Xi(\Omega)}(\Omega_{\|1}, \Omega) \leq \frac{1}{M^{k+l}} < \frac{1}{V}.
\]
The first part of the theorem follows at once.
For the proof of the second part, we refer to Sec. 4.4.

Let $\Sigma_V = A^\infty$. This is simply the code space corresponding to the IFS $\Phi$ defined in Eq. (4.9). From Theorem 5, there exists a continuous onto mapping $\Phi: \Sigma_V \to \Omega_V$ defined by

$$
\Phi(a_1a_2a_3\ldots) = \lim_{k \to \infty} \eta^a \circ \eta^b \circ \cdots \circ \eta^c(\omega)
$$

for all $a_1a_2a_3\ldots \in \Sigma_V$, for any $\omega$. In the terminology of Sec. 2.3 the sequence $a_1a_2a_3\ldots \in \Sigma_V$ is an address of the $V$-grove $\Phi(a_1a_2a_3\ldots) \in \Omega_V$ and $\Sigma_V$ is the code space for the set of $V$-groves $\Omega_V$. In general $\Phi: \Sigma_V \to \Omega_V$ is not one-to-one, as we will see in Sec. 4.2.

4.2. Compositions of the Mappings $\eta^a$

Compositions of the mappings $\eta^a$: $\Omega^V \to \Omega^V$, $a \in A$, represented by $V$-tuples of level-1 function trees, as in Fig. 17, can be represented by higher level trees that we call level-$k$ function trees.

First we illustrate the ideas, then we formalize. In Fig. 19, we illustrate the idea of composing $V$-tuples of function trees. In this example $V = 3$, $N = 5$ and $M = 2$. The top row shows the level-1 function trees corresponding to $a, b, c \in \mathcal{A}$ given by

$$
a = ((1; 2, 3), (3; 1, 3), (5; 2, 3)), 
b = ((4; 3, 1), (2; 1, 2), (3; 3, 2)),
$$

and

$$
c = ((1; 1, 2), (5; 1, 3), (4; 2, 3)).
$$

The first entry in the second row shows the $3$-tuple of level-$2$ function trees $a \circ b$. The bottom bracketed expression shows the $3$-tuple of level-$2$ function trees $a \circ b \circ c$. Then in Fig. 20, we have represented $\eta^a \circ \eta^b \circ \eta^c(\omega)$ for $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega^3$. The key ideas are (i) a function $\eta$ can be converted into a mapping $\eta^{a,b,c} : \Omega^3 \to \Omega^3$ and (ii)

$$
\eta^{a,b,c} = \eta^a \circ \eta^b \circ \eta^c.
$$

The mapping $\eta^{a,b,c}(\omega)$ is defined to be the $3$-tuple of code trees obtained by attaching the tree $\omega_i$ to each of the top limbs of each level-$3$ function tree in $a \circ b \circ c$ with label $e$ for all $e \in \{1, 2, 3\}$ then dropping all of the labels on the limbs.

Next we formalize. Let $k \in \mathbb{N}$. Define a level-$k$ function tree to be a level-$k$ labeled tree with the nodes of the first $k - 1$ levels labeled from $\{1, 2, \ldots, N\}$ and limbs (i.e. second nodal values) of all $k$ levels labeled from $\{1, 2, \ldots, V\}$. We define a grove of level-$k$ function trees, say $g$, to be a $V$-tuple of level-$k$ function trees, with trunks labeled according to the component number, and we define $G_k$ to be the set of such $g$. Let $\mathcal{G} := \bigcup_{k \in \mathbb{N}} G_k$. We will refer to a component of an element of $\mathcal{G}$ simply as a function tree. For $g \in \mathcal{G}$ we will write $|g| = k$, where $k \in \mathbb{N}$ is the unique number such that $g \in G_k$.

![Fig. 19 Illustrations of compositions of function trees to produce higher level function trees. Here $V = 3$, $N = 5$ and $M = 2$. We compose the level-1 function trees corresponding to $a = ((1; 2, 3), (3; 1, 3), (5; 2, 3))$, $b = ((4; 3, 1), (2; 1, 2), (3; 3, 2))$ and $c = ((1; 1, 2), (5; 1, 3), (4; 2, 3))$. The top row shows the separate level-1 function trees, $a$, $b$ and $c$. The second row shows the level-2 function tree $a \circ b$, and the function tree $c$. The last row shows the level-3 function tree $a \circ b \circ c.$](image-url)
Then, for all $g = (g_1, g_2, \ldots, g_V) \in G$, for all $v \in \{1, 2, \ldots, V\}$,

$$g_v(\text{node } i) \in \{1, 2, \ldots, N\} \quad \forall \ i \in T \text{ with } |i| \leq |g|-1,$$

and

$$g_v(\text{limb } i) \in \{1, 2, \ldots, V\} \quad \forall \ i \in T \text{ with } |i| \leq |g|.$$  

For all $g, h \in G$ we define the composition $g \circ h$ to be the grove of $(|g| + |h|)$-level function trees given by the following expressions.

$$(g \circ h)_v(\text{node } i) = g_v(\text{node } i) \quad \forall i \in T \text{ with } |i| \leq |g| - 1;$$

$$(g \circ h)_v(\text{limb } i) = g_v(\text{limb } i) \quad \forall i \in T \text{ with } |i| \leq |g|;$$

$$(g \circ h)_v(\text{node } ij) = h_{g_v(\text{limb } i)}(\text{node } j) \quad \forall ij \in T \text{ with } |i| = |g|, \ |j| \leq |h| - 1;$$

$$(g \circ h)_v(\text{limb } ij) = h_{g_v(\text{limb } i)}(\text{limb } j) \quad \forall ij \in T \text{ with } |i| = |g|, \ |j| \leq |h|.$$  

For all $g \in G$ we define $\eta^g : \Omega^V \rightarrow \Omega^V$ by

$$(\eta^g(\omega))_v(i) = g_v(\text{node } i) \quad \forall i \in T \text{ with } |i| \leq |g| - 1,$$

and

$$(\eta^g(\omega))_v(ij) = \omega_{g_v(\text{limb } i)}(\text{node } j) \quad \forall ij \in T \text{ with } |i| = |g|, \ |j| \geq 0.$$  

This is consistent with the definition of $\eta^g : \Omega^V \rightarrow \Omega^V$ for $g \in A$, as the following theorem shows. We will write $\eta^g$ to denote both the mapping $\eta^g : \Omega^V \rightarrow \Omega^V$ and the corresponding unique $V$-tuple of level-$k$ code trees for all $g \in G$.

**Theorem 13.** For all $g, h \in G_k$ we have

$$\eta^{g \circ h} = \eta^g \circ \eta^h.$$  

It follows that the operation $\circ$ between ordered pairs of elements of $G$ is associative. In particular, if $(a_1, a_2, \ldots, a_k) \in A^k$ then

$$\eta^{a_1} \circ \eta^{a_2} \circ \cdots \circ \eta^{a_k} = \eta^{a_1 a_2 \cdots a_k}.$$  

**Proof.** This is a straightforward exercise in substitutions and is omitted here.  

We remark that Eqs. (4.15) and (4.14) allow us to work directly with function trees to construct, count and track compositions of mappings $\eta^g \in A$. The space $G$ also provides a convenient setting for contrasting the forwards and backwards algorithms associated with the IFS $\Phi$. For example, by composing function trees in such a way as to build from the bottom up, which corresponds to a backwards algorithm, we find that we can construct a sequence of cylinder set approximations to the first component $\omega_1$ of $\omega \in \Omega_V$ without having to compute approximations to the other components.  

Let $G_k(V) \subseteq G_k$ denote the set of elements of $G_k$ that can be written in the form $a_1 \circ a_2 \circ \cdots \circ a_k$ for some $a = (a_1, a_2, \ldots, a_k) \in A^k$. (We remark that $G_k(V)$ is $V$-variable by a similar argument to the proof of Theorem 11.) Then we are able to estimate the measures of the cylinder sets $[(\eta^{a_1 \circ a_2 \cdots \circ a_k})_V]_{\omega}$ by computing the probabilities of occurrence of function trees $g \in G_k(V)$ such that $\eta^g_1 = (\eta^{a_1 \circ a_2 \cdots \circ a_k})_V$. Built up starting from level-0 trees, with probabilities given by Eq. (4.6), as we do in the Sec. 4.4.

The labeling of limbs in the approximating grove of function trees of level-k of $\Phi(a_1 a_2 \cdots)$ defines the basic $V$-variable dependence structure.
of $\Phi(a_1a_2\ldots)$. We call the code tree of limbs of a function tree the associated dependence tree.

The grove of code trees for $\Phi(a_1a_2\ldots)$ is by definition totally determined by the labels of the nodes. Nevertheless its grove of dependence trees contains all information concerning its $V$-variable structure. The dependence tree is the characterizing skeleton of $V$-variable fractals.

4.3. A Direct Characterization of the Measure $\nu_V$

Let $(F_n)_{n=0}^{\infty} = (F_0,\ldots,F_m)_{m=0}^\infty$ be a sequence of random groves of level-$1$ function trees. Each random function tree can be expressed as $F_m = (N_m^1,L_m^1(1),\ldots,L_m^1(M))$ where the $N_m^1$s and $L_m^1$s correspond to random labelings of nodes and limbs, respectively.

We assume that the function trees $\{F_m\}$ are independent with $Pr(N_m^1 = j) = 1/V$ for any $v,k \in \{1,\ldots,V\}$, $n \in \mathbb{N}$, $j \in \{1,\ldots,M\}$, and $n \in \{1,\ldots,M\}$.

First the family $\{F_0,\ldots,F_m\}$ generates a random dependence tree $K: T \rightarrow \{1,\ldots,V\}$, in the following way. Let $K(0) = 1$. If $K(i_1,\ldots,i_n) = j$, for some $j = 1,\ldots,V$, then we define $K(i_1,\ldots,i_{n+1}) = L_j^{k(n+1)}$.

Given the dependence tree, let $I_i = N_{K(i)}$ if $|i| = n$.

The following theorem gives an alternative definition of $\rho_V$:

**Theorem 14.** $\rho_V((\tau)) = Pr(I_i = \tau(i), \forall |i| \leq |\tau|)$, where $(I_i)_{i \in T}$ is defined as above, and $\tau$ is a finite level code tree.

**Proof.** It is a straightforward exercise to check that $(F_0 o F_1 o \cdots o F_{k-1})$ is a level-$k$ function tree with nodes given by $(I_i)_{i \in \Delta k-1}$, and limbs given by $(K(i))_{i \in \Delta k}$.

Thus

$$Pr(I_i = \tau(i), \forall i \text{ with } |i| \leq k) = Pr((F_0 o F_1 o \cdots o F_{k-1})_{(\text{node } i)}) = \tau(i), \forall i \text{ with } |i| \leq k - 1. \quad (4.16)$$

For $a = (a_1,a_2,\ldots,a_k) \in A^k$, let

$$a^n = a_1^{n_1}a_2^{n_2}\cdots a_k^{n_k}$$

denote the probability of selection of $\eta^n = \eta^{a_1}\eta^{a_2}\cdots\eta^{a_k} = \eta^{a_1^{n_1}a_2^{n_2}\cdots a_k^{n_k}}$.

By the invariance of $\mu_V$

$$\mu_V = \sum_{a \in A^k} p_a\eta^a(\mu_V).$$

Now let $\tau$ be a level-$(k-1)$ code tree. Then

$$\rho_V(\tau) = \mu_V(\tau) = \sum_{a \in A^k} p_a\mu_V((\eta^n)^{-1}(\tau))$$

$$= \sum_{(a \in A^k)_{(\tau(i), \forall i)}} p_a\mu_V((\eta^n)^{-1}(\tau)).$$

From this and (4.16), it follows that $\rho_V((\tau)) = Pr(I_i = \tau(i), \forall i \text{ with } |i| \leq k - 1)$.

4.4. Proof of Theorem 12 Equation (4.13)

**Proof.** We say that a dependence tree is *free up to level $k$*, if at each level $j$, for $1 \leq j \leq k$, the $M^j$ limbs have distinct labels. If $V$ is much bigger than $M$ and $k$ then it is clear that the probability of being free up to level $k$ is close to unity. More precisely, if $F$ is the event that the dependence tree of $(F_0 o F_1 o \cdots o F_{k-1})_{(\text{node } i)}$ is free and $V \geq M^k$, then

$$\rho_V((\tau)) = \mu_V(\tau) = \sum_{a \in A^k} p_a\mu_V(\rho_{a^n})$$

$$= \sum_{a \in A^k} p_a\mu_V((\eta^n)^{-1}(\tau))$$

$$= \sum_{(a \in A^k)_{(\tau(i), \forall i)}} p_a\mu_V((\eta^n)^{-1}(\tau)).$$

$Pr(F) = \prod_{i=1}^{M-1} \left(1 - \frac{1}{V}\right) \prod_{i=1}^{M^k-1} \left(1 - \frac{i}{V}\right)$

$$\geq 1 - \frac{1}{V} \left(\sum_{i=1}^{M-1} + \sum_{i=1}^{M^2-1} + \cdots + \sum_{i=1}^{M^k-1}\right)$$

$$\geq 1 - \frac{1}{2V}(M^2 + M^4 + \cdots + M^2k)$$

$$\geq 1 - \frac{M^2k+1}{2V(M^2-1)} \geq 1 - \frac{2M^{2k}}{3V}.$$
Hence
\[ \nu(\omega) = \Pr(S) = \Pr(F) \Pr(S|F) + \Pr(F^c) \Pr(S|F^c) \leq \Pr(S|F) + \Pr(F^c) \leq \rho(\omega) + \frac{2M^{2k}}{3V}. \]

Similarly,
\[ \nu(\omega) = \Pr(S) \Pr(S|F) \geq \rho(\omega) - \frac{2M^{2k}}{3V}. \]

Hence
\[ |\nu(\omega) - \rho(\omega)| \leq \frac{2M^{2k}}{3V}. \tag{4.17} \]

In order to compute the Monge Kantorovitch distance \( d_{\text{M}}(\nu, \rho) \), suppose \( f : \Omega \rightarrow \mathbb{R} \) is Lipschitz with \( \text{Lip} f \leq 1 \), i.e. \( |f(\omega) - f(\omega')| \leq d_\Omega(\omega, \omega') \forall \omega, \omega' \in \Omega \). Since \( \text{diam}(\Omega) = 1 \), we subtract a constant from \( f \) and so can assume \( |f| \leq \frac{1}{2} \) without changing the value of \( \int \text{Lip} f dp - \int f dp \nu \).

For each level-\( k \) code tree \( \tau \in T_k \) choose some \( \omega_\tau \in [\tau] \subseteq \Omega \). It then follows that
\[
\left| \int_\Omega f dp - \int_\Omega f dp \nu \right| = \sum_{\tau \in T_k} \left( \int_{[\tau]} f dp - \int_{[\tau]} f dp \nu \right) \\
= \sum_{\tau \in T_k} \int_{[\tau]} \left( f - f(\omega_\tau) \right) dp \\
- \sum_{\tau \in T_k} \int_{[\tau]} \left( f - f(\omega_\tau) \right) dp \nu \\
+ \sum_{\tau \in T_k} \left( f(\omega_\tau) \left( \rho(\omega_\tau) - \nu(\omega_\tau) \right) \right) \\
\leq \frac{1}{M^{2k}} \sum_{\tau \in T_k} \rho(\omega_\tau) + \frac{1}{M^{2k}} \sum_{\tau \in T_k} \nu(\omega_\tau) \\
+ \sum_{\tau \in T_k} \frac{M^{2k}}{3V} \\
= \phi(k) := \frac{2}{M^{2k} + \frac{2M^{2k}}{3V} + \frac{1}{3V}}.
\]

since \( \text{diam} [\tau] \leq \frac{1}{M^{2k+1}} \) from Eq. (3.2), \( |f(\omega_\tau)| \leq \frac{1}{2} \).

Lip \( f \leq 1 \), \( \omega_\tau \in [\tau] \), and using Eq. (4.17). Choose \( x \) so that \( \frac{\rho(x)}{M^{2k}} = \frac{1}{M^{2k}} \). This is the value of \( x \) which minimizes \( \left( \frac{\rho(x)}{M^{2k}} + \frac{1}{M^{2k}} \right) \). Choose \( k \) so that \( k \leq x \leq k+1 \).

Then
\[
\phi(k) \leq \frac{2}{M^{2k+1}} + \frac{2M^{2k}}{3V} \leq \frac{2}{M^{2k}} \left( 1 + \frac{1}{3V} \right) \\
\leq \frac{7}{27V} \left( \frac{M}{3V} \right)^k, \quad (M \geq 2).
\]

Hence Eq. (4.13) is true.

5. SUPERFRACTALS

5.1. Contraction Mappings on \( \mathbb{H}^V \) and the Superfractal Set \( \mathcal{S}_{V,1} \).

Definition 9. Let \( V \in \mathbb{N} \), let \( \mathcal{A} \) be the index set introduced in Eq. (4.1), let \( \mathcal{F} \) be given as in Eq. (3.1), and let probabilities \( \{p_a | a \in \mathcal{A}\} \) be given as in Eq. (4.5). Define
\[
f^a : \mathbb{H}^V \rightarrow \mathbb{H}^V
\]
by
\[
f^a(K) = \left( \bigcup_{m=1}^M f_m^a(K_{v,m}), \bigcup_{m=1}^M f_m^a(K_{v,m}) \right) \bigcup_{m=1}^M f_m^a(K_{v,m}) \bigcup_{m=1}^M f_m^a(K_{v,m})
\tag{5.1}
\]
\( \forall K = (K_1, K_2, \ldots, K_V) \in \mathbb{H}^V \), \( \forall a \in \mathcal{A} \).

Let \( \mathcal{F}_V := \{\mathbb{H}^V; f^a, a \in \mathcal{A}\} \). \tag{5.2}

Theorem 15. \( \mathcal{F}_V \) is an IFS with contraction factor \( l \).

Proof. We only need to prove that the mapping \( f^a : \mathbb{H}^V \rightarrow \mathbb{H}^V \) is contractive with contraction factor \( l, \forall a \in \mathcal{A} \). Note that, \( \forall K = (K_1, K_2, \ldots, K_M), \mathcal{L} = (L_1, L_2, \ldots, L_M) \in \mathbb{H}^V \),
\[
d_{\mathcal{L}} \left( \bigcup_{m=1}^M f_m^a(K_m), \bigcup_{m=1}^M f_m^a(L_m) \right) \leq \max_m \left\{ d_{\mathcal{L}}(f_m^a(K_m), f_m^a(L_m)) \right\} \\
\leq \max_m \left\{ \| f_m^a(K_m, L_m) \| \right\} \\
= \| f_m^a(K, L) \| = l \cdot d_{\mathcal{L}}(K, L).
\]
Hence, \( \forall (K_1, K_2, \ldots, K_V), (L_1, L_2, \ldots, L_V) \in H_V, \)
\[
d_{H^V}(f^*(K_1, K_2, \ldots, K_V), f^*(L_1, L_2, \ldots, L_V)) = \max_v \left\{ d_V \left( \bigcup_{m=1}^{M} f^*_m(K_{v,m}), \bigcup_{m=1}^{M} f^*_m(L_{v,m}) \right) \right\}
\]
\[
\leq \max_v \left\{ l \cdot d_v((K_{v,1}, K_{v,2}, \ldots, K_{v,M})), \right\}
\]
\[
\leq l \cdot d_v((K_1, K_2, \ldots, K_V), (L_1, L_2, \ldots, L_V)).
\]

The theory of IFS in Sec. 2.1 applies to the IFS \( F_1 \). It possesses a unique set attractor \( \delta_{V,1} \in E(X) \), and a unique measure attractor \( \Psi_{V,1} \in \mathcal{P}(E(X)) \). The random iteration algorithm corresponding to the IFS \( F_1 \) may be used to approximate sequences of points \((V\)-tuples of compact sets) in \( \delta_{V,1} \) distributed according to the probability measure \( \Psi_{V,1} \).

However, the individual components of these vectors in \( \delta_{V,1} \), certain special subsets of \( \mathcal{X} \), are the objects in which we are interested. Accordingly, for all \( v \in \{1, 2, \ldots, V\} \), let us define \( \delta_{V,v} \subset \mathbb{R} \) to be the set of \( v \)-th components of points in \( \delta_{V,1} \).

**Theorem 16.** For all \( v \in \{1, 2, \ldots, V\} \), we have \( \delta_{V,v} = \delta_{V,1} \).

When the probabilities in the superIFS \( F_V \) are given by (4.6), then starting from any initial \( V \)-tuple of non-empty compact subsets of \( \mathcal{X} \), the random distribution of the sets \( K \in H \) that occur in the \( v \)-th component of vectors produced by the random iteration algorithm after \( n \) initial steps converge weakly to the marginal probability measure:
\[
\Psi_{V,1}(B) = \Psi_{V}(B, H, \ldots, H) \quad \forall B \in \mathcal{B}(H),
\]

independently of \( v \), almost always, as \( n \to \infty \).

**Proof.** The direct way to prove this theorem is to parallel the proof of Theorem 10, using the maps \((f^*_v): H^V \to H^V | v \in A\) in place of the maps \((\tau_v^*): \Omega^V \to \Omega^V | v \in A\).

However, an alternate proof follows with the aid of the map \( F: \Omega \to E(X) \) introduced in Theorem 7. We have put this alternate proof at the end of the proof of Theorem 17.

**Definition 10.** We call \( \delta_{V,1} \) a superfractal set. Points in \( \delta_{V,1} \) are called \( V \)-variable fractal sets.

**Example 4.** See Fig. 21. This example is similar to the one in Sec. 1.2. It shows some of the images produced in a realization of random iteration of a superIFS with \( M = N = V = 2 \). Projective transformations are used in both IFSs, specifically...
One of the goals of this example is to illustrate how closely similar images can be produced, with “random” variations, so the two IFSs are quite similar. Let us refer to images (or, more precisely, the sets of points that they represent) such as the ones at the bottom middle and at the bottom right, as “ti-trees.” Then each transformation maps approximately the unit square \( \square := \{(x, y) | 0 \leq x \leq 1, 0 \leq y \leq 1\} \), in which each ti-tree lies, into itself. Both \( f_1(x, y) \) and \( f_2(x, y) \) map ti-trees to the lower right branches of ti-trees. Both \( f_1(x, y) \) and \( f_2(x, y) \) map ti-trees to a ti-tree minus the lower right branch. The initial image for each component, or “screen,” is illustrated at the top left. It corresponds to an array of pixels of dimensions 400 \( \times \) 400, some of which are red, some green, and the rest white. Upon iteration, images of the red pixels and green pixels are combined as in Example 1. The number of iterations increases from left to right, and from top to bottom, corresponding to the fifth iteration. Both images at the bottom middle and bottom right correspond to more than 30 iterations, and are representative of typical images produced after more than 30 iterations. (We carried out more than 50 iterations.) They represent images selected from the superfractal \( \tilde{\mathcal{H}}_y \) according to the invariant measure \( \tilde{\mathcal{P}}_{2,1} \). Note that it is the support of the red and green pixels that corresponds to an element of \( \tilde{\mathcal{H}}_y \). Note too the “texture effect,” similar to the one described in Example 3.

By Theorem 5, there is a continuous mapping \( \mathcal{F}_\psi : \Sigma_V \to \tilde{\mathcal{H}}_y \) that assigns to each address in the code space \( \Sigma_V = \mathcal{A}^\omega \) a V-tuple of compact sets in \( \tilde{\mathcal{H}}_y \). But this mapping is not helpful for characterizing \( \tilde{\mathcal{H}}_y \) because \( \mathcal{F}_\psi : \Sigma_V \to \tilde{\mathcal{H}}_y \) is not in general one-to-one, for the same reason that \( \Phi : \Sigma_V \to \tilde{\mathcal{H}}_y \) is not one-to-one, as explained in Sec. 4.2. The following result is closer to the point. It tells us in particular that the set of V-trees is a useful code space for V-variable fractals, because the action of the IFS \( \Phi \) on the space of V-tuples of code trees is conjugate to the action of the IFS \( \mathcal{F}_\psi \) acting on V-tuples of compact sets. (We are concerned here with the mappings that provide the correspondences between V-groves and V-trees, on the one hand, and points and probability distributions on \( \tilde{\mathcal{H}}_V \) and \( \tilde{\mathcal{H}}_{V,1} \), on the other.)

**Theorem 17.** Let \( a \in A \) and \( \eta^c : \Omega^V \to \Omega^V \) be defined as in Theorem 9. Let \( \mathcal{F} : \tilde{\mathcal{H}}^+_Y \to \tilde{\mathcal{H}}^+_Y \) be the mapping introduced in Theorem 7. Define \( \tilde{\mathcal{F}} : \Omega^V \to (\tilde{\mathcal{H}}^+_Y)^{(n)} \) by

\[
\tilde{\mathcal{F}}(\omega_1, \omega_2, \ldots, \omega_V) = (\mathcal{F}(\omega_1), \mathcal{F}(\omega_2), \ldots, \mathcal{F}(\omega_V)),
\]

for all \((\omega_1, \omega_2, \ldots, \omega_V) \in \Omega^V \). Then

\[
\tilde{\mathcal{F}}(\eta^c(\omega)) = \mathcal{F}(\mathcal{F}(\omega)) \quad \forall \omega \in A, \omega \in \Omega^V.
\]

Also

\[
\tilde{\mathcal{F}}(\Omega_V) = \tilde{\mathcal{H}}_V \quad \text{and} \quad \tilde{\mathcal{F}}(\Omega_{V,1}) = \tilde{\mathcal{H}}_{V,1},
\]

where \( \Omega_V \) denotes the set of \( t \)th components of members of \( \tilde{\mathcal{H}}_V \). Similarly, when the probabilities in the IFS \( \mathcal{F}_V \) of Eq. (5.2.), are given by Eq. (4.16), we have

\[
\mathcal{F}(\mu^t) = \widetilde{\mathcal{P}}_{V,1} \quad \text{and} \quad \mathcal{F}(\mu_V) = \Psi_{V,1},
\]

where \( \mu_V \) is the marginal probability distribution given by Eq. (4.10).

**Proof.** We begin by establishing the key Eq. (5.3). Note that from Theorem 7, for any \( \kappa \in (\Omega^V) \),

\[
\mathcal{F}(\omega) = (\mathcal{F}(\omega_1), \mathcal{F}(\omega_2), \ldots, \mathcal{F}(\omega_V))
\]

\[
= \left\{ \lim_{k \to \infty} (\mathcal{F}_k(\omega_1)(K)), \lim_{k \to \infty} (\mathcal{F}_k(\omega_2)(K)), \ldots \right\}.
\]

The first component here exemplifies the others; and using Eq. (3.4), it can be written

\[
\lim_{k \to \infty} (\mathcal{F}_k(\omega_1)(K)) = \lim_{k \to \infty} \left\{ \bigcup_{\nu \in \mathcal{F}_V(\kappa)} \mathcal{F}_k^{(\nu)} \circ \mathcal{F}_k^{(\nu)}(K) \right\}
\]

(5.7)
Since the convergence is uniform and all of the functions involved are continuous, we can interchange the limit with function operation at will. Look at

$$f^k(\mathcal{F}(\omega))$$

$$= \lim_{k \to \infty} \{F_k(\omega_1)(K), \lim_{k \to \infty} \{F_k(\omega_2)(K), \ldots, \lim_{k \to \infty} \{F_k(\omega)(K)\}\}$$

By the definition in Theorem 15, Eq. (5.1), we have

$$f^k(\mathcal{F}_k(\omega_1)(K), \mathcal{F}_k(\omega_2)(K), \ldots, \mathcal{F}_k(\omega)(K))$$

$$= \left(\bigcup_{m=1}^{M} f^{m}(\mathcal{F}_k(\omega_1,\omega))(K), \bigcup_{m=1}^{M} f^{m}(\mathcal{F}_k(\omega_2,\omega))(K), \ldots, \bigcup_{m=1}^{M} f^{m}(\mathcal{F}_k(\omega,\omega))(K)\right).$$

By Eq. (3.4), the first component here is

$$\lim_{m \to \infty} \{f^{m}(\mathcal{F}_k(\omega_1,\omega))(K)\}$$

$$= \lim_{m \to \infty} \left(\bigcup_{i \in \{0\} \cup \{1\}} f^{i_{1}}(\omega_{1},\omega_{2},\omega_{3},\ldots) \circ f^{i_{2}}(\omega_{1},\omega_{2},\omega_{3},\ldots) \circ \cdots \circ f^{i_{k}}(\omega_{1},\omega_{2},\omega_{3},\ldots) \circ f^{i_{k+1}}(\omega_{1},\omega_{2},\omega_{3},\ldots)\right)(K)$$

$$= \mathcal{F}_k \ast \{f_1(\omega_1,\omega_2,\omega_3,\ldots)\}(K),$$

where we have used the definition in Eq. (4.7).

Hence

$$f^k(\mathcal{F}(\omega))$$

$$= \lim_{k \to \infty} \{f^{k}(\mathcal{F}(\omega_1)(K), \mathcal{F}_k(\omega_2)(K), \ldots, \mathcal{F}_k(\omega)(K))\}$$

$$= \lim_{k \to \infty} \{f^{k}(\mathcal{F}_k(\omega_1,\omega_2,\omega_3,\ldots))(K), \mathcal{F}_k(\omega_1,\omega_2,\omega_3,\ldots))(K), \ldots, \mathcal{F}_k(\omega_1,\omega_2,\omega_3,\ldots))(K)\}.$$

Comparing with Eqs. (5.6) and (5.7), we find that the right-hand side here converges to \(f^k(\mathcal{F}(\omega))\) as \(k \to \infty\). So Eq. (5.3) is true.

Now consider the set \(\mathcal{F}(\Omega_1)\). We have

$$\mathcal{F}(\Omega_1) = \mathcal{F} \left( \bigcup_{a \in A} \mathcal{F}_a \right) = \bigcup_{a \in A} f^a(\Omega_1)$$

It follows by uniqueness that \(\mathcal{F}(\Omega_1)\) must be the set attractor of the IFS \(\mathcal{F}_1\). Hence \(\mathcal{F}(\Omega_1) = \Omega_1\)

which is the first statement in Eq. (5.4). Now

$$\mathcal{F}(\Omega_1) = \{f(\omega_1,\omega_2,\ldots,\omega)(\Omega_1)\}$$

= first component of \(\{f(\omega_1,\omega_2,\ldots,\omega)(\Omega_1)\} \subset \Omega_1\}

= first component of \(\mathcal{F}(\Omega_1)\)

= first component of \(\mathcal{F}(\Omega_1) = \Omega_1\),

which contains the second statement in Eq. (5.4).

In a similar manner we consider the push-forward under \(\mathcal{F} : \Omega^V \to \Omega^V\) of the invariant measure \(\mu_V\) of the IFS \(\Phi = \{\Omega^V, \mathcal{F}_a, a \in A\}\). \(\mathcal{F}(\mu_V)\) is normalized, i.e. \(\mathcal{F}(\mu_V) \in \mathcal{P}(\Omega^V)\), because \(\mathcal{F}(\mu_V)(\Omega^V) = \mu_V(\Omega^V)\). We now show that \(\mathcal{F}(\mu_V)\) is the measure attractor the IFS \(\mathcal{F}\). The measure attractor of the IFS \(\Phi\) obeys

$$\mu_V = \sum_{a \in A} f^a \mathcal{F}(\mu_V).$$

Applying \(\mathcal{F}\) to both sides (i.e. constructing the push-forwards) we obtain

$$\mathcal{F}(\mu_V) = \mathcal{F} \left( \sum_{a \in A} f^a \mathcal{F}(\mu_V) \right) = \sum_{a \in A} f^a(\mathcal{F}(\mu_V)) = \sum_{a \in A} f^a \mathcal{F}(\mu_V),$$

where in the last step we have used the key Eq. (5.3). So \(\mathcal{F}(\mu_V)\) is the measure attractor of the IFS \(\mathcal{F}_1\). Using uniqueness, we conclude \(\mathcal{F}(\mu_V) = \mu_V\) which is the first equation in Eq. (5.5).

Finally, observe that, for all \(B \in \mathcal{B}(\Omega)\),

$$\mathcal{F}(B) = f(\mu_V)(B, \Omega, H, \ldots) = \mu_V(\mathcal{F}^{-1}(B, \Omega, H, \ldots))$$

= \(\mu_V((\mathcal{F}^{-1}(B, \Omega, H, \ldots), \mathcal{F}^{-1}(H), \ldots, \mathcal{F}^{-1}(H)))\) [using Eq. (5.6)]

= \(\mu_V((\mathcal{F}^{-1}(B, \Omega, H, \ldots))\)

(since \(\mathcal{F} : \Omega \to H\))

= \(\mu_V((\mathcal{F}^{-1}(B, \Omega, H, \ldots))\)

[by Definition (4.10)] = \(\mathcal{F}(\mu_V)(B)\).

This contains the second equation in Eq. (5.5).
In a similar way, we obtain the alternate proof of
Theorem 16. Simply lift Theorem 10 to the domain
of the IFS $F_1$ using $F : \Omega^V \rightarrow \Omega^V$.

The code tree $\Phi(a_1, a_2, \ldots) = \Sigma$ is called a tree
address of the $V$-variable fractals as follows. At any “magnification",
you can identify the $V$-variable fractal set is made of $V$ “forms” or
“shapes”:

**Theorem 17.** Let $M \in \mathcal{M}_V$ be any $V$-variable
fractal set. Let $\epsilon > 0$ be given. Then $M$ is a finite
union of continuous transformations of at most $V$
distinct compact subsets of $X$, and the diameter of
each of these transformed sets is at most $\epsilon$.

**Proof.** Choose $n$ so that $\ell^n < \epsilon$. Note that
$\mathcal{M}_V = \bigcup_{a \in A} f_a(\mathcal{M}_V)$
$= \bigcup_{a_1, a_2, \ldots, a_n \in A} f_{a_1} \circ f_{a_2} \circ \cdots \circ f_{a_n}(\mathcal{M}_V)$.

Hence, since $(M, M, \ldots, M) \in \mathcal{M}_V$ it follows that
there exists $(K_1, K_2, \ldots, K_V) \in \mathcal{M}_V$ such that
$M \in$ first component of
$\bigcup_{(a_1, a_2, \ldots, a_n) \in A} f_{a_1} \circ f_{a_2} \circ \cdots \circ f_{a_n}(K_1, K_2, \ldots, K_V)$.

Each set in the union has diameter at most $\epsilon$.

**Example 5.** This example is similar to Example 4,
with $M = N = V = 2$. The goal is to illustrate
Theorem 18. Figure 22 shows, from left to
top, from top to bottom, a sequence of six succes-
sive images, illustrating successive $2$-variable frac-
tals, corresponding to a superIFS of two IFSs. Each
IFS consists of two projective transformations, each
mapping the unit square $\square$ into itself. The images
were obtained by running the random iteration
algorithm, as described in Sec. 1.2. The initial image
on each screen was a blue convex region contained
in a $400 \times 400$ array representing $\square$, and the images
shown correspond to one of the discretized screens
after 40, 41, 42, 43, 44, and 45 iterations. The key
features of the transformations can be deduced from
the images. (For example, one of the transforma-
tions of one of the IFSs, interpreted as a mapping
from one screen to the next, maps the top middle
image to the top two objects in the top left image.)
Each of these images, at several scales, looks as
though it is the union of projective transformations
of at most two distinct sets.

**Theorem 19.** The set of $V$-variable fractal sets
associated with the superIFS $\mathcal{F}_V$ converges to the set
of fractal sets associated with the superIFS $F$ intro-
duced in Sec. 3.3; that is, in the metric of $H(\Omega(X))$,
$$\lim_{V \rightarrow \infty} \mathcal{F}_V = \mathcal{F}.$$ (5.8)

Moreover, if the probabilities $\{P^a|a \in A\}$ obey
Eq. (4.6), then in the metric of $H(\Omega(X))$
$$\lim_{V \rightarrow \infty} \mathcal{F}_V = \Psi,$$
where $\Psi$ is the stationary measure on random frac-
tal sets associated with the superIFS $F$.

**Proof.** We have, using the mapping
$F : \Omega \rightarrow H(\Omega)$,
$$\lim_{V \rightarrow \infty} \mathcal{F}_V = \lim_{V \rightarrow \infty} \mathcal{F}(\Omega_V)$$
$$= F\left(\lim_{V \rightarrow \infty} \Omega_V\right)$$ (since $F : \Omega \rightarrow H(\Omega)$
is continuous by Theorem 7)
$$= F(\Omega)$$ (by Theorem 12)
$$= \mathcal{F}$$ [by Eq. (3.7)].

Similarly, using the mapping $F : \Omega \rightarrow P(\Omega)$, have
$$\lim_{V \rightarrow \infty} \mathcal{F}_V = \lim_{V \rightarrow \infty} F(\rho_V)$$ (by Theorem 17)
$$= F\left(\lim_{V \rightarrow \infty} \rho_V\right)$$ (since $F : \Omega \rightarrow P(\Omega)$
is continuous by Theorem 7)
$$= F(\rho)$$ (by Theorem 12)
$$= \Psi$$ [by Eq. (3.8)].

### 5.2. Contraction Mappings on $P^V$ and the Superfractal
Measures $\mathcal{F}_V$

Recall that $P = P(\Omega)$. Let $P^V = (P(\Omega))^V$. In this
section we follow the same lines as in Sec. 5.1, con-
structing an IFS using the individual IFSs of the
superIFS $F$, except that here the underlying space
is $P^V$ instead of $H^V$.

**Definition 11.** Let $V \in \mathcal{N}$, let $A$ be the index
set introduced in Eq. (4.1), let $F$ be given as in
Eq. (3.1), and let probabilities $\{P^a|a \in A\}$ be given
as in Eq. (4.5). Define
$$f^a : P^V \rightarrow P^V$$
May 31, 2005 9:51 00279

by
\[
\begin{align*}
 f^a(\mu) & = \left( \sum_{m=1}^{M} p_{m}^a f_m^a (\mu_{a,1}), \sum_{m=1}^{M} p_{m}^a f_m^a (\mu_{a,2}), \ldots, \sum_{m=1}^{M} p_{m}^a f_m^a (\mu_{a,2^l}) \right) \\
 & = \left( \sum_{m=1}^{M} p_{m}^a f_m^a (\mu_{a,1}), \sum_{m=1}^{M} p_{m}^a f_m^a (\mu_{a,2}), \ldots, \sum_{m=1}^{M} p_{m}^a f_m^a (\mu_{a,2^l}) \right) \\
 & \leq \max \left\{ \sum_{m=1}^{M} p_{m}^a f_m^a (\varphi_{v,m}) \right\} \\
 & \leq \max \{ 1 \cdot d_p (\mu_{v,1}, \mu_{v,2}, \ldots, \mu_{v,2^l}) \} \\
 & \leq 1, \quad \left( \varphi_{v,1}, \varphi_{v,2}, \ldots, \varphi_{v,2^l} \right) \\
\end{align*}
\]  

(5.9)

Theorem 20. \( \tilde{F}_V \) is an IFS with contractivity factor \( l \).

Proof. We only need to prove that the mapping \( f^a : \mathbb{P}^l \to \mathbb{P}^l \) is contractive with contractivity factor \( l, \forall a \in A \). Note that, \( \forall \mu = (\mu_1, \mu_2, \ldots, \mu_M), \varphi = (\varphi_1, \varphi_2, \ldots, \varphi_M) \in \mathbb{P}^M \),

\[
\begin{align*}
 d_p (\mu_1, \varphi_1, \varphi_2, \ldots, \varphi_M) & \leq \max\left\{ \sum_{m=1}^{M} d_p (\mu_{a,m}, \varphi_{a,m}) \right\} \\
 & = 1, \quad \left( \varphi_{v,1}, \varphi_{v,2}, \ldots, \varphi_{v,2^l} \right) \\
\end{align*}
\]

(5.10)

The set attractor of the IFS \( \tilde{F}_V \) is \( \tilde{H}_V \) is a subset of \( \mathbb{P}^l \), a set of \( V \)-tuples of probability measures on \( X \). As we will see, each of these measures is supported on a \( V \)-variable fractal set belonging to the superfractal \( \tilde{H}_V \). The measure attractor of the IFS \( \tilde{F}_V \) is a probability measure \( \tilde{P}_V \) in \( P(\mathbb{P}^l) \), namely a probability measure on a set of \( V \)-tuples of normalized measures, each one a random fractal measure. The random iteration algorithm corresponding to the IFS \( \tilde{F}_V \) may be used to approximate sequences of points in \( \tilde{H}_V \), namely vectors of measures on \( X \), distributed according to the probability measure \( \tilde{P}_V \).

As in Sec. 5.1, we define \( \tilde{H}_{V,v} \) to be the set of \( v \)th components of sets in \( \tilde{H}_V \), for \( v \in \{1,2,\ldots,V\} \).

Theorem 21. For all \( v \in \{1,2,\ldots,V\} \) we have

\[
\tilde{H}_{V,v} = \tilde{H}_{V,v}. 
\]

When the probabilities in the IFS \( \tilde{F}_V \) are given by Eq. (4.6), then starting at any initial \( V \)-tuple of probability measures on \( X \), the probability measures \( \mu \in P(X) \) that occur in the \( v \)th component of points...
Fractal Valued Random Iteration Algorithm and Fractal Hierarchy 139

Fig. 23 Three successive 2-variable fractal measures computed using the random iteration algorithm in Theorem 21 applied to the superIFS in Example 6. The pixels in the support of each measure are colored either black or a shade of green, using a similar technique to the one used in Example 2. The intensity of the given of a pixel is a monotonic increasing function of the measure of the pixel.

produced by the random iteration algorithm after \( n \) steps converge weakly to the marginal probability measure

\[
\tilde{\Phi}_{V,1}(B) := \tilde{\Phi}_V(B, P, P, \ldots, P) \quad \forall B \in \mathcal{B}(P),
\]

independently of \( n \), almost always, as \( n \to \infty \).

**Proof.** The direct way to prove this theorem is to parallel the proof of Theorem 10, using the maps \( \{ f^a: \mathbb{P}^V \to \mathbb{P}^V \mid a \in A \} \) in place of the maps \( \{ \eta^a: \Omega^V \to \Omega^V \mid a \in A \} \). It is simpler however to lift Theorem 10 to the domain of the IFS \( \mathcal{F}_V \) using \( \mathcal{F}: \Omega^V \to \mathbb{P} = \mathcal{P}(\mathcal{X}) \) which is defined in Theorem 22 with the aid of the mapping \( \tilde{\mathcal{F}}: \Omega \to \mathbb{P} = \mathcal{P}(\mathcal{X}) \) introduced in Theorem 7. We omit the details as they are straightforward.

We call \( \tilde{\Phi}_{V,1} \) a superfractal set of measures (of variability \( V \)). Points in \( \tilde{\Phi}_{V,1} \) are called \( V \)-variable fractal measures.

**Example 6.** See Fig. 23. This example corresponds to the same superIFS as in Example 4. The probabilities of the functions in the IFSs are \( p_1^V = p_2^V = 0.74 \), and \( p_3^V = p_4^V = 0.26 \). The IFSs are assigned probabilities \( P_1 = P_2 = 0.5 \).

The following theorem tells us in particular that the set of \( V \)-trees is a useful code space for \( V \)-variable fractal measures, because the action of the IFS \( \Phi \) on the space of \( V \)-tuples of code trees is conjugate to the action of the IFS \( \tilde{\mathcal{F}}_V \) acting on \( V \)-tuples of normalized measures.

**Theorem 22.** Let \( a \in A \) and \( \eta^a: \Omega^V \to \Omega^V \) be defined as in Theorem 9. Let \( f^a: \mathbb{P}^V \to \mathbb{P}^V \) be defined as in Theorem 20. Let \( \tilde{\mathcal{F}}: \Omega \to \mathbb{P} = \mathcal{P}(\mathcal{X}) \) be the mapping introduced in Theorem 7. Define \( \mathcal{F}: \Omega^V \to \mathbb{P}^V = (\mathcal{P}(\mathcal{X}))^V \) by

\[
\mathcal{F}(\omega_1, \omega_2, \ldots, \omega_V) = (\mathcal{F}(\omega_1), \mathcal{F}(\omega_2), \ldots, \mathcal{F}(\omega_V)),
\]

for all \( (\omega_1, \omega_2, \ldots, \omega_V) \in \Omega^V \). Then

\[
\mathcal{F}(\eta^a(\omega)) = f^a(\mathcal{F}(\omega)) \quad \forall a \in A, \quad \omega \in \Omega^V.
\]

Also

\[
\mathcal{F}(\Omega^V) = \tilde{\Phi}_V \quad \text{and} \quad \tilde{\mathcal{F}}(\Omega^V) = \tilde{\Phi}_{V,1},
\]

where \( \Omega^V \) denotes the set of \( V \) components of members of \( \Omega^V \). Similarly, when the probabilities in the IFS \( \mathcal{F}_V \) of Eq. (5.10), are given by Eq. (4.6), we have

\[
\mathcal{F}(\rho^V) = \tilde{\Phi}_V \quad \text{and} \quad \tilde{\mathcal{F}}(\rho^V) = \tilde{\Phi}_{V,1},
\]

where \( \rho^V \) is the marginal probability distribution given by Eq. (4.10).

**Proof.** The proof is entirely parallel to the proof of Theorem 17, using \( \tilde{\mathcal{F}} \) in place of \( \mathcal{F} \) and is omitted.

**Definition 12.** The code tree \( \Phi(a_{123\ldots}) \) is called a tree address of the \( V \)-variable fractal measure \( \mathcal{F}_V(a_{123\ldots}) \).

The mapping \( \tilde{\mathcal{F}}: \Omega^V \to \tilde{\Phi}_{V,1} \) together with Theorem 11 allows us to characterize \( V \)-variable fractals as follows:

**Theorem 23.** Let \( \tilde{\mu} \in \tilde{\Phi}_{V,1} \) be any \( V \)-variable fractal measure. Let \( \epsilon > 0 \) be given. Then \( \tilde{\mu} \) is a finite weighted superposition of continuous transformations of at most \( V \) distinct normalized measures supported on compact subsets of \( \mathcal{X} \), and the diameter of the support of each of these transformed measures is at most \( \epsilon \).
Proof. Choose \( n \) so that \( p^n < \epsilon \). Note that
\[
\hat{h}_V = \bigcup_{a \in A} f^n(\hat{h}_V)
\]
\[
= \bigcup_{(a_1, a_2, \ldots, a_n) \in A} f^{a_1} \circ f^{a_2} \circ \cdots \circ f^{a_n}(\hat{h}_V).
\]
Hence, since \((\hat{\mu}, \hat{\rho}, \ldots, \hat{\nu}) \in \hat{h}_V\) if it follows that there exists \((\varpi_1, \varpi_2, \ldots, \varpi_V) \in \hat{h}_V\) such that \(\hat{\mu}\) is first component of
\[
\bigcup_{(a_1, a_2, \ldots, a_n) \in A} f^{a_1} \circ f^{a_2} \circ \cdots \circ f^{a_n}(\varpi_1, \varpi_2, \ldots, \varpi_V).
\]
Inspection of Eq. (5.9) shows that each of the measures in the set of measures on the right-hand side here is as stated in the theorem. \(\square\)

Example 7. See Fig. 24. This corresponds to the same superIFS as used in Example 5, but here the measure is rendered in shades of blue to provide a pictorial illustration of Theorem 23. The three successive images were computed with the aid of the random iteration algorithm in Theorem 21, a new rendered measure theoretic image being produced at each iteration. At each discernable scale, approximately, each picture appears to have the property that is a superposition of a number of “little pictures” belonging to one of two equivalence classes. Pictures belonging to an equivalence class in this case are related by a projective transformation here with a scaling of brightness.

Theorem 24. The set of \( V \)-variable fractal measures associated with the superIFS \( F_V \) converges to the set of fractal measures introduced in Sec. 3.3; that is, in the metric of \( H(P(X)) \)
\[
\lim_{V \to \infty} \hat{h}_{V,1} = \hat{\mu}.
\]
If the probabilities \( P[a_1 \in A] \) obey Eq. (4.6), then in the metric of \( H(P(X)) \)
\[
\lim_{V \to \infty} \hat{\Psi}_{V,1} = \hat{\Psi},
\]
where \(\hat{\Psi}\) is the stationary measure on fractal sets associated with the superIFS \( F \).

Proof. We have, using the mapping \( \hat{F} : \Omega \to P = P(X) \),
\[
\lim_{V \to \infty} \hat{h}_{V,1} = \hat{h}(\Omega) \] (by Theorem 22)
\[
= \hat{F} \left( \lim_{V \to \infty} \Omega_{V,1} \right) \] (since \( \hat{F} : \Omega \to P = P(X) \) is continuous by Theorem 7)
\[
= \hat{F}(\Omega) \] (by Theorem 12)
\[
= \hat{\mu} \] [by Eq. (3.7)].

We have, using the mapping \( \hat{F} : \Omega \to P(P) \),
\[
\lim_{V \to \infty} \hat{\Psi}_{V,1} = \hat{\Psi}(\rho(V)) \] (by Theorem 22)
\[
= \hat{F} \left( \lim_{V \to \infty} \rho(V) \right) \] (since \( \hat{F} : \Omega \to P(P) \) is continuous by Theorem 7)
\[
= \hat{F}(\rho) \] (by Theorem 12)
\[
= \hat{\Psi} \] [by Eq. (3.8)]. \(\square\)

5.3. Fractal Dimensions

Here we quantify and compare the Hausdorff dimensions of fractals corresponding to a (super) IFS of similitudes on \( \mathbb{R}^N \) for some \( K \in \mathbb{N} \) that obeys the open set condition in the following four cases: deterministic fractals, standard random fractals, homogeneous random fractals (\( V = 1 \)), and \( V \)-variable fractals (\( V > 1 \)). The functions of the IFS \( F^n \) are of the form \( f^n(x) = x^n \rho^n \), where \( \rho^n \) is an orthonormal transformation, \( x^n \in [0, 1) \), and \( \rho^n \in \mathbb{R}^N \), for all \( n \in \{1, 2, \ldots, N\} \) and \( m \in \{1, 2, \ldots, M\} \).

5.3.1. Deterministic fractals

In this case there is only one IFS, say \( F^1 \). By Theorem 4, the Hausdorff dimension of the corresponding fractal set \( A \) is \( D \), the unique solution of
\[
\sum_{m=1}^{M} (s_m^1)^D = 1.
\]

Fig. 24 Three successive 2-variable fractal measures, in shades of blue. Illustrates the “shapes” and “forms” theorem. See Example 7.
Example 8. Suppose $K \geq 2$. Let the IFS $F^1$ consists of three similitudes with $s_1^1 = s_2^1 = s_3^1 = \frac{1}{3}$ and that the fixed points are not collinear. Then the set attractor of $F^1$ is the Sierpinski triangle with vertices at the three fixed points. Its fractal dimension $D_1$ is given by $3^{1/\log 2} = 1$ which implies $D_1 = \frac{\log 3}{\log 2} = 1.585$.

Let the IFS $F^2$ consist of three similitudes with $s_1^2 = s_2^2 = s_3^2 = \frac{1}{2}$ and the same fixed points as $F^1$. Then the fractal dimension $D_2$ of the set attractor of $F^2$ is given by $\frac{1}{2} \log (3^2) = 1$ which implies $D_2 = 1$.

5.3.2. Random fractals

By Theorem 8, the Hausdorff dimension $D_R$ of \( \Psi \)-almost all of the random fractal sets for the superIFS $F$ is given by

$$\sum_{n=1}^{N} P_n \sum_{m=1}^{M} (s_m^n)^{D_R} = 1.$$  

Example 9. Let the superIFS be $\{\square, F^1, F^2; P_1 = P_2 = 0.5\}$ where the IFS’s are defined in Example 8. Then the fractal dimension $D_R$ of \( \Psi \)-almost all of the random fractals in the set is given by $\frac{1}{2} 3^{1/\log 2} + \frac{1}{2} 3^{1/\log 2} = 1$ which implies $D_R = 1.262$.

5.3.3. Homogeneous random fractals ($V = 1$)

The case of homogeneous random fractals corresponds to $V = 1$. Each run of the experiment gives a different random Sierpinski triangle.

Theorem 25. Let the superIFS $F$ be as specified as in Theorem 8. Let $V = 1$. Then for $\Psi_{1.3}$ almost all $A \in \partial_{1.1}$

$$\text{dim}_H A = D$$

where $D$ is the unique solution of

$$\sum_{n=1}^{N} P_n \log \sum_{m=1}^{M} (s_m^n)^D = 1.$$  

Example 10. For the case of the superIFS in Example 9, whose 1-variable fractal sets we refer to as homogeneous random Sierpinski triangles, the Hausdorff dimension $D$ of almost all of them is given by $\frac{1}{2} \log (3^2) + \frac{1}{2} \log (3^2) = 0$, which implies $D = 2 \log 3/(\log 2 + \log 3) = 1.226$.

5.3.4. $V$-variable fractals ($V \geq 1$)

Let $(a_1, a_2, \ldots) \in \mathbb{A}^\infty$ denote an i.i.d. sequence of indices, with probabilities $\{p^m[a] \in \mathbb{A}\}$ according to Eq. (4.6). Define, for $\alpha \in [0, 1]$ and $a \in \mathbb{A}$, the $V \times V$ flow matrix

$$M_{\alpha, w}(a) = \sum_{m \in \{v, w\}} (s_m^a)^{\alpha},$$

and let us write

$$M_k^{\alpha} = M_k^{\alpha, w}(a) = M_{\alpha, w}(a).$$

We think of $s_m^a$ as being the “flow” through the $m^{th}$ channel from screen $v$ to screen $w$, where $v, w = w$. The sequence of random matrices $M_0^{\alpha, w}, M_1^{\alpha, w}, \ldots$ is i.i.d., again with probabilities induced from $\{P_1, P_2, \ldots, P_V\}$. For any real square matrix $M$ we define the norm $\|M\|$ to be the sum of the absolute values of its entries. By the Furstenberg-Kesten theorem,\ref{footnote:FK}

$$\gamma(\alpha) := \lim_{k \to \infty} k^{-1} \log \|M_k^{\alpha}(\alpha) \circ \cdots \circ M_0^{\alpha}(\alpha)\|$$

exists and has the same value with probability one. Provided that the superIFS obeys the open set condition, we have shown in Baransley et al.\ref{footnote:Baransley} that the unique value of $D \in [0, \infty)$ such that

$$\gamma(D) = 0$$

is the Hausdorff dimension of $\Psi_{1.1}$ almost all $A \in \partial_{V.1}$.

Kingman remarks that in general the calculation of $\gamma$ “has pride of place among the unsolved problems of subadditive ergodic theory”.\ref{footnote:Kingman}. However it is possible to estimate numerically. Namely, generate random copies of $M^k$ and iteratively compute $M^1, M^2, \ldots, M^k$ and hence $k^{-1} \log \|M_k^{\alpha}(\alpha) \circ \cdots \circ M_0^{\alpha}(\alpha)\|$ for $k = 1, 2, \ldots$. The limit will give $\gamma(\alpha)$. (Even for large $V$ this will be quick since the $M^k$ are sparse.) One could now use the bisection method to estimate $D$.

6. APPLICATIONS

Fractal geometry plays some role in many application areas, including the following: in biology: breast tissue patterns, structure and development of plants, blood vessel patterns, and morphology of fern fronds. In chemistry: pattern-forming alloy solidification, and diffusion processes. In physics: transport in porous media, patterns formed during phase transitions in statistical mechanics, dynamical systems, turbulence and wave propagation. In
probabilities, for example. In many of these areas it is clearly desirable to use random fractals; for example, random fractals can be used in connection with diverse mathematical modeling application areas including Brownian motion, oil-wells, critical phenomena in statistical physics, for example, associated with lattice gases and percolation, stock-market prices in finance, and in computer graphics they can be used to represent diverse picture types including natural images and textures. But random fractals are hard to compute, which may have held up the development of some applications, while deterministic fractals, which can be computed relatively easily, may not be rich enough to provide convenient models for the applications to which one would want to apply them.

Thus we believe that \( V \)-variable fractals could find many applications; they can be computed easily, with rapid access to many examples, contain a controllable amount of “randomness”, and have many of the advantages of fractals in general: for similitudes, with an open set condition, their fractal dimension may be computed, they are resolution independent, and they are in general geometrically complex at all levels of magnification, while being expressed with relatively small amounts of information, coefficients of affine transformations and some probabilities, for example.

### 6.1. Space-Filling Curves

Space-filling curves can be constructed with the aid of IFS theory, see, for example, Sagan. These curves have many applications, including adaptive multigrid methods for numerical computation of solutions of PDEs and hierarchical watermarking of digital images. Here we note that interesting \( V \)-variable space-filling curves, and finite resolution approximants to them, can be produced.

#### Example 11

Let \( M = 3, V = 2, N = 2 \). The IFS \( F_1 = \{ \square; f_1^1, f_1^2 \} \) consists of affine maps whose actions we explain with the aid of the left-hand diagram in Fig. 25. \( \square \) is the unit square in the diagram, while \( f_1^1(\square) \) is the lower left square, \( f_1^2(\square) \) is the upper left square, and \( f_1^3(\square) \) is the rectangle on the right. The transformations are chosen so that

\[
\begin{align*}
    f_1^1(\square) &= \square, \quad f_1^2(\square) = AB, \quad f_1^3(\square) = BC.
\end{align*}
\]

Specifically, \( f_1^1(x, y) = (\frac{1}{2}x, \frac{1}{2}y), f_1^2(x, y) = (\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}y + 1), f_1^3(x, y) = (\frac{1}{2}x - \frac{1}{2}, \frac{1}{2}y - 1). \)

The IFS \( F_2 = \{ \square; f_2^1, f_2^2, f_2^3 \} \) is explained with the aid of the right-hand diagram in Fig. 25; \( f_2^1(\square) \) is the lower left rectangle, \( f_2^2(\square) \) is the upper left rectangle, and \( f_2^3(\square) \) is the rectangle on the right; such that \( f_2^1(\square) = OA, f_2^2(\square) = AB, f_2^3(\square) = BC \), and \( f_2^3(\square) = BC \). Specifically, \( f_2^1(x, y) = (\frac{1}{2}x, \frac{1}{2}y), f_2^2(x, y) = (-\frac{1}{2}x + \frac{1}{2}, -\frac{1}{2}y + 1), f_2^3(x, y) = (\frac{1}{2}x - \frac{1}{2}, -\frac{1}{2}y - 1). \)

Neither of the IFSs here is strictly contractive, but each is contractive “on the average”, for any assignment of positive probabilities to the constituent functions. We assign probabilities \( P_1 = P_2 = 0.5 \) to the individual IFSs. An initial image consisting of the line segment \( \square \) is chosen on both screens, and the random iteration algorithm

![Fig. 25](image-url)
is applied; typical images produced after five iterations are illustrated in Fig. 26; an image produced after seven iterations is shown in Fig. 27. Each of these images consists of line segments that have been assigned colors according to the address of the line segment, in such a way as to provide some consistency from one image to the next.

6.2. Computer Graphics

New techniques in computer graphics are playing an increasingly important role in the digital content creation industry, as evidenced by the succession of successes of computer generated films, from “Toy Story” to “Finding Nemo”. Part of the appeal of such films is the artistic quality of the graphics.

Here we point out that \( V \)-variable fractals are able to provide new types of rendered graphics, significantly extending standard IFS graphics.

Example 12. Here \( N = 2 \), \( V = 2 \), \( M = 4 \). The two IFSs are given by

\[
F^n = \{□; f_1^n, f_2^n, f_3^n, f_4^n \} \quad \text{for} \quad n \in \{1, 2\},
\]

where \( □ \subset R^2 \), and each \( f_m^n : □ \rightarrow □ \) is a projective transformation. The colors were obtained as follows.

A computer graphics rendering of the set attractor of \( F^1 \) is shown in Fig. 28, and of \( F^2 \) in Fig. 29.

The coloring of each of these two figures was obtained with the aid of an auxiliary IFS acting on the cube \( C := [0, 255]^3 \subset R^3 \) given by

\[
G := \{C; g_1^1, g_2^1, g_3^1, g_4^1 \}
\]

where each \( g_m^1 \) is a contractive (in the Euclidean metric) affine transformation,
represented by a $3 \times 3$ matrix and a $3 \times 1$ vector. For $n \in \{1, 2\}$ discretized approximations, of the same resolution, to the attractors of both IFSs $F^n$ and $G$ were calculated via the deterministic algorithm (Corollary 1); each pixel on the attractor of the IFS $F^n$ was assigned the color whose red, green and blue components, each an integer from 0 to 255, were the three coordinates of the point on the attractor of $G$ with the same code space address. At those points in the attractor of $F^n$ with multiple code space addresses, the lowest address was chosen.

The superIFS we use is

$$F = \{ \emptyset; F^1, F^2; P^1 = 0.5, P^2 = 0.5 \}$$

with $V = 2$. Then Figs. 30 and 31 show two examples, from among many different but similar ones, all equally visually complex, of computer graphics of 2-variable fractals for this superIFS, computed using the new random iteration algorithm. The images were rendered in much the same way as the images of the attractor sets of $F^1$ and $F^2$ were rendered above. The essential difference is the meaning of a “code space address” of a point on a $V$-variable fractal, which we define to be the sequence of lower indices of a sequence of functions that converges to the point; for example, the point

$$\lim_{k \to \infty} f^1_2 \circ f^2_1 \circ f^1_2 \circ f^2_1 \circ \cdots \circ f^m_k(x)$$
corresponds to the address 212122 \ldots m_k \ldots, in the obvious notation.

6.3. V-Variable Fractal Interpolation

The technique of fractal interpolation has many applications including modeling of speech signals, altitude maps in geophysics, and stock-market indices. A simple version of this technique is as follows. Let a set of real interpolation points in \{(x_i, y_i) \in \mathbb{R}^2 | i = 0, 1, \ldots, I\} be given. It is desired to find a continuous function \( f : [x_0, x_I] \rightarrow \mathbb{R} \) such that \( f(x_i) = y_i, \forall i \in \{0, 1, \ldots, M\} \), such that its graph \( G = \{(x, y) \in \mathbb{R}^2 : y = f(x)\} \) is a fractal, possibly with specified fractal dimension. Introduce the IFS

\[
F = \{\mathbb{R}^2; f_1, f_2, \ldots, f_M\}
\]

with

\[
f_m(x, y) = (a_{m1}x + c_{m1}y + y_m, a_{m2}x + c_{m2}y + y_m),
\]

where the real coefficients \( a_{m1}, c_{m1}, c_{m2}, d_{m1}, \) and \( e_m \) are chosen so that

\[
f_m(x_0, y_0) = y_{m-1}; \quad f_m(x_0, y_0) = y_m,
\]

and \( a_{m1} \in [0, 1], \) for \( m \in \{1, 2, \ldots, M\} \). Then the attractor of the IFS is the graph of a function \( f \) with the desired properties, its dimension being a function of the free parameters \( \{d_{m1}, m = 1, 2, \ldots, M\} \).

Now let the superIFS \( F = \{\square; F_1, F_2, F_3\} \) for some \( V \), consist of two IFSs both of which provide fractal interpolations of the data. Then all of the elements of the corresponding superfractal will be graphs of continuous functions that interpolate the data, have the property of \( V \)-variability, and may be sampled using the random iteration algorithm.

7. GENERALIZATIONS

It is natural to extend the notions of \( V \)-variable fractals, superIFS and superfractal to include the case of maps contractive on the average, more than a finite number of maps, more than a finite number of IFSs, IFSs with a variable number of maps, IFSs operating on sets which are not necessarily induced by point maps, other methods of constructing the probabilities for a superIFS, probabilities that are dependent upon position, etc. But for reasons of simplicity and in order to illustrate key features we have not treated these generalizations at any length.

ACKNOWLEDGMENTS

This work was carried out mainly at the Australian National University and was partially supported by the Australian Research Council.

REFERENCES