

# A FRACTAL VALUED RANDOM ITERATION ALGORITHM AND FRACTAL HIERARCHY

MICHAEL BARNSLEY, JOHN HUTCHINSON, AND ÖRJAN STENFLO

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 memory usage. The algorithm is illustrated by some computed examples. Some variants, special cases, generalizations of the framework, and potential applications are mentioned.

## 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 [16]. A different approach using a “forward” algorithm was developed in Barnsley and Demko [4].

Falconer [11], Graf [14] and Mauldin and Williams [21] randomized each step in the backward construction algorithm to obtain random fractal sets. Arbeiter [1] introduced and studied random fractal measures; see also Olsen [23]. Hutchinson and Rùchendorff [17] and [18] introduced new probabilistic techniques which allowed one to consider more general classes of random fractals. For further material see Zähle [28], Patzschke and Zähle [24], 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.

---

*Date:* October 14, 2003.

*2000 Mathematics Subject Classification.* Primary 28A80, 65C05; Secondary 60J05, 60G57, 68U05.

*Key words and phrases.* Iterated Function Systems, Random Fractals, Markov Chain Monte Carlo.

Deterministic fractal sets and measures are defined as the attractors of certain iterated function systems (IFSs), as reviewed in Section 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 modelling 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 superfractals. We also explain why we think these objects are of special interest and deserve attention.

We start with two pairs of contractive affine transformations,  $\{f_1^1, f_2^1\}$  and  $\{f_1^2, f_2^2\}$ , where  $f_m^n : \square \rightarrow \square$  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  $(\square_{1'}, \square_{2'})$ .

Initialize by placing an image on each of the Input Screens, as illustrated in Figure 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  $\{f_1^1, f_2^1\}$  or  $\{f_1^2, f_2^2\}$ , say  $\{f_1^{n_1}, f_2^{n_1}\}$ . Apply  $f_1^{n_1}$  to one of the images on  $\square_1$  or  $\square_2$ , selected randomly, to make an image on  $\square_{1'}$ . Then apply  $f_2^{n_1}$  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  $\square_{1'}$ . (For example, if black-and-white images are used, simply take the union of the black region of  $I$  with the black region on  $\square_{1'}$ , and put the result back onto  $\square_{1'}$ .)

(ii) Again pick randomly one of the pairs of functions  $\{f_1^1, f_2^1\}$  or  $\{f_1^2, f_2^2\}$ , say  $\{f_1^{n_2}, f_2^{n_2}\}$ . Apply  $f_1^{n_2}$  to one of the images on  $\square_1$ , or  $\square_2$ , selected randomly, to make an image on  $\square_{2'}$ . Also apply  $f_2^{n_2}$  to one of the images on  $\square_1$ , or  $\square_2$ , also

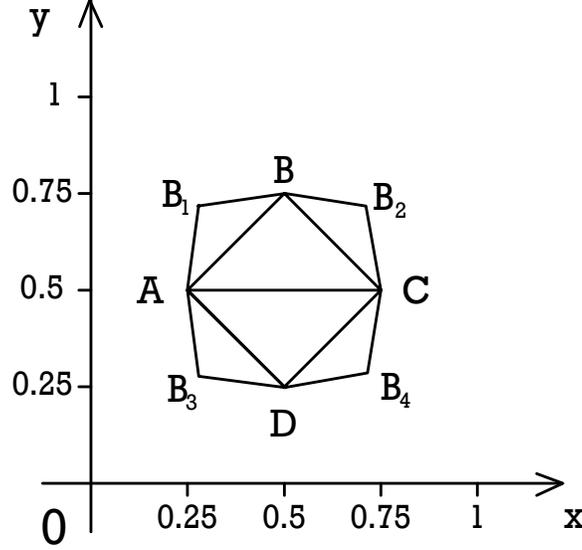


FIGURE 1. Triangles used to define the four transformations  $f_1^1, f_2^1, f_1^2,$  and  $f_2^2$ .

selected randomly, and overlay the resulting image on the image now already on  $\square_{2'}$ .

(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

$$(1.1) \quad f_1^1(x, y) = \left(\frac{1}{2}x - \frac{3}{8}y + \frac{5}{16}, \frac{1}{2}x + \frac{3}{8}y + \frac{3}{16}\right),$$

$$(1.2) \quad f_2^1(x, y) = \left(\frac{1}{2}x + \frac{3}{8}y + \frac{3}{16}, -\frac{1}{2}x + \frac{3}{8}y + \frac{11}{16}\right),$$

$$f_1^2(x, y) = \left(\frac{1}{2}x - \frac{3}{8}y + \frac{5}{16}, -\frac{1}{2}x - \frac{3}{8}y + \frac{13}{16}\right),$$

$$(1.3) \quad f_2^2(x, y) = \left(\frac{1}{2}x + \frac{3}{8}y + \frac{3}{16}, \frac{1}{2}x - \frac{3}{8}y + \frac{5}{16}\right).$$

We describe how these transformations act on the triangle  $ABC$  in the diamond  $ABCD$ , where  $A = (\frac{1}{4}, \frac{1}{2})$ ,  $B = (\frac{1}{2}, \frac{3}{4})$ ,  $C = (\frac{3}{4}, \frac{1}{2})$ , and  $D = (\frac{1}{2}, \frac{1}{4})$ . Let  $B_1 = (\frac{9}{32}, \frac{23}{32})$ ,  $B_2 = (\frac{23}{32}, \frac{23}{32})$ ,  $B_3 = (\frac{9}{32}, \frac{9}{32})$ , and  $B_4 = (\frac{23}{32}, \frac{9}{32})$ . See Figure 1. Then we have

$$\begin{aligned} f_1^1(A) &= A, f_1^1(B) = B_1, f_1^1(C) = B; \\ f_2^1(A) &= B, f_2^1(B) = B_2, f_2^1(C) = C; \end{aligned}$$

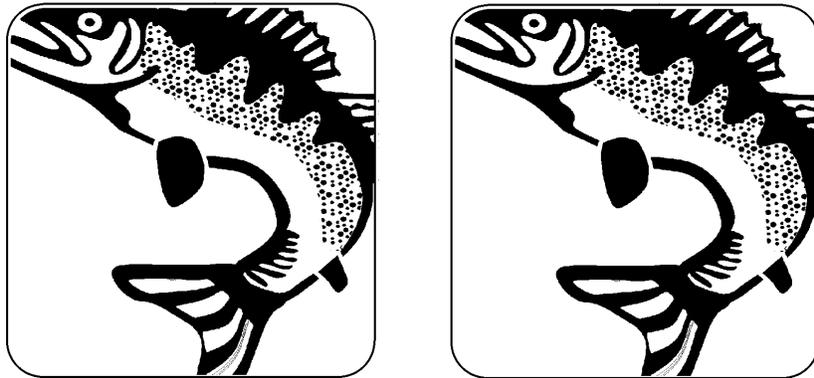


FIGURE 2. An initial image of a jumping fish on each of the two screens  $\square_1$  and  $\square_2$ .

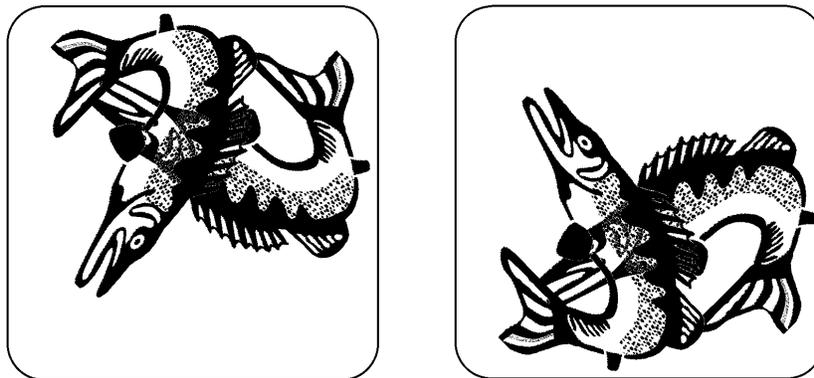


FIGURE 3. The pair of images after one iteration.

$$f_1^2(A) = A, f_1^2(B) = B_3, f_1^2(C) = D;$$

$$f_2^2(A) = D, f_2^2(B) = B_4, f_2^2(C) = C.$$

In Figure 2 we show an initial pair of images, two jumping fish, one on each of the two screens  $\square_1$  and  $\square_2$ . In Figures 3, 4, 5, 6, 7, 8, and 9, we show the start of the sequence of pairs of images obtained in a particular trial, for the first seven iterations. Then in Figures 10, 11, and 12, we show three successive pairs of computed screens, obtained after more than twenty iterations. These latter images are typical of those obtained after twenty 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 Figure 11 consist of the union of shrunken copies of the images in Figure 10, while the curves in Figure 12 are made from two shrunken copies of the curves in Figure 11.

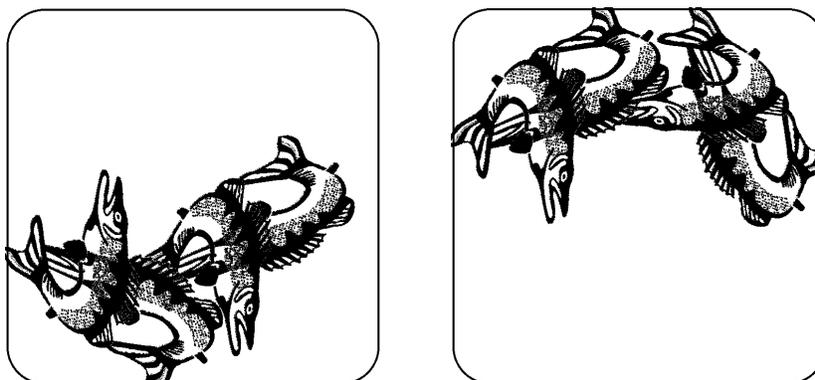


FIGURE 4. The two images after two iterations.

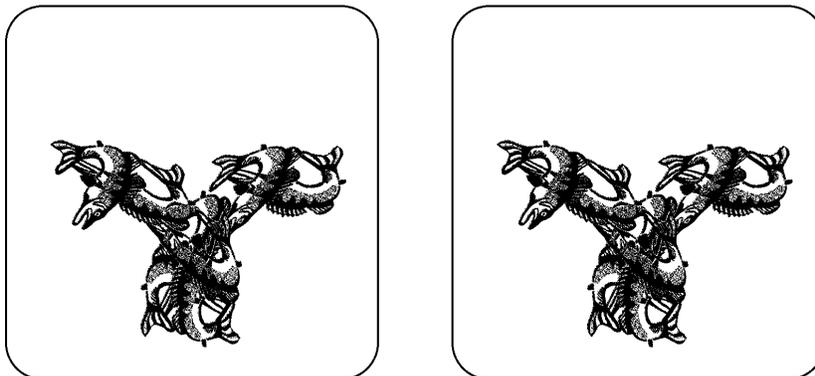


FIGURE 5. The two images after three iterations. Both images are the same.

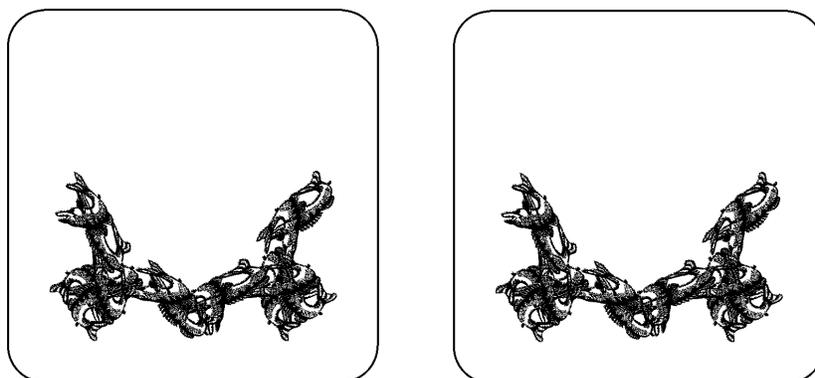


FIGURE 6. The two images after four iterations. Both images are again the same, a braid of fish.

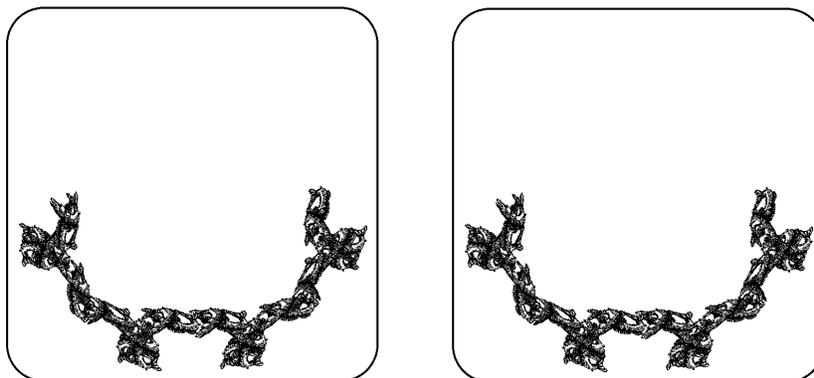


FIGURE 7. The two images after five iterations. The two images are the same.

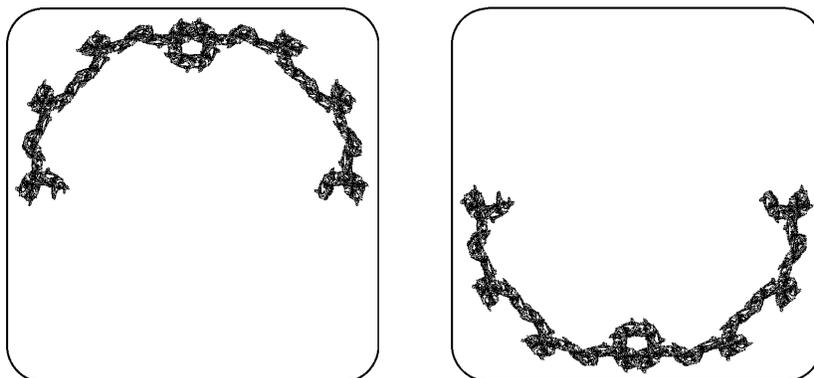


FIGURE 8. The two images after six iterations.

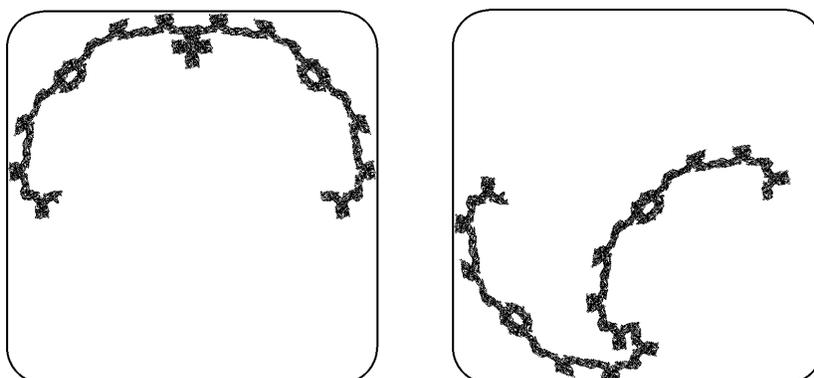


FIGURE 9. The two images after seven iterations.

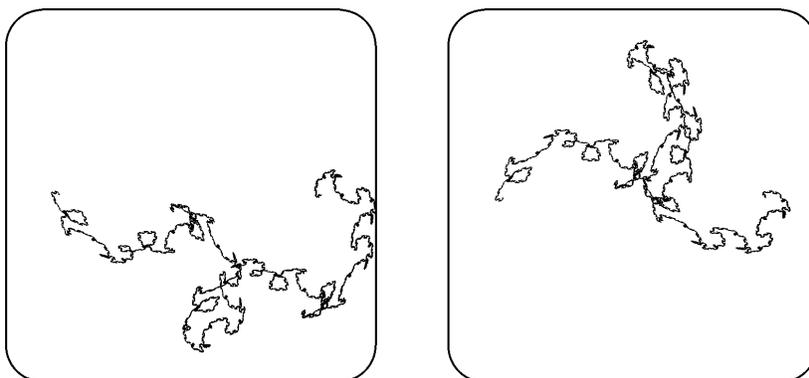


FIGURE 10. Images on the two screens  $\square_1$  and  $\square_2$  after a certain number  $L > 20$  of iterations. Such pictures are typical of the "stationary state" at the printed resolution.

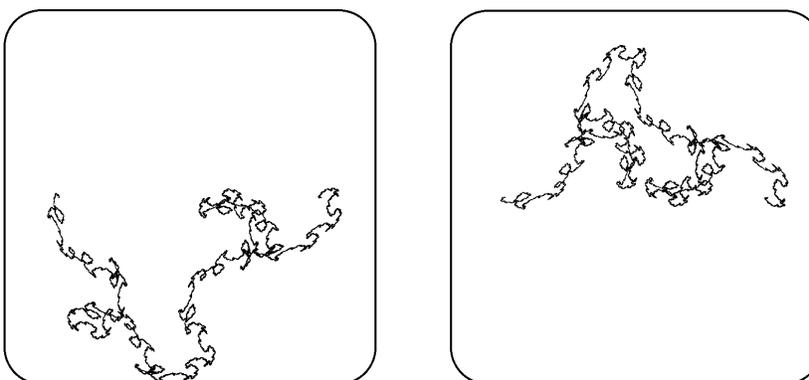


FIGURE 11. Images on the two screens after  $L + 1$  iterations.

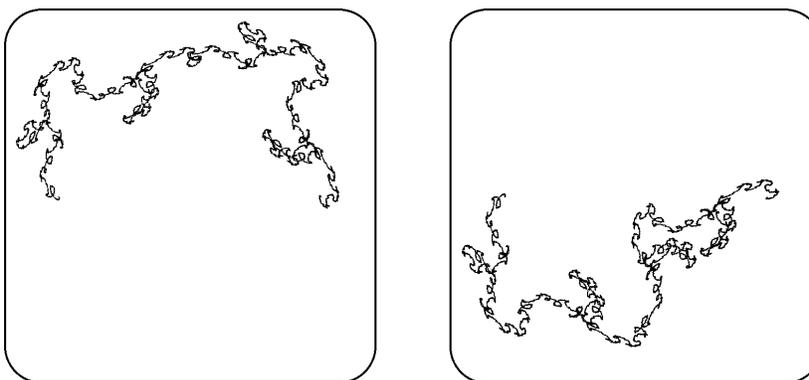


FIGURE 12. Images on the two screens after  $L + 2$  iterations.

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 Figures 10, 11, and 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 modelling 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.

The core machinery that we use is basic IFS theory, as described in [16] and [4]. So in Section 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 Section 3 with the definition of a *superIFS*, namely an IFS made of IFSs. We then introduce associated trees, in particular labelled trees, the space of code trees  $\Omega$ , and construction trees; then we review standard random fractals using the terminology of trees and superIFSs.

In Section 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  $\{\Omega^V; \eta^a, \mathcal{P}^a, a \in \mathcal{A}\}$  where  $\mathcal{A}$  is a finite index set,  $\mathcal{P}^a$ 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 labelled trees. This, in turn, enables us to establish the convergence, as  $V \rightarrow \infty$ , 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 Section 5 the discussion of trees in Section 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 Section 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  $V \rightarrow \infty$ , 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 Section 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 Section 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 Hambly and others [15], [2], [19], [27].

In Section 6 we describe some potential applications of the theory including new types of space-filling curves for digital imaging, geometric modelling and texture rendering in digital content creation, and random fractal interpolation for computer aided design systems. In Section 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 [4].

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

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

$$\underbrace{\mathbb{X} \times \mathbb{X} \times \dots \times \mathbb{X}}_{V \text{ TIMES}}$$

with metric

$$d(x, y) = d_{\mathbb{X}^V}(x, y) = \max \{d_{\mathbb{X}}(x_v, y_v) \mid v = 1, 2, \dots, V\}, \forall x, y \in \mathbb{X}^V,$$

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

In some applications, to computer graphics for example,  $(\mathbb{X}, d_{\mathbb{X}})$  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  $\mathbb{S} = \mathbb{S}(\mathbb{X})$  denote the set of all subsets of  $\mathbb{X}$ , and let  $C \in \mathbb{S}$ . We extend the definition of a function  $f : \mathbb{X} \rightarrow \mathbb{X}$  to  $f : \mathbb{S} \rightarrow \mathbb{S}$  by

$$f(C) = \{f(x) \mid x \in C\}$$

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

$$(1.4) \quad \mathcal{D}(A, B) = \max\{\min\{d_{\mathbb{X}}(x, y) \mid y \in B\} \mid x \in A\},$$

and define the *Hausdorff metric* by

$$d_{\mathbb{H}}(A, B) = \max\{\mathcal{D}(A, B), \mathcal{D}(B, A)\}.$$

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

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

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

$$d_{\mathbb{P}}(\mu, \nu) = \sup \left\{ \int_{\mathbb{X}} f d\mu - \int_{\mathbb{X}} f d\nu \mid f : \mathbb{X} \rightarrow \mathbb{R}, |f(x) - f(y)| \leq d_{\mathbb{X}}(x, y) \forall x, y \in \mathbb{X} \right\}.$$

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  $\mathbb{X}$ , [9]. We define the push-forward map  $f : \mathbb{P}(\mathbb{X}) \rightarrow \mathbb{P}(\mathbb{X})$  by

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

Again here we have extended the domain of action of the function  $f : \mathbb{X} \rightarrow \mathbb{X}$ .

We will use such spaces as  $\mathbb{P}(\mathbb{H}^V)$  and  $\mathbb{P}((\mathbb{P}(\mathbb{X}))^V)$  (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  $\mathbb{X}$ , and often we drop the subscript on the metric without ambiguity. So for example, we will write

$$d(A, B) = d_{\mathbb{H}((\mathbb{P}(\mathbb{X}))^V)}(A, B) \quad \forall A, B \in \mathbb{H}((\mathbb{P}(\mathbb{X}))^V).$$

We also use the following common symbols:

$\mathbb{N} = \{1, 2, 3, \dots\}$ ,  $\mathbb{Z} = \{\dots - 2, -1, 0, 1, 2, \dots\}$ , and  $\mathbb{Z}^+ = \{0, 1, 2, \dots\}$ .  
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 Section 5 how these ideas can be generalized. The machinery and ideas introduced here are applied repeatedly later on in more elaborate settings.

Let

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

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

$$d(f_m(x), f_m(y)) \leq l \cdot d(x, y) \quad \forall x, y \in \mathbb{X}, \forall m \in \{1, 2, \dots, M\}.$$

The  $p_m$ 's are probabilities, with

$$\sum_{m=1}^M p_m = 1, p_m \geq 0 \quad \forall m.$$

We define mappings  $F : \mathbb{H}(\mathbb{X}) \rightarrow \mathbb{H}(\mathbb{X})$  and  $F : \mathbb{P}(\mathbb{X}) \rightarrow \mathbb{P}(\mathbb{X})$  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.** [16] *The mappings  $F : \mathbb{H}(\mathbb{X}) \rightarrow \mathbb{H}(\mathbb{X})$  and  $F : \mathbb{P}(\mathbb{X}) \rightarrow \mathbb{P}(\mathbb{X})$  are both contractions with factor  $0 \leq l < 1$ . That is,*

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

and

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

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

$$F(A) = A,$$

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

$$F(\mu) = \mu.$$

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, c.f. [8]. 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  $A_0 \in \mathbb{H}(\mathbb{X})$ , or  $\mu_0 \in \mathbb{P}(\mathbb{X})$ , and define recursively*

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

*respectively; then*

$$(2.2) \quad \lim_{k \rightarrow \infty} A_k = A, \text{ or } \lim_{k \rightarrow \infty} \mu_k = \mu,$$

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

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

In practical applications to two-dimensional computer graphics, the transformations and the spaces upon which they act must be discretized. The precise behaviour 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 [25], Chapter 4. The following example gives the flavour of such applications.

**Example 1.** In Figure 13 we illustrate a practical deterministic algorithm, based on the first formula in Equation (2.2) starting from a simple IFS on the unit square  $\square \subset \mathbb{R}^2$ . The IFS is  $F = \{\square; f_1^1, f_2^1, f_2^2; 0.36, 0.28, 0.36\}$  where the transformations are defined in Equations (1.1), (1.2), and (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 colours red and green. Each successive computed image is made of pixels belonging to a discrete model for  $\square$  and consists of red pixels 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 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 colour, 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 colour 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 colours themselves do not settle down, as illustrated in Figure 15. We call this “the texture effect”, and

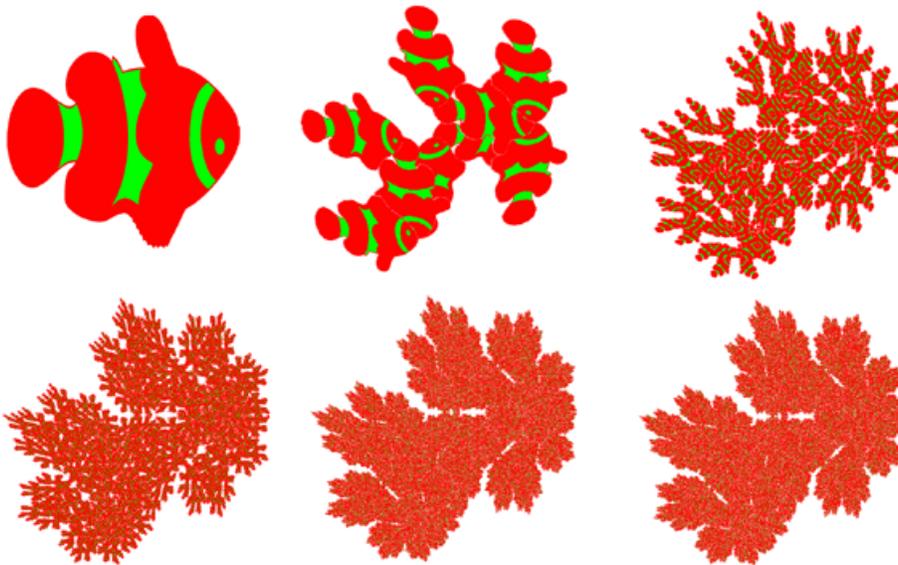


FIGURE 13. An illustration of the deterministic algorithm.

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 [10].

**Theorem 2.** *Specify a starting point  $x_1 \in \mathbb{X}$ . 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:*

$$(2.3) \quad \mu(B) = \lim_{l \rightarrow \infty} \frac{|\{B \cap \{x_1, x_2, \dots, x_l\}\}|}{l}.$$

for all  $B \in \mathbb{B}(\mathbb{X})$  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  $x_1 \in \mathbb{X}$  and almost all random orbits the sequence of point measures  $\frac{1}{l}(\delta_{x_1} + \delta_{x_2} + \dots + \delta_{x_l})$  converges in the weak sense to  $\mu$ , see for example [?], pages 11 and 12. (Weak convergence of probability measures is the same as convergence in the Monge Kantorovitch metric, see [9], pages 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 computed image — and it allows the efficient computation of zooms into small parts of an image. However, as in the case of deterministic algorithms, the images produced depend on the computational details of image resolution, the precision to which the points  $\{x_1, x_2, \dots, x_l\}$

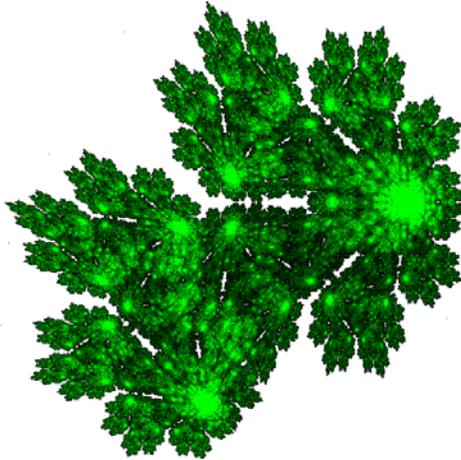


FIGURE 14. “Picture” of the measure attractor of an IFS with probabilities produced by the random iteration algorithm. The measure is depicted in shades of green, from 0 (black) to 255 (bright green).

are computed, the contractivity of the transformations, the way in which Equation (2.3) is implemented, choices of colours, etc. Different implementations can produce different results.

**Example 2.** Figure 14 shows a “picture” of the invariant measure of the IFS in Example 1, computed using a discrete implementation of the random iteration algorithm, as follows. Pixels corresponding to a discrete model for  $\square \subset \mathbb{R}^2$  are assigned the colour white. Successive floating point coordinates of points in  $\square$  are computed by random iteration and the first (say) one hundred points are discarded. Thereafter, as each new point is calculated, the pixel to which it belongs is set to black. This phase of the computation continues until the pixels cease to change, and produces a black image of the support of the measure, the set attractor of the IFS, against a white background. Then the random iteration process is continued, and as each new point is computed the green component of the pixel to which the latest point belongs is brightened by a fixed amount. Once a pixel is at brightest green, its value is not changed when later points belong to it. The computation is continued until a balance is obtained between that part of the image which is brightest green and that which is lightest green, and then stopped.

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 [10]. 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  $B \in \mathbb{B}(\mathbb{X})$  is such that  $f_m(B) \subset B \quad \forall m \in \{1, 2, \dots, M\}$ . 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  $E \subset B^C$ , for any  $m$ , since  $f_m(B) \subset B$ ,

$$f_m(\mu|_B)(E) = \mu(B \cap f_m^{-1}(E)) = \mu(\emptyset) = 0.$$

Moreover,

$$(2.4) \quad \mu(B) = f_m(\mu|_B)(\mathbb{X}) = f_m(\mu|_B)(B).$$

It follows that

$$\begin{aligned} \mu(B) &= \sum_{m=1}^M p_m f_m \mu(B) = \sum_{m=1}^M p_m f_m(\mu|_B)(B) + \sum_{m=1}^M p_m f_m(\mu|_{B^C})(B) \\ &= \mu(B) + \sum_{m=1}^M p_m f_m(\mu|_{B^C})(B) \quad (\text{from (2.4)}). \end{aligned}$$

Hence

$$(2.5) \quad \sum_{m=1}^M p_m f_m(\mu|_{B^C})(B) = 0.$$

Hence for any measurable set  $E \subset \mathbb{X}$

$$\begin{aligned} (\mu|_B)(E) &= \mu(B \cap E) = \sum_{m=1}^M p_m f_m \mu(B \cap E) \\ &= \sum_{m=1}^M p_m f_m(\mu|_B)(B \cap E) + \sum_{m=1}^M p_m f_m(\mu|_{B^C})(B \cap E) \\ &= \sum_{m=1}^M p_m f_m(\mu|_B)(E) + 0 \quad (\text{using (2.5)}). \end{aligned}$$

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  $\mu(B) = 1$ .  $\square$

**Example 3.** Figure 15 shows close-ups on the two images at the bottom left in Figure 13, see Example 1. At each iteration it is observed that the pattern of red and green pixels changes in a seemingly random manner. 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.)

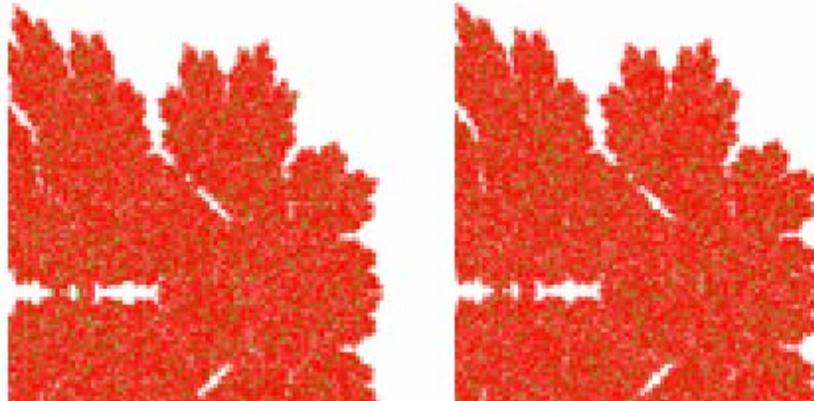


FIGURE 15. Close-up on the same small region in each of the bottom left two images in Figure 13, showing the texture effect; the distribution of red and green pixels changes with each iteration.

**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-Lipshitz transformation; that is, if  $f : \mathbb{X} \rightarrow \mathbb{X}$  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 modelling: 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 [12] pp. 25 et seq.

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

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

where  $|U_i|^s$  denotes the  $s^{\text{th}}$  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

$$H^s(S) = \lim_{\delta \rightarrow 0} H_\delta^s(S).$$

The  $s$ -dimensional Hausdorff measure is a Borel measure but is not normally even  $\sigma$ -finite.

**Definition 3.** *The Hausdorff dimension of the set  $S \subset \mathbb{X}$  is defined to be*

$$\dim_H S = \inf\{s \mid H^s(S) = 0\}.$$

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  $S \subset \mathbb{X}$  is defined to be*

$$\dim_B S = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(S)}{\log(1/\delta)}$$

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

$$f_m(O) \cap f_l(O) = \emptyset \text{ if } m \neq l.$$

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  $O_m$  is an orthonormal transformation on  $\mathbb{R}^K$ ,  $s_m \in (0, 1)$ , 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

$$(2.6) \quad \sum_{m=1}^M s_m^D = 1.$$

Moreover,

$$0 < H^D(A) < \infty.$$

*Proof.* This theorem, in essence, was first proved by Moran in 1946, [22]. A full proof is given in [12] p.118.  $\square$

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 Equation 2.6 is the unique value which makes  $(p_1, p_2, \dots, p_M)$  into a probability vector.

**2.3. Code Space.** A good way of looking at an IFS  $F$  as in (2.1) is in terms of the associated *code space*  $\Sigma = \{1, 2, \dots, M\}^\infty$ . Members of  $\Sigma$  are infinite sequences from the alphabet  $\{1, 2, \dots, M\}$  and indexed by  $\mathbb{N}$ . We equip  $\Sigma$  with the metric  $d_\Sigma$  defined for  $\omega \neq \varkappa$  by

$$d_\Sigma(\omega, \varkappa) = \frac{1}{M^k},$$

where  $k$  is the index of the first symbol at which  $\omega$  and  $\varkappa$  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  $F : \Sigma \rightarrow A$ , defined for all  $\sigma_1\sigma_2\sigma_3\dots \in \Sigma$  by*

$$F(\sigma_1\sigma_2\sigma_3\dots) = \lim_{k \rightarrow \infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \dots \circ f_{\sigma_k}(x).$$

*The limit is independent of  $x \in \mathbb{X}$  and the convergence is uniform in  $x$ .*

*Proof.* This result is contained in [16] Theorem 3.1(3). □

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

Note that  $F : \Sigma \rightarrow A$  is not in general one-to-one.

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

**Theorem 6.** *For each  $m \in \{1, 2, \dots, M\}$  define the shift operator  $s_m : \Sigma \rightarrow \Sigma$  by*

$$s_m(\sigma_1\sigma_2\sigma_3\dots) = m\sigma_1\sigma_2\sigma_3\dots$$

*$\forall \sigma_1\sigma_2\sigma_3\dots \in \Sigma$ . Then  $s_m$  is a contraction mapping with contractivity factor  $\frac{1}{M}$ . Consequently*

$$S := \{\Sigma; s_1, s_2, \dots, s_M; p_1, p_2, \dots, p_M\}$$

*is an IFS. Its set attractor is  $\Sigma$ . Its measure attractor is the unique measure  $\pi \in \mathbb{P}(\Sigma)$  such that*

$$\pi\{\omega_1\omega_2\omega_3\dots \in \Sigma \mid \omega_1 = \sigma_1, \omega_2 = \sigma_2, \dots, \omega_k = \sigma_k\} = p_{\sigma_1} \cdot p_{\sigma_2} \cdot \dots \cdot p_{\sigma_k}$$

*$\forall k \in \mathbb{N}, \forall \sigma_1, \sigma_2, \dots, \sigma_k \in \{1, 2, \dots, M\}$ .*

*If  $\mu$  is the measure attractor of the IFS  $F$ , with  $F : \Sigma \rightarrow A$  defined as in Theorem 5, then*

$$\mu = F(\pi).$$

*Proof.* This result is [16] 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 [7], and results can often be lifted to the IFS  $F$ . For example, when the IFS is non-overlapping, the entropy (see [5] for the definition) of the stationary stochastic process associated with  $F$  is the same as that associated with the corresponding shift IFS, namely:  $-\sum p_m \log p_m$ .

## 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  $n \in \{1, 2, \dots, N\}$  let  $F^n$  denote the IFS

$$F^n = \{\mathbb{X}; f_1^n, f_2^n, \dots, f_M^n; p_1^n, p_2^n, \dots, p_M^n\}$$

where each  $f_m^n : X \rightarrow X$  is a Lipschitz function with Lipschitz constant  $0 \leq l < 1$  and the  $p_m^n$ 's are probabilities with

$$\sum_{m=1}^M p_m^n = 1, p_m^n \geq 0 \quad \forall m, n.$$

Let

$$(3.1) \quad \mathcal{F} = \{\mathbb{X}; F^1, F^2, \dots, F^N; P_1, P_2, \dots, P_N\},$$

where the  $P_n$ 's are probabilities with

$$\sum_{n=1}^N P_n = 1, P_n \geq 0 \quad \forall n \in \{1, 2, \dots, N\}.$$

$P_n > 0 \quad \forall n \in \{1, 2, \dots, N\}$  and  $\sum_{n=1}^N P_n = 1$ .

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

**3.2. Trees.** We associate various collections of trees with  $\mathcal{F}$  and the parameters  $M$  and  $N$ .

Let  $T$  denote the ( $M$ -fold) set of finite sequences from  $\{1, 2, \dots, M\}$ , including the empty sequence  $\emptyset$ . Then  $T$  is called a *tree* and the sequences are called the *nodes* of the tree. For  $i = i_1 i_2 \dots i_k \in T$  let  $|i| = k$ . The number  $k$  is called the *level* of the *node*  $\sigma$ . The bottom node  $\emptyset$  is at level zero. If  $j = j_1 j_2 \dots j_l \in T$  then  $ij$  is the concatenated sequence  $i_1 i_2 \dots i_k j_1 j_2 \dots j_l$ .

We define a *level- $k$*  ( $M$ -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 *labelled 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  $m \in \{1, 2, \dots, M\}$ , or the pair of nodes  $(\emptyset, \emptyset)$ , which is also called the *trunk*. In representations of labelled trees, as in Figures 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, or possibly to both the nodes and limbs when a labelled tree is multivalued. For a two-valued labelled tree  $\tau$  we will write

$$\tau(i) = (\tau(\text{node } i), \tau(\text{limb } i)) \text{ for } i \in T,$$

to denote the two components.

A *code tree* is a labelled tree whose range is  $\{1, 2, \dots, N\}$ . We write

$$\Omega = \{\tau \mid \tau : T \rightarrow \{1, 2, \dots, N\}\}$$

for the set of all infinite code trees.

We define the *subtree*  $\tilde{\tau} : T \rightarrow \{1, 2, \dots, N\}$  of a labelled tree  $\tau : T \rightarrow \{1, 2, \dots, N\}$ , corresponding to a node  $i = i_1 i_2 \dots i_k \in T$ , by

$$\tilde{\tau}(j) = \tau(ij) \quad \forall j \in T.$$

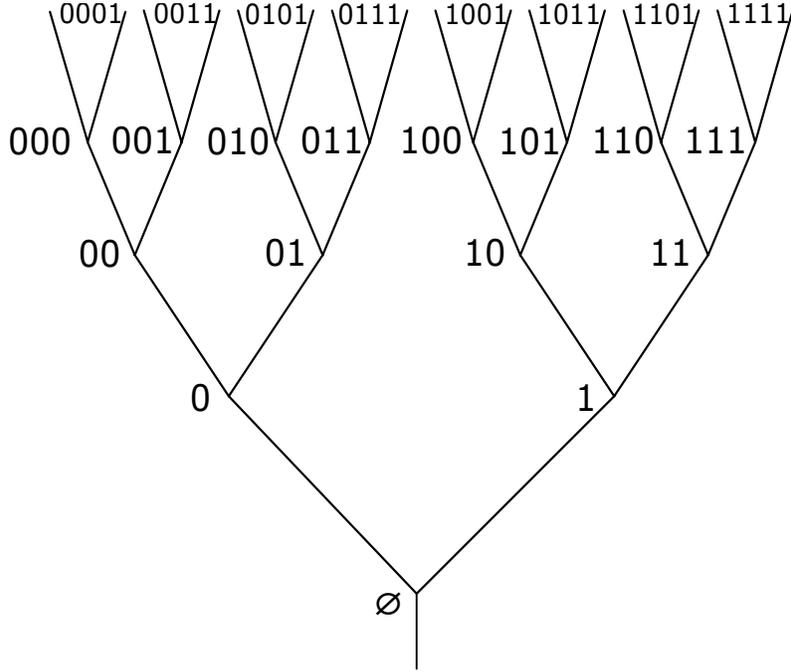


FIGURE 16. Pictorial representation of a level-4 2-fold tree labelled by the sequences corresponding to its nodes. The labels on the fourth level are shown for every second node. The line segments between the nodes and the line segment below the bottom node are referred to as *limbs*. The bottom limb is also called the *trunk*.

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 labelled trees with  $|\tau| \leq |\sigma|$ , where we allow  $|\sigma| = \infty$ . 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

$$\tau \prec \sigma.$$

If  $\tau$  is a level- $k$  code tree, the corresponding *cylinder set* is defined by

$$[\tau] = [\tau]_{\Omega} := \{\sigma \in \Omega : \tau \prec \sigma\}.$$

We define a metric on  $\Omega$  by, for  $\omega \neq \varkappa$ ,

$$d_{\Omega}(\omega, \varkappa) = \frac{1}{M^k}$$

if  $k$  is the least integer such that  $\omega(i) \neq \varkappa(i)$  for some  $i \in T$  with  $|i| = k$ . Then  $(\Omega, d_{\Omega})$  is a compact metric space. Furthermore,

$$\text{diam}(\Omega) = 1,$$

and

$$(3.2) \quad \text{diam}([\tau]) = \frac{1}{M^{k+1}}$$

whenever  $\tau$  is a level- $k$  code tree.

The probabilities  $P_n$  associated with  $\mathcal{F}$  in Equation (3.1) induce a natural probability distribution

$$\rho \in \mathbb{P}(\Omega)$$

on  $\Omega$ . It is defined on cylinder sets  $[\tau]$  by

$$(3.3) \quad \rho([\tau]) = \prod_{1 \leq |i| \leq |\tau|} P_{\tau(i)}.$$

That is, the random variables  $\tau(i)$ , with nodal values in  $\{1, 2, \dots, N\}$ , are chosen i.i.d. via  $\Pr(\tau(i) = n) = P_n$ . Then  $\rho$  is extended in the usual way to the  $\sigma$ -algebra  $\mathbb{B}(\Omega)$  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  $\mathcal{F}$  is a code tree wherein the symbols  $1, 2, \dots$ , and  $N$  are replaced by the respective IFSSs  $F^1, F^2, \dots$ , and  $F^N$ . A construction tree consists of nodes and limbs, where each node is labelled by one of the IFSSs belonging to  $\mathcal{F}$ . 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  $\mathcal{F}(\Omega)$  for the set of construction trees for  $\mathcal{F}$ . For  $\sigma \in \Omega$  we write  $\mathcal{F}(\sigma)$  to denote the corresponding construction tree. We will use the *same notation*  $\mathcal{F}(\sigma)$  to denote the random fractal set associated with the construction tree  $\mathcal{F}(\sigma)$ , as described in the next section.

**3.3. Random Fractals.** In this section we describe the canonical random fractal sets and measures associated with  $\mathcal{F}$  in (3.1).

Let  $\mathcal{F}$  be given as in (3.1), let  $k \in \mathbb{N}$ , and define

$$\mathcal{F}_k : \Omega \times \mathbb{H}(\mathbb{X}) \rightarrow \mathbb{H}(\mathbb{X}),$$

by

$$(3.4) \quad \mathcal{F}_k(\sigma)(K) = \bigcup_{\{i \in T \mid |i|=k\}} f_{i_1}^{\sigma(\emptyset)} \circ f_{i_2}^{\sigma(i_1)} \circ \dots \circ f_{i_k}^{\sigma(i_1 i_2 \dots i_{k-1})}(K)$$

$\forall \sigma \in \Omega$  and  $K \in \mathbb{H}(\mathbb{X})$ . (The set  $\mathcal{F}_k(\sigma)(K)$  is obtained by taking the union of the compositions of the functions occurring on the branches of the construction tree  $\mathcal{F}(\sigma)$  starting at the bottom and working up to the  $k^{\text{th}}$  level, acting upon  $K$ .)

In a similar way, with measures in place of sets, and unions of sets replaced by sums of measures weighed by probabilities, we define

$$\tilde{\mathcal{F}}_k : \Omega \times \mathbb{P}(\mathbb{X}) \rightarrow \mathbb{P}(\mathbb{X})$$

by

$$(3.5) \quad \tilde{\mathcal{F}}_k(\sigma)(\varsigma) = \sum_{\{i \in T \mid |i|=k\}} (p_{i_1}^{\sigma(\emptyset)} \cdot p_{i_2}^{\sigma(i_1)} \cdot \dots \cdot p_{i_k}^{\sigma(i_1 i_2 \dots i_{k-1})}) \tilde{f}_{i_1}^{\sigma(\emptyset)} \circ \tilde{f}_{i_2}^{\sigma(i_1)} \circ \dots \circ \tilde{f}_{i_k}^{\sigma(i_1 i_2 \dots i_{k-1})}(\varsigma)$$

$\forall \sigma \in \Omega$  and  $\varsigma \in \mathbb{P}(\mathbb{X})$ . Note that the  $\tilde{\mathcal{F}}_k(\sigma)(\varsigma)$  all have unit mass because the  $p_m^n$  sum (over  $m$ ) to unity.

**Theorem 7.** *Let sequences of functions  $\{\mathcal{F}_k\}$  and  $\{\tilde{\mathcal{F}}_k\}$  be defined as above. Then both the limits*

$$\mathcal{F}(\sigma) = \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\sigma)(K)\}, \text{ and } \tilde{\mathcal{F}}(\sigma) = \lim_{k \rightarrow \infty} \{\tilde{\mathcal{F}}_k(\sigma)(\varsigma)\},$$

exist, are independent of  $K$  and  $\varsigma$ , and the convergence (in the Hausdorff and Monge Kantorovitch metrics, respectively,) is uniform in  $\sigma$ ,  $K$ , and  $\varsigma$ . The resulting functions

$$\mathcal{F} : \Omega \rightarrow \mathbb{H}(\mathbb{X}) \text{ and } \tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P}(\mathbb{X})$$

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.  $\square$

Let

$$(3.6) \quad \mathfrak{H} = \{\mathcal{F}(\sigma) \in \mathbb{H}(\mathbb{X}) | \sigma \in \Omega\}, \text{ and } \tilde{\mathfrak{H}} = \{\tilde{\mathcal{F}}(\sigma) \in \mathbb{P}(\mathbb{X}) | \sigma \in \Omega\}.$$

Similarly let

$$(3.7) \quad \mathfrak{P} = \mathcal{F}(\rho) = \rho \circ \mathcal{F}^{-1} \in \mathbb{P}(\mathfrak{H}), \text{ and } \tilde{\mathfrak{P}} = \tilde{\mathcal{F}}(\rho) = \rho \circ \tilde{\mathcal{F}}^{-1} \in \mathbb{P}(\tilde{\mathfrak{H}}).$$

**Definition 7.** *The sets  $\mathfrak{H}$  and  $\tilde{\mathfrak{H}}$  are called the sets of fractal sets and fractal measures, respectively, associated with  $\mathcal{F}$ . These random fractal sets and measures are said to be distributed according to  $\mathfrak{P}$  and  $\tilde{\mathfrak{P}}$ , 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  $\{F^n : n = 1, 2, \dots, N\}$  and the superIFS  $\mathcal{F}$  are said to obey the (uniform) open set condition if there exists a non-empty open set  $O$  such that for each  $n \in \{1, 2, \dots, N\}$*

$$F^n(O) = \bigcup_{m=1}^M f_m^n(O) \subset O,$$

and

$$f_m^n(O) \cap f_l^n(O) = \emptyset \quad \forall m, l \in \{1, 2, \dots, M\} \text{ with } m \neq l.$$

For the rest of this section we restrict attention to  $(\mathbb{X}, d)$  where  $\mathbb{X} \subseteq \mathbb{R}^K$  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, see [11], [14] and [21].

**Theorem 8.** *Let the set of IFSs  $\{F^n : n = 1, 2, \dots, N\}$  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  $n \in \{1, 2, \dots, N\}$  and  $m \in \{1, 2, \dots, M\}$ . Also let  $\{F^n : n = 1, 2, \dots, N\}$  obey the uniform open set condition. Then for  $\mathfrak{P}$ -almost all  $A \in \mathfrak{H}$*

$$\dim_H A = \dim_B A = D$$

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 [12] 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

$$\Omega^V = \underbrace{\Omega \times \Omega \times \dots \times \Omega}_{V \text{ TIMES}}$$

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  $\Omega_V$ : 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  $\{1, 2, \dots, V\}$ .

One reason that we are interested in  $\Omega_V$  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  $a \in \mathcal{A}$  where

$$(4.1) \quad \mathcal{A} := \{\{1, 2, \dots, N\} \times \{1, 2, \dots, V\}^M\}^V.$$

A typical index  $a \in \mathcal{A}$  will be denoted

$$(4.2) \quad a = (a_1, a_2, \dots, a_V)$$

where

$$(4.3) \quad a_v = (n_v; v_{v,1}, v_{v,2}, \dots, v_{v,M})$$

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

Specifically, algebraically, the mapping  $\eta^a$  is defined in Equation (4.8) below. But it is very useful to represent the indices and the mappings with trees. See Figure 17. Each map  $\eta^a$  in Equation (4.8) and each index  $a \in \mathcal{A}$  may be represented by a  $V$ -tuple of labelled 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 labelled by the index of the component  $v \in \{1, 2, \dots, V\}$  to which it corresponds. The node of each function tree is labelled by the IFS number  $n_v$  (shown circled) of the corresponding component of the mapping. The  $m^{\text{th}}$  limb of the  $v^{\text{th}}$  tree is labelled by the number  $v_{v,m} \in \{1, 2, \dots, V\}$ , for  $m \in \{1, 2, \dots, M\}$ . 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  $a \in \mathcal{A}$  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 Figure 18 an arbitrary grove  $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega^3$  is represented by a triple of coloured tree pictures, one blue, one orange, and one magenta, with trunks labelled one, two, and three respectively. The top left of Figure 18 shows the map  $\eta^a$ , and the index  $a \in \mathcal{A}$ , where

$$(4.4) \quad \eta^a(\omega_1, \omega_2, \omega_3) = (\xi_1(\omega_1, \omega_2), \xi_5(\omega_3, \omega_2), \xi_4(\omega_3, \omega_1)),$$

and

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

represented by function trees. The functions  $\{\xi_n : n = 1, 2, \dots, 5\}$  are defined in Equation (4.7) below. The result of the action of  $\eta^a$  on  $\omega$  is represented, in the bottom part of Figure 18, by a grove whose lowest nodes are labelled by the IFS

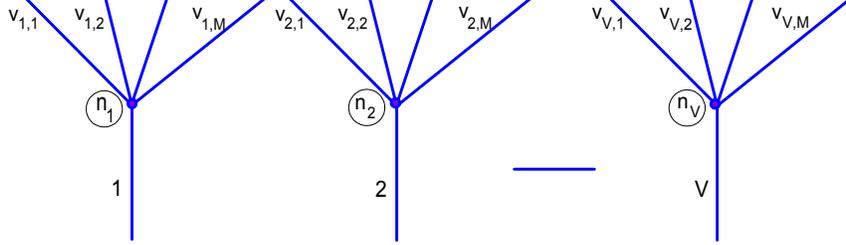


FIGURE 17. Each map  $\eta^a$  in Equation (4.8) and each index  $a \in \mathcal{A}$  may be represented by a  $V$ -tuple of labelled level-1 trees that we call (*level-1*) *function trees*. Each function tree has a trunk, a node, and  $M$  limbs. The function trees correspond to the components of the mapping. The trunk of each function tree is labelled by the index of the component  $v \in \{1, 2, \dots, V\}$  to which it corresponds. The node of each function tree is labelled by the IFS number  $n_v$  (shown circled) of the corresponding component of the mapping. The  $m^{\text{th}}$  limb of the  $v^{\text{th}}$  tree is labelled by the domain number  $v_{v,m} \in \{1, 2, \dots, V\}$ , for  $m \in \{1, 2, \dots, M\}$ .

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 Figure 17 on a grove  $\omega \in \Omega^V$  (represented by  $V$  trees with trunks labelled 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.

We are also going to need a set of probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$ , with

$$(4.5) \quad \sum_{a \in \mathcal{A}} \mathcal{P}^a = 1, \mathcal{P}^a \geq 0 \forall a \in \mathcal{A}.$$

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

$$(4.6) \quad \mathcal{P}^a = \frac{P_{n_1} P_{n_2} \dots P_{n_V}}{V^{MV}},$$

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

or each  $n \in \{1, 2, \dots, N\}$  define the  $n^{\text{th}}$  shift mapping  $\xi_n : \Omega^M \rightarrow \Omega$  by

$$(4.7) \quad (\xi_n(\omega))(\emptyset) = n \text{ and } (\xi_n(\omega))(mi) = \omega_m(i) \forall i \in T, m \in \{1, 2, \dots, M\},$$

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

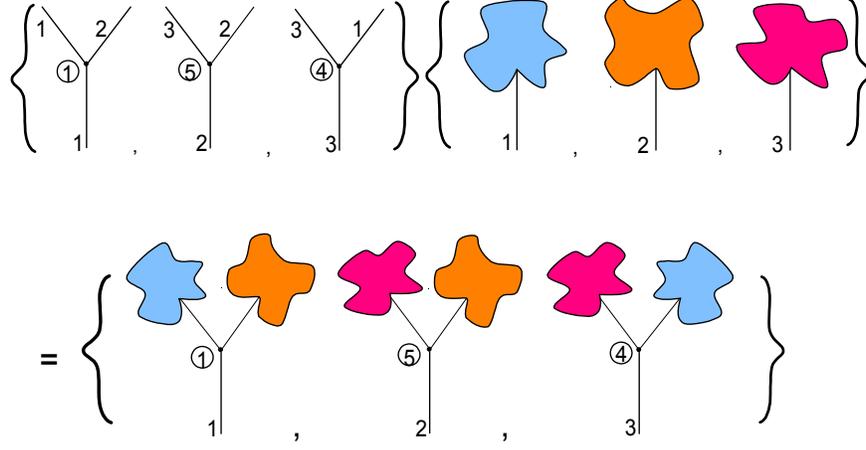


FIGURE 18. The top right portion of the image represents a grove  $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega^3$  by a triple of coloured tree pictures. The top left portion represents the map  $\eta^a$  in Equation (4.4) using level-1 function trees. The bottom portion represents the image  $\eta^a(\omega)$ .

**Theorem 9.** For each  $a \in \mathcal{A}$  define  $\eta^a : \Omega^V \rightarrow \Omega^V$  by

$$(4.8) \quad \eta^a(\omega_1, \omega_2, \dots, \omega_V) := (\xi_{n_1}(\omega_{v_{1,1}}, \omega_{v_{1,2}}, \dots, \omega_{v_{1,M}}), \\ \xi_{n_2}(\omega_{v_{2,1}}, \omega_{v_{2,2}}, \dots, \omega_{v_{2,M}}), \dots, \xi_{n_V}(\omega_{v_{V,1}}, \omega_{v_{V,2}}, \dots, \omega_{v_{V,M}}))$$

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

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

$$\begin{aligned} d_{\Omega^V}(\eta^a(\omega_1, \omega_2, \dots, \omega_V), \eta^a(\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_V)) &= \\ \max_{v \in \{1, 2, \dots, V\}} d_{\Omega}(\xi_{n_v}(\omega_{v_{v,1}}, \omega_{v_{v,2}}, \dots, \omega_{v_{v,M}}), \xi_{n_v}(\tilde{\omega}_{v_{v,1}}, \tilde{\omega}_{v_{v,2}}, \dots, \tilde{\omega}_{v_{v,M}})) & \\ \leq \frac{1}{M} \cdot \max_{v \in \{1, 2, \dots, V\}} d_{\Omega^M}((\omega_{v_{v,1}}, \omega_{v_{v,2}}, \dots, \omega_{v_{v,M}}), (\tilde{\omega}_{v_{v,1}}, \tilde{\omega}_{v_{v,2}}, \dots, \tilde{\omega}_{v_{v,M}})) & \\ \leq \frac{1}{M} \cdot \max_{v \in \{1, 2, \dots, V\}} d_{\Omega}(\omega_v, \tilde{\omega}_v) = \frac{1}{M} \cdot d_{\Omega^V}((\omega_1, \omega_2, \dots, \omega_V), (\tilde{\omega}_1, \tilde{\omega}_2, \dots, \tilde{\omega}_V)). & \end{aligned}$$

□

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

$$(4.9) \quad \Phi = \{\Omega^V; \eta^a, \mathcal{P}^a, a \in \mathcal{A}\}.$$

Let the set attractor and the measure attractor of the IFS  $\Phi$  be denoted by  $\Omega_V$  and  $\mu_V$  respectively. Clearly,  $\Omega_V \in \mathbb{H}(\Omega^V)$  while  $\mu_V \in \mathbb{P}(\Omega^V)$ . We call  $\Omega_V$  the space of  $V$ -groves. The elements of  $\Omega_V$  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  $v \in \{1, 2, \dots, V\}$ , let us define  $\Omega_{V,v} \subset \Omega_V$  to be the set of  $v^{\text{th}}$  components of groves in  $\Omega_V$ . Also let  $\rho_V \in \mathbb{P}(\Omega)$  denote the marginal probability measure defined by

$$(4.10) \quad \rho_V(B) := \mu_V(B, \Omega, \Omega, \dots, \Omega) \forall B \in \mathbb{B}(\Omega).$$

**Theorem 10.** *For all  $v \in \{1, 2, \dots, V\}$  we have*

$$(4.11) \quad \Omega_{V,v} = \Omega_{V,1} := \{\text{set of all } V\text{-trees}\}.$$

When the probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  obey Equation (4.6), then starting at any initial grove, the random distribution of trees  $\omega \in \Omega$  that occur in the  $v^{\text{th}}$  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  $n \rightarrow \infty$ .

*Proof.* Let  $\Xi : \Omega^V \rightarrow \Omega^V$  denote any map that permutes the coordinates. Then the IFS  $\Phi = \{\Omega^V; \eta^a, \mathcal{P}^a, a \in \mathcal{A}\}$  is invariant under  $\Xi$ , that is  $\Phi = \{\Omega^V; \Xi\eta^a\Xi^{-1}, \mathcal{P}^a, a \in \mathcal{A}\}$ . It follows that  $\Xi\Omega_V = \Omega_V$  and  $\Xi\mu_V = \mu_V$ . It follows that Equation (4.11) holds, and also that, in the obvious notation, for any  $(B_1, B_2, \dots, B_V) \in (\mathbb{B}(\Omega))^V$  we have

$$(4.12) \quad \mu_V(B_1, B_2, \dots, B_V) = \mu_V(B_{\sigma_1}, B_{\sigma_2}, \dots, B_{\sigma_V}).$$

In particular

$$\rho_V(B) = \mu_V(B, \Omega, \Omega, \dots, \Omega) = \mu_V(\Omega, \Omega, \dots, \Omega, B, \Omega, \dots, \Omega) \forall B \in \mathbb{B}(\Omega),$$

where the “ $B$ ” on the right-hand-side is in the  $v^{\text{th}}$  position. 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  $\{\mu_V^{(l)} : l = 1, 2, 3, \dots\}$ , that converge weakly to  $\mu_V$  almost always. In particular  $\mu_V^{(l)}(B, \Omega, \Omega, \dots, \Omega)$  converges to  $\rho_V$  almost always.  $\square$

Let  $L$  denote a set of fixed length vectors of labelled 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  $\omega \in L$ , 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.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

$$S = \{\omega \in \Omega^V \mid |\{\text{subtrees of components of } \omega \text{ at level } k\}| \leq V, \forall k \in \mathbb{N}\}.$$

Then  $S$  is closed: Let  $\{s_n \in S\}$  converge 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.  $d_{\Omega^V}(s_n, s^*) < \frac{1}{M^{(k+l)}}$ ). 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  $(\sigma, \sigma, \dots, \sigma) \in S$ .

Also  $S$  is invariant under the IFS  $\Phi$ : Clearly any  $s \in S$  can be written  $s = \eta^a(\tilde{s})$  for some  $a \in \mathcal{A}$  and  $\tilde{s} \in S$ . Also, if  $s \in S$  then  $\eta^a(s) \in S$ . So  $S = \cup\{\eta^a(S) \mid a \in \mathcal{A}\}$ .

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  $(\sigma, \sigma, \dots, \sigma) \in \Omega_V$ , and so  $\sigma \in \Omega_{V,1}$ .  $\square$

**Theorem 12.** *For all*

$$d_{\mathbb{H}(\Omega)}(\Omega_{V,1}, \Omega) < \frac{1}{V},$$

which implies that

$$\lim_{V \rightarrow \infty} \Omega_{V,1} = \Omega,$$

where the convergence is in the Hausdorff metric. Let the probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  obey Equation (4.6). Then

$$(4.13) \quad d_{\mathbb{P}(\Omega)}(\rho_V, \rho) \leq 1.4 \left( \frac{M}{V} \right)^{\frac{1}{4}}$$

which implies

$$\lim_{V \rightarrow \infty} \rho_V = \rho,$$

where  $\rho$  is the stationary measure on trees introduced in Section 3.2, and convergence is in the Monge Kantorovitch metric.

*Proof.* To prove the first part, let  $M^{k+1} > V \geq M^k$ . Let  $\tau$  be any level- $k$  code tree. Then  $\tau$  is clearly  $V$ -variable and it can be extended to an infinite  $V$ -variable code tree. It follows that  $[\tau] \cap \Omega_{V,1} \neq \emptyset$ . The collection of such cylinder sets  $[\tau]$  forms a disjoint partition of  $\Omega$  by subsets of diameter  $\frac{1}{M^{k+1}}$ , see Equation (3.2)), from which it follows that

$$d_{\mathbb{H}(\Omega)}(\Omega_{V,1}, \Omega) \leq \frac{1}{M^{k+1}} < \frac{1}{V}.$$

The first part of the theorem follows at once.

For the proof of the second part, we refer to Section 4.4.  $\square$

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

$$\Phi(a_1 a_2 a_3 \dots) = \lim_{k \rightarrow \infty} \eta^{a_1} \circ \eta^{a_2} \circ \dots \circ \eta^{a_k}(\omega)$$

for all  $a_1 a_2 a_3 \dots \in \Sigma_V$ , for any  $\omega$ . In the terminology of section 2.3 the sequence  $a_1 a_2 a_3 \dots \in \Sigma_V$  is an address of the  $V$ -grove  $\Phi(a_1 a_2 a_3 \dots) \in \Omega_V$  and  $\Sigma_V$  is the code space for the set of  $V$ -groves  $\Omega_V$ . In general  $\Phi : \Sigma_V \rightarrow \Omega_V$  is not one-to-one, as we will see in Section 4.2.

**4.2. Compositions of the Mappings  $\eta^a$ .** Compositions of the mappings  $\eta^a : \Omega^V \rightarrow \Omega^V$ ,  $a \in \mathcal{A}$ , represented by  $V$ -tuples of level-1 function trees, as in Figure 17, can be represented by higher level trees that we call *level- $k$  function trees*.

First we illustrate the ideas, then we formalize. In Figure 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

$$\begin{aligned} a &= ((1; 2, 3), (3; 1, 3), (5; 2, 3)), \\ b &= ((4; 3, 1), (2; 1, 2), (3; 3, 2)), \end{aligned}$$

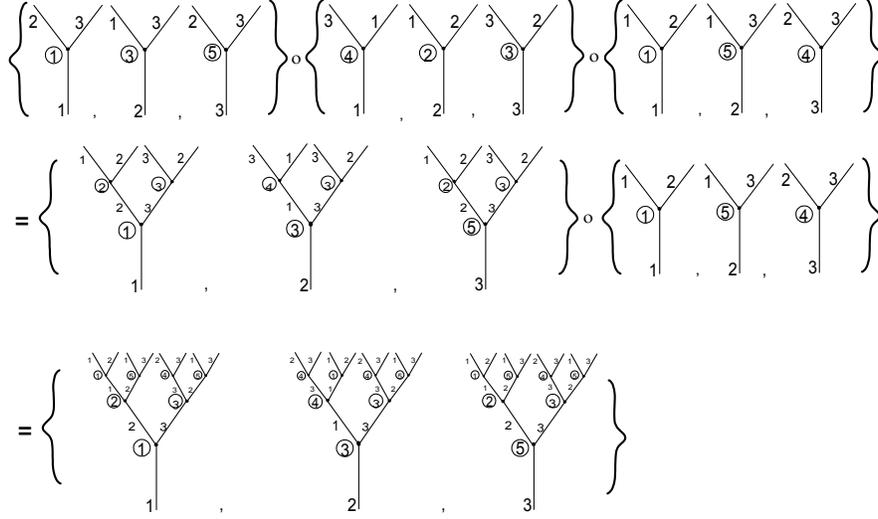


FIGURE 19. Illustrations of compositions of function trees to produce higher level function trees. Here  $V = 3$ ,  $N = 5$ , and  $M = 2$ . We compose the level-1 function trees corresponding to  $a = ((1; 2, 3), (3; 1, 3), (5; 2, 3))$ ,  $b = ((4; 3, 1), (2; 1, 2), (3; 3, 2))$ , and  $c = ((1; 1, 2), (5; 1, 3), (4; 2, 3))$ . 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  $a \circ b \circ c$ .

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 Figure 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)  $a \circ b \circ c$  can be converted into a mapping  $\eta^{a \circ b \circ c} : \Omega^3 \rightarrow \Omega^3$  and (ii)

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

The mapping  $\eta^{a \circ b \circ c}(\omega)$  is defined to be the 3-tuple of code trees obtained by attaching the tree  $\omega_v$  to each of the top limbs of each level-3 function tree in  $a \circ b \circ c$  with label  $v$  for all  $v \in \{1, 2, 3\}$  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- $k$  labelled tree with the nodes of the first  $k - 1$  levels labelled from  $\{1, 2, \dots, N\}$  and limbs (i.e. second nodal values) of all  $k$  levels labelled from  $\{1, 2, \dots, V\}$ . We define a *grove* of level- $k$  function trees, say  $g$ , to be a  $V$ -tuple of level- $k$  function trees, with trunks labelled according to the component number, and we define  $G_k$  to be the set of such  $g$ . Let  $G := \bigcup_{k \in \mathbb{N}} G_k$ . We will refer to a component of an element of

$G$  simply as a *function tree*. For  $g \in G$  we will write  $|g| = k$ , where  $k \in \mathbb{N}$  is the unique number such that  $g \in G_k$ .

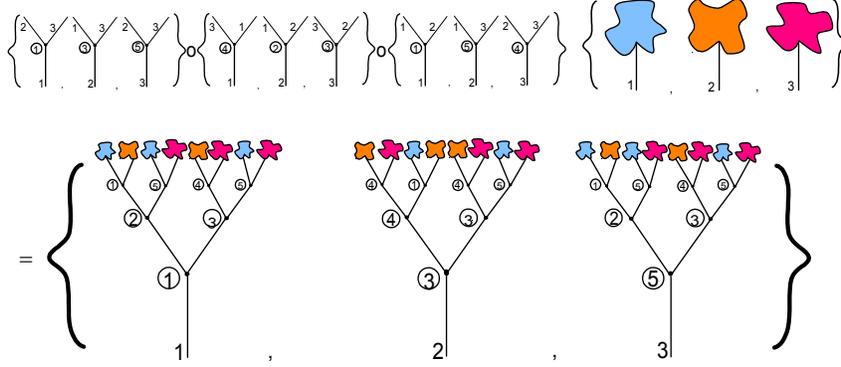


FIGURE 20. Illustration of the composition of the three mappings,  $\eta^a \circ \eta^b \circ \eta^c = \eta^{a \circ b \circ c}$  applied to  $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega^3$ . See also Figure 19 and the text.

Then, for all  $g = (g_1, g_2, \dots, g_V) \in G$ , for all  $v \in \{1, 2, \dots, V\}$ ,

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

and

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

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.

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

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

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

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

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

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

and

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

This is consistent with the definition of  $\eta^a : \Omega^V \rightarrow \Omega^V$  for  $a \in \mathcal{A}$ , 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

$$(4.14) \quad \eta^{g \circ h} = \eta^g \circ \eta^h.$$

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

$$(4.15) \quad \eta^{a_1} \circ \eta^{a_2} \circ \dots \circ \eta^{a_k} = \eta^{a_1 \circ a_2 \circ \dots \circ a_k}.$$

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

We remark that Equations (4.15) and (4.14) allow us to work directly with function trees to construct, count, and track compositions of mappings  $\eta^a \in \mathcal{A}$ . 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  $a_1 \circ a_2 \circ \dots \circ a_k$  for some  $\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathcal{A}^k$ . (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  $[(\eta^{a_1 \circ a_2 \circ \dots \circ a_k})_1]$  by computing the probabilities of occurrence of function trees  $g \in G_k(V)$  such that  $(\eta^g)_1 = (\eta^{a_1 \circ a_2 \circ \dots \circ a_k})_1$  built up starting from level-0 trees, with probabilities given by Equation (4.6), as we do in the Section 4.4.

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

The grove of code trees for  $\Phi(a_1 a_2 \dots)$  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  $\rho_V$ .** Let  $\{\mathbf{F}_n\}_{n=0}^\infty = \{F_1^n, \dots, F_V^n\}_{n=0}^\infty$  be a sequence of random groves of level-1 function trees. Each random function tree can be expressed as  $F_v^n = (N_v^n, L_v^n(1), \dots, L_v^n(M))$  where the  $N_v^n$ 's and  $L_v^n$ 's correspond to random labellings of nodes and limbs respectively.

We assume that the function trees  $\{F_v^n\}$ , are independent with  $Pr(N_v^n = j) = P_j$  and  $Pr(L_v^n(m) = k) = 1/V$  for any  $v, k \in \{1, \dots, V\}$ ,  $n \in \mathbb{N}$ ,  $j \in \{1, \dots, N\}$ , and  $m \in \{1, \dots, M\}$ .

First the family  $\{L_1^n, \dots, L_V^n\}$  generates a random dependence tree,  $K : T \rightarrow \{1, \dots, V\}$ , in the following way. Let  $K(\emptyset) = 1$ . If  $K(i_1, \dots, i_n) = j$ , for some  $j = 1, \dots, V$ , then we define  $K(i_1, \dots, i_n, i_{n+1}) = L_j^n(i_{n+1})$ .

Given the dependence tree, let  $I_i = N_{K(i)}^n$  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  $(\mathbf{F}_0 \circ \mathbf{F}_1 \circ \dots \circ \mathbf{F}_{k-1})_1$  is a level- $k$  function tree with nodes given by  $\{I_i\}_{|i| \leq k-1}$ , and limbs given by  $\{K(i)\}_{|i| \leq k}$ .

Thus

$$(4.16) \quad \begin{aligned} & Pr(I_i = \tau(i), \forall i \text{ with } |i| \leq k-1) \\ & = Pr((\mathbf{F}_0 \circ \mathbf{F}_1 \circ \dots \circ \mathbf{F}_{k-1})_1(\text{node } i) = \tau(i), \forall i \text{ with } |i| \leq k-1). \end{aligned}$$

For

$$\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathcal{A}^k.$$

let

$$\mathcal{P}^{\mathbf{a}} = \mathcal{P}^{a_1} \mathcal{P}^{a_2} \dots \mathcal{P}^{a_k}$$

denote the probability of selection of

$$\eta^{\mathbf{a}} = \eta^{a_1} \circ \eta^{a_2} \circ \dots \circ \eta^{a_k} = \eta^{a_1 \circ a_2 \circ \dots \circ a_k}.$$

By the invariance of  $\mu_V$

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

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

$$\begin{aligned} \rho_V([\tau]) &= \mu_V([\tau], \Omega, \dots, \Omega) = \sum_{\mathbf{a} \in \mathcal{A}^k} \mathcal{P}^{\mathbf{a}} \mu_V((\eta^{\mathbf{a}})^{-1}([\tau], \Omega, \dots, \Omega)) \\ &= \sum_{\{\mathbf{a} \in \mathcal{A}^k \mid (\eta^{\mathbf{a}})_1(\text{node } i) = \tau(i), \forall i\}} \mathcal{P}^{\mathbf{a}}. \end{aligned}$$

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

#### 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 dependence tree of  $(\mathbf{F}_0 \circ \mathbf{F}_1 \circ \dots \circ \mathbf{F}_{k-1})_1$  is free and  $V \geq M^k$ , then

$$\begin{aligned} \rho_V([\tau]) &= \mu_V([\tau], \Omega, \dots, \Omega) = \sum_{\mathbf{a} \in \mathcal{A}^k} \mathcal{P}^{\mathbf{a}} \mu_V((\eta^{\mathbf{a}})^{-1}([\tau], \Omega, \dots, \Omega)) \\ \Pr(F) &= \prod_{i=1}^{M-1} \left(1 - \frac{i}{V}\right) \prod_{i=1}^{M^2-1} \left(1 - \frac{i}{V}\right) \dots \prod_{i=1}^{M^k-1} \left(1 - \frac{i}{V}\right) \\ &\geq 1 - \frac{1}{V} \left( \sum_{i=1}^{M-1} i + \sum_{i=1}^{M^2-1} i + \dots + \sum_{i=1}^{M^k-1} i \right) \\ &\geq 1 - \frac{1}{2V} (M^2 + M^4 + \dots + M^{2k}) \\ &\geq 1 - \frac{M^{2(k+1)}}{2V(M^2-1)} \geq 1 - \frac{2M^{2k}}{3V}. \end{aligned}$$

In the last steps we have assumed  $M \geq 2$ .

Let  $S$  be the event that  $(\eta^{a_1 \circ a_2 \circ \dots \circ a_k})_1 = \tau$ . Then, using the independence of the random variables labelling the nodes of a free function tree, and using Equation (3.3), we see that

$$\Pr(S|F) = \prod_{\{i \in T \mid 1 \leq |i| \leq k\}} \mathcal{P}^{\tau(i)} = \rho([\tau]).$$

Hence

$$\begin{aligned} \rho_V([\tau]) &= \Pr(S) = \Pr(F) \Pr(S|F) + \Pr(F^C) \Pr(S|F^C) \\ &\leq \Pr(S|F) + \Pr(F^C) \leq \rho([\tau]) + \frac{2M^{2k}}{3V}. \end{aligned}$$

Similarly,

$$\rho_V([\tau]) \geq \Pr(F) \Pr(S|F)$$

$$\geq \rho([\tau]) - \frac{2M^{2k}}{3V}.$$

Hence

$$(4.17) \quad |\rho_V([\tau]) - \rho([\tau])| \leq \frac{2M^{2k}}{3V}.$$

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

For each level- $k$  code tree  $\tau \in T_k$  choose some  $\omega_\tau \in [\tau] \subseteq \Omega$ . It then follows that

$$\begin{aligned} \left| \int_{\Omega} f d\rho - \int_{\Omega} f d\rho_V \right| &= \left| \sum_{\tau \in T_k} \left( \int_{[\tau]} f d\rho - \int_{[\tau]} f d\rho_V \right) \right| = \\ & \left| \sum_{\tau \in T_k} \int_{[\tau]} (f - f(\omega_\tau)) d\rho - \sum_{\tau \in T_k} \int_{[\tau]} ((f - f(\omega_\tau)) d\rho_V + \sum_{\tau \in T_k} f(\omega_\tau)(\rho([\tau]) - \rho_V([\tau])) \right| \\ & \leq \frac{1}{M^{k+1}} \sum_{\tau \in T_k} \rho([\tau]) + \frac{1}{M^{k+1}} \sum_{\tau \in T_k} \rho_V([\tau]) + \sum_{\tau \in T_k} \frac{M^{2k}}{3V} \\ & = \varphi(k) := \frac{2}{M^{k+1}} + \frac{M^{3k}}{3V}, \end{aligned}$$

since  $diam [\tau] \leq \frac{1}{M^{k+1}}$  from Equation (3.2),  $|f(\omega_\tau)| \leq \frac{1}{2}$ ,  $Lip f \leq 1$ ,  $\omega_\tau \in [\tau]$ , and using Equation 4.17. Choose  $x$  so that  $\frac{2V}{M} = M^{4x}$ . This is the value of  $x$  which minimizes  $\left( \frac{2}{M^{x+1}} + \frac{M^{3x}}{3V} \right)$ . Choose  $k$  so that  $k \leq x \leq k+1$ . Then

$$\begin{aligned} \varphi(k) &\leq 2 \left( \frac{M}{2V} \right)^{\frac{1}{4}} + \frac{1}{3V} \left( \frac{2V}{M} \right)^{\frac{3}{4}} = 2^{\frac{3}{4}} \left( \frac{M}{V} \right)^{\frac{1}{4}} \left( 1 + \frac{1}{3M} \right) \\ &\leq \frac{7}{2^{\frac{1}{4}} 3} \left( \frac{M}{V} \right)^{\frac{1}{4}}, \quad (M \geq 2). \end{aligned}$$

Hence Equation (4.13) is true.  $\square$

## 5. SUPERFRACTALS

### 5.1. Contraction Mappings on $\mathbb{H}^V$ and the Superfractal Set $\mathfrak{H}_{V,1}$ .

**Definition 9.** Let  $V \in \mathbb{N}$ , let  $\mathcal{A}$  be the index set introduced in Equation (4.1), let  $\mathcal{F}$  be given as in Equation (3.1), and let probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  be given as in Equation (4.5). Define

$$f^a : \mathbb{H}^V \rightarrow \mathbb{H}^V$$

by

$$(5.1) \quad f^a(K) = \left( \bigcup_{m=1}^M f_m^{n_1}(K_{v_{1,m}}), \bigcup_{m=1}^M f_m^{n_2}(K_{v_{2,m}}), \dots, \bigcup_{m=1}^M f_m^{n_V}(K_{v_{V,m}}) \right)$$

$\forall K = (K_1, K_2, \dots, K_V) \in \mathbb{H}^V, \forall a \in \mathcal{A}$ . Let

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

**Theorem 15.**  $\mathcal{F}_V$  is an IFS with contractivity factor  $l$ .

*Proof.* We only need to prove that the mapping  $f^a : \mathbb{H}^V \rightarrow \mathbb{H}^V$  is contractive with contractivity factor  $l$ ,  $\forall a \in \mathcal{A}$ . Note that,  $\forall K = (K_1, K_2, \dots, K_M)$ ,  $L = (L_1, L_2, \dots, L_M) \in \mathbb{H}^M$ ,

$$\begin{aligned} & d_{\mathbb{H}}\left(\bigcup_{m=1}^M f_m^n(K_m), \bigcup_{m=1}^M f_m^n(L_m)\right) \\ & \leq \max_m \{d_{\mathbb{H}}(f_m^n(K_m), f_m^n(L_m))\} \\ & \leq \max_m \{l \cdot d_{\mathbb{H}}(K_m, L_m)\} \\ & = l \cdot d_{\mathbb{H}^M}(K, L). \end{aligned}$$

Hence,  $\forall (K_1, K_2, \dots, K_V), (L_1, L_2, \dots, L_V) \in \mathbb{H}^V$ ,

$$\begin{aligned} & d_{\mathbb{H}^V}(f^a(K_1, K_2, \dots, K_V), f^a(L_1, L_2, \dots, L_V)) \\ & = \max_v \{d_{\mathbb{H}}\left(\bigcup_{m=1}^M f_m^{n_v}(K_{v,m}), \bigcup_{m=1}^M f_m^{n_v}(L_{v,m})\right)\} \\ & \leq \max_v \{l \cdot d_{\mathbb{H}^M}((K_{v,1}, K_{v,2}, \dots, K_{v,M}), \\ & \quad (L_{v,1}, L_{v,2}, \dots, L_{v,M}))\} \\ & \leq l \cdot d_{\mathbb{H}^V}((K_1, K_2, \dots, K_V), (L_1, L_2, \dots, L_V)). \end{aligned}$$

□

The theory of IFS in Section 2.1 applies to the IFS  $\mathcal{F}_V$ . It possesses a unique set attractor  $\mathfrak{H}_V \in \mathbb{H}(\mathbb{H}^V)$ , and a unique measure attractor  $\mathfrak{P}_V \in \mathbb{P}(\mathbb{H}^V)$ . The random iteration algorithm corresponding to the IFS  $\mathcal{F}_V$  may be used to approximate sequences of points ( $V$ -tuples of compact sets) in  $\mathfrak{H}_V$  distributed according to the probability measure  $\mathfrak{P}_V$ .

However, the individual components of these vectors in  $\mathfrak{H}_V$ , certain special subsets of  $\mathbb{X}$ , are the objects we are interested in. Accordingly, for all  $v \in \{1, 2, \dots, V\}$ , let us define  $\mathfrak{H}_{V,v} \subset \mathbb{H}$  to be the set of  $v^{\text{th}}$  components of points in  $\mathfrak{H}_V$ .

**Theorem 16.** For all  $v \in \{1, 2, \dots, V\}$  we have

$$\mathfrak{H}_{V,v} = \mathfrak{H}_{V,1}.$$

When the probabilities in the superIFS  $\mathcal{F}_V$  are given by (4.6), then starting from any initial  $V$ -tuple of non-empty compact subsets of  $\mathbb{X}$ , the random distribution of the sets  $K \in \mathbb{H}$  that occur in the  $v^{\text{th}}$  component of vectors produced by the random iteration algorithm after  $n$  initial steps converge weakly to the marginal probability measure

$$\mathfrak{P}_{V,1}(B) := \mathfrak{P}_V(B, \mathbb{H}, \mathbb{H}, \dots, \mathbb{H}) \forall B \in \mathbb{B}(\mathbb{H}),$$

independently of  $v$ , almost always, as  $n \rightarrow \infty$ .

*Proof.* The direct way to prove this theorem is to parallel the proof of Theorem 10, using the maps  $\{f^a : \mathbb{H}^V \rightarrow \mathbb{H}^V \mid a \in \mathcal{A}\}$  in place of the maps  $\{\eta^a : \Omega^V \rightarrow \Omega^V \mid a \in \mathcal{A}\}$ .

However, an alternate proof follows with the aid of the map  $\mathcal{F} : \Omega \rightarrow \mathbb{H}(\mathbb{X})$  introduced in Theorem 7. We have put this alternate proof at the end of the proof of Theorem 17. □

**Definition 10.** We call  $\mathfrak{H}_{V,1}$  a superfractal set. Points in  $\mathfrak{H}_{V,1}$  are called  $V$ -variable fractal sets.

**Example 4.** See Figure 21. This example is similar to the one in Section 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

$$\begin{aligned} f_1^1(x, y) &= \left( \frac{1.629x + 0.135y - 1.99}{-0.780x + 0.864y - 2.569}, \frac{0.505x + 1.935y - 0.216}{0.780x - 0.864y + 2.569} \right), \\ f_2^1(x, y) &= \left( \frac{1.616x - 2.758y + 3.678}{1.664x - 0.944y + 3.883}, \frac{2.151x + 0.567y + 2.020}{1.664x - 0.944y + 3.883} \right), \\ f_1^2(x, y) &= \left( \frac{1.667x + .098y - 2.005}{-0.773x + 0.790y - 2.575}, \frac{0.563x + 2.064y - 0.278}{0.773x - 0.790y + 2.575} \right), \\ f_2^2(x, y) &= \left( \frac{1.470x - 2.193y + 3.035}{2.432x - 0.581y + 2.872}, \frac{1.212x + 0.686y + 2.059}{2.432x - 0.581y + 2.872} \right). \end{aligned}$$

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  $\square := \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$ , in which each ti-tree lies, into itself. Both  $f_2^1(x, y)$  and  $f_2^2(x, y)$  map ti-trees to lower right branches of ti-trees. Both  $f_1^1(x, y)$  and  $f_1^2(x, y)$  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. The top middle image corresponds to the fifth iteration. Both the images at the bottom middle and bottom left correspond to more than thirty iterations, and are representative of typical images produced after more than thirty iterations. (We carried out more than fifty iterations.) They represent images selected from the superfractal  $\mathfrak{H}_{2,1}$  according to the invariant measure  $\mathfrak{P}_{2,1}$ . Note that it is the support of the red and green pixels that corresponds to an element of  $\mathfrak{H}_{2,1}$ . Note too the “texture effect”, similar to the one discussed in Example 3.

By Theorem 5 there is a continuous mapping  $\mathcal{F}_V : \Sigma_V \rightarrow \mathfrak{H}_V$  that assigns to each address in the code space  $\Sigma_V = \mathcal{A}^\infty$  a  $V$ -tuple of compact sets in  $\mathfrak{H}_V$ . But this mapping is not helpful for characterizing  $\mathfrak{H}_V$  because  $\mathcal{F}_V : \Sigma_V \rightarrow \mathfrak{H}_V$  is not in general one-to-one, for the same reason that  $\Phi : \Sigma_V \rightarrow \Omega_V$  is not one-to-one, as explained in Section 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  $\mathcal{F}_V$  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  $\mathfrak{H}_V$  and  $\mathfrak{H}_{V,1}$ , on the other.)

**Theorem 17.** Let  $a \in \mathcal{A}$  and  $\eta^a : \Omega^V \rightarrow \Omega^V$  be defined as in Theorem 9. Let  $f^a : \mathbb{H}^V \rightarrow \mathbb{H}^V$  be defined as in Theorem 15. Let  $\mathcal{F} : \Omega \rightarrow \mathbb{H}(\mathbb{X})$  be the mapping

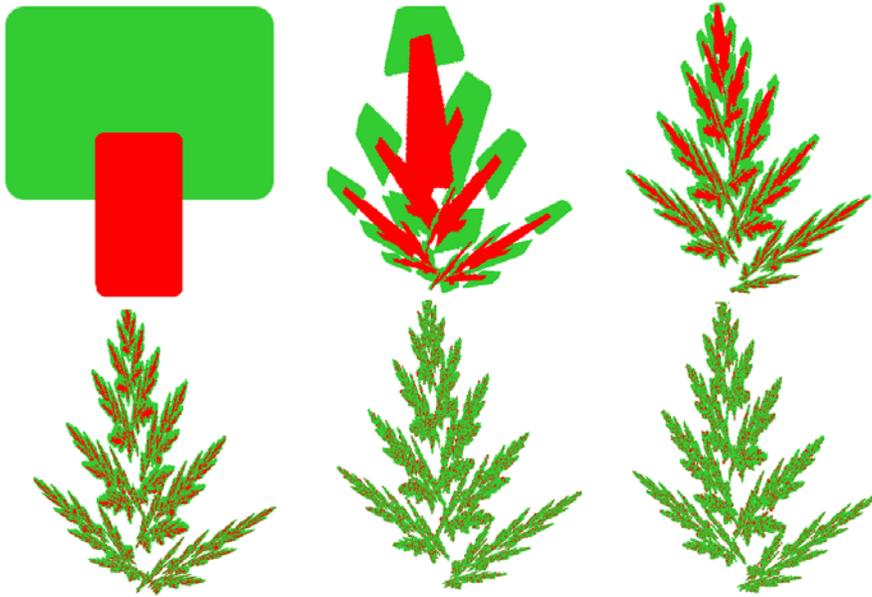


FIGURE 21. Sequence of images converging to 2-variable fractals, see Example 4. Convergence to within the numerical resolution has occurred in the bottom left two images. Note the subtle but real differences between the silhouettes of these two sets. A variant of the “texture effect” can also be seen. The red points appear to dance forever on the green ti-trees, while the ti-trees dance forever on the superfractal.

introduced in Theorem 7. Define  $\mathcal{F} : \Omega^V \rightarrow (\mathbb{H}(\mathbb{X}))^V$  by

$$\mathcal{F}(\omega_1, \omega_2, \dots, \omega_V) = (\mathcal{F}(\omega_1), \mathcal{F}(\omega_2), \dots, \mathcal{F}(\omega_V)),$$

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

$$(5.3) \quad \mathcal{F}(\eta^a(\omega)) = f^a(\mathcal{F}(\omega)) \quad \forall a \in \mathcal{A}, \omega \in \Omega^V.$$

Also

$$(5.4) \quad \mathcal{F}(\Omega_V) = \mathfrak{H}_V \text{ and } \mathcal{F}(\Omega_{V,1}) = \mathfrak{H}_{V,1},$$

where  $\Omega_{V,v}$  denotes the set of  $v^{\text{th}}$  components of members of  $\Omega_V$ . Similarly, when the probabilities in the IFS  $\mathcal{F}_V$  of Equation (5.2), are given by Equation (4.6) we have

$$(5.5) \quad \mathcal{F}(\mu_V) = \mathfrak{P}_V, \text{ and } \mathcal{F}(\rho_V) = \mathfrak{P}_{V,1},$$

where  $\rho_V$  is the marginal probability distribution given by Equation (4.10).

*Proof.* We begin by establishing the key Equation (5.3). Note that from Theorem 7, for any  $K \in \mathbb{H}(\mathbb{X})$ ,

$$(5.6) \quad \begin{aligned} \mathcal{F}(\omega) &= (\mathcal{F}(\omega_1), \mathcal{F}(\omega_2), \dots, \mathcal{F}(\omega_V)) \\ &= \left( \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_1)(K)\}, \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_2)(K)\}, \dots, \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_V)(K)\} \right). \end{aligned}$$

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

$$(5.7) \quad \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_1)(K)\} \\ = \lim_{k \rightarrow \infty} \left\{ \bigcup_{\{i \in T \mid |i|=k\}} f_{i_1}^{\omega_1(\emptyset)} \circ f_{i_2}^{\omega_1(i_1)} \circ \dots \circ f_{i_k}^{\omega_1(i_1 i_2 \dots i_{k-1})}(K) \right\}.$$

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

$$f^a(\mathcal{F}(\omega)) = f^a\left(\lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_1)(K)\}, \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_2)(K)\}, \dots, \lim_{k \rightarrow \infty} \{\mathcal{F}_k(\omega_V)(K)\}\right) \\ = \lim_{k \rightarrow \infty} \{f^a(\mathcal{F}_k(\omega_1)(K), \mathcal{F}_k(\omega_2)(K), \dots, \mathcal{F}_k(\omega_V)(K))\}.$$

By the definition in Theorem 15, Equation 5.1, we have

$$f^a(\mathcal{F}_k(\omega_1)(K), \mathcal{F}_k(\omega_2)(K), \dots, \mathcal{F}_k(\omega_V)(K)) \\ = \left( \bigcup_{m=1}^M f_m^{n_1}(\mathcal{F}_k(\omega_{v_{1,m}})(K)), \bigcup_{m=1}^M f_m^{n_2}(\mathcal{F}_k(\omega_{v_{2,m}})(K)), \dots, \bigcup_{m=1}^M f_m^{n_V}(\mathcal{F}_k(\omega_{v_{V,m}})(K)) \right).$$

By equation (3.4) the first component here is

$$\bigcup_{m=1}^M f_m^{n_1}(\mathcal{F}_k(\omega_{v_{1,m}})(K)) \\ = \bigcup_{m=1}^M f_m^{n_1} \left( \bigcup_{\{i \in T \mid |i|=k\}} f_{i_1}^{\omega_{v_{1,m}}(\emptyset)} \circ f_{i_2}^{\omega_{v_{1,m}}(i_1)} \circ \dots \circ f_{i_k}^{\omega_{v_{1,m}}(i_1 i_2 \dots i_{k-1})}(K) \right) \\ = \mathcal{F}_{k+1}(\xi_{n_1}(\omega_{v_{1,1}}, \omega_{v_{1,2}}, \dots, \omega_{v_{1,M}}))(K),$$

where we have used the definition in Equation (4.7). Hence

$$f^a(\mathcal{F}(\omega)) = \lim_{k \rightarrow \infty} \{(f^a(\mathcal{F}_k(\omega_1)(K), \mathcal{F}_k(\omega_2)(K), \dots, \mathcal{F}_k(\omega_V)(K)))\} = \\ \lim_{k \rightarrow \infty} \{(\mathcal{F}_{k+1}(\xi_{n_1}(\omega_{v_{1,1}}, \omega_{v_{1,2}}, \dots, \omega_{v_{1,M}}))(K), \mathcal{F}_{k+1}(\xi_{n_2}(\omega_{v_{2,1}}, \omega_{v_{2,2}}, \dots, \omega_{v_{2,M}}))(K), \\ \dots, \mathcal{F}_{k+1}(\xi_{n_V}(\omega_{v_{V,1}}, \omega_{v_{V,2}}, \dots, \omega_{v_{V,M}}))(K))\}.$$

Comparing with Equations (5.6) and (5.7), we find that the right hand side here converges to  $\mathcal{F}(\eta^a(\omega))$  as  $k \rightarrow \infty$ . So Equation (5.3) is true.

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

$$\mathcal{F}(\Omega_V) = \mathcal{F}\left(\bigcup_{a \in \mathcal{A}} \eta^a(\Omega_V)\right) = \bigcup_{a \in \mathcal{A}} \mathcal{F}(\eta^a(\Omega_V)) = \bigcup_{a \in \mathcal{A}} f^a(\mathcal{F}(\Omega_V)).$$

It follows by uniqueness that  $\mathcal{F}(\Omega_V)$  must be *the* set attractor of the IFS  $\mathcal{F}_V$ . Hence  $\mathcal{F}(\Omega_V) = \mathfrak{H}_V$  which is the first statement in Equation (5.4). Now

$$\mathcal{F}(\Omega_{V,1}) = \{\mathcal{F}(\omega_1) \mid (\omega_1, \omega_2, \dots, \omega_V) \in \Omega_V\} \\ = \text{first component of } \{(\mathcal{F}(\omega_1), \mathcal{F}(\omega_2), \dots, \mathcal{F}(\omega_V)) \mid (\omega_1, \omega_2, \dots, \omega_V) \in \Omega_V\} \\ = \text{first component of } \mathcal{F}(\Omega_V) = \text{first component of } \mathfrak{H}_V = \mathfrak{H}_{V,1},$$

which contains the second statement in Equation (5.4).

In a similar manner we consider the push-forward under  $\mathcal{F} : \Omega^V \rightarrow \mathbb{H}^V$  of the invariant measure  $\mu_V$  of the IFS  $\Phi = \{\Omega^V; \eta^a, \mathcal{P}^a, a \in \mathcal{A}\}$ .  $\mathcal{F}(\mu_V)$  is normalized, i.e.  $\mathcal{F}(\mu_V) \in \mathbb{P}(\mathbb{H}^V)$ , because  $\mathcal{F}(\mu_V)(\mathbb{H}^V) = \mu_V(\mathcal{F}^{-1}(\mathbb{H}^V)) = \mu_V(\Omega^V)$ . We now

show that  $\mathcal{F}(\mu_V)$  is the measure attractor the IFS  $\mathcal{F}_V$ . The measure attractor of the IFS  $\Phi$  obeys

$$\mu_V = \sum_{a \in \mathcal{A}} \mathcal{P}^a \eta^a(\mu_V).$$

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

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

where in the last step we have used the key Equation (5.3). So  $\mathcal{F}(\mu_V)$  is the measure attractor of the IFS  $\mathcal{F}_V$ . Using uniqueness, we conclude  $\mathcal{F}(\mu_V) = \mathfrak{P}_V$  which is the first equation in Equation (5.5). Finally, observe that, for all  $B \in \mathbb{B}(\mathbb{H})$ ,

$$\begin{aligned} \mathfrak{P}_{V,1}(B) &= \mathcal{F}(\mu_V)(B, \mathbb{H}, \mathbb{H}, \dots, \mathbb{H}) = \mu_V(\mathcal{F}^{-1}(B, \mathbb{H}, \mathbb{H}, \dots, \mathbb{H})) \\ &= \mu_V((\mathcal{F}^{-1}(B), \mathcal{F}^{-1}(\mathbb{H}), \mathcal{F}^{-1}(\mathbb{H}), \dots, \mathcal{F}^{-1}(\mathbb{H}))) \text{ (using Equation (5.6))} \\ &= \mu_V((\mathcal{F}^{-1}(B), \Omega, \Omega, \dots, \Omega)) \text{ (since } \mathcal{F} : \Omega \rightarrow \mathbb{H}) \\ &= \rho_V(\mathcal{F}^{-1}(B)) \text{ (by definition (4.10))} = \mathcal{F}(\rho_V)(B). \end{aligned}$$

This contains the second equation in Equation (5.5).

In a similar way, we obtain the alternate proof of Theorem 16. Simply lift Theorem 10 to the domain of the IFS  $\mathcal{F}_V$  using  $\mathcal{F} : \Omega^V \rightarrow \mathbb{H}^V$ .  $\square$

The code tree  $\Phi(a_1 a_2 a_3 \dots)$  is called a tree address of the V-variable fractal  $\mathcal{F}(a_1 a_2 a_3 \dots)$ .

The mapping  $\mathcal{F} : \Omega_{V,1} \rightarrow \mathfrak{H}_{V,1}$  together with Theorem 11 provides a characterization of V-variable fractals as follows. At any ‘‘magnification’’, any V-variable fractal set is made of  $V$  ‘‘forms’’ or ‘‘shapes’’:

**Theorem 18.** *Let  $M \in \mathfrak{H}_{V,1}$  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  $\mathbb{X}$ , and the diameter of each of these transformed sets is at most  $\epsilon$ .*

*Proof.* Choose  $n$  so that  $l^n < \epsilon$ . Note that

$$\mathfrak{H}_V = \bigcup_{a \in \mathcal{A}} f^a(\mathfrak{H}_V) = \bigcup_{(a_1, a_2, \dots, a_n) \in \mathcal{A}} f^{a_1} \circ f^{a_2} \circ \dots \circ f^{a_n}(\mathfrak{H}_V).$$

Hence, since  $(M, M, \dots, M) \in \mathfrak{H}_V$  it follows that there exists  $(K_1, K_2, \dots, K_V) \in \mathfrak{H}_V$  such that

$$M \in \text{first component of } \bigcup_{(a_1, a_2, \dots, a_n) \in \mathcal{A}} f^{a_1} \circ f^{a_2} \circ \dots \circ f^{a_n}(K_1, K_2, \dots, K_V).$$

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

**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 right, from top to bottom, a sequence of six successive images, illustrating successive 2-variable fractals, 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 Section 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 forty, forty-one, forty-two, forty-three, forty-four, and forty-five iterations. The key features of the transformations can be deduced from the images. (For example,

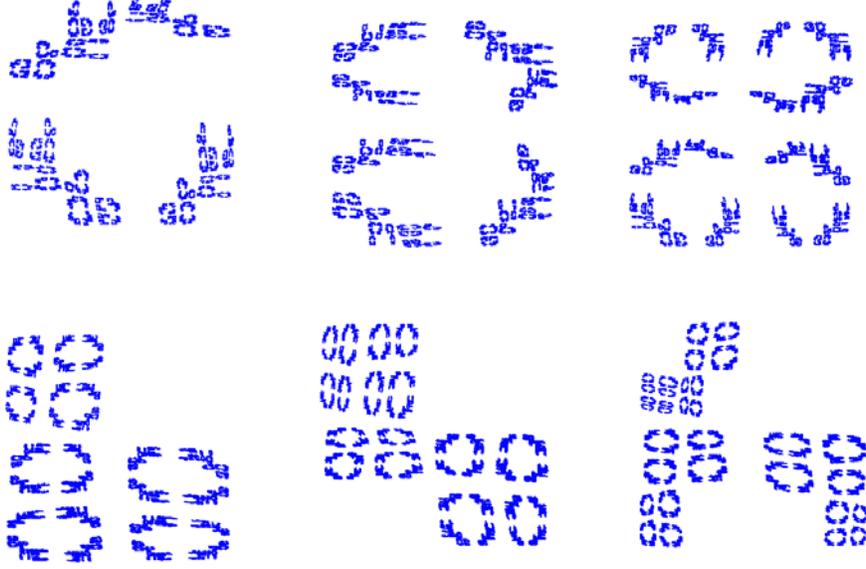


FIGURE 22. A sequence of 2-variable fractal sets (roughly accurate to viewing resolution), corresponding to  $M=N=2$ . Can you spot the projective transformations? See Example 4. Each of these six images exhibits “2-variability”: at several scales, each looks as though it is the union of projective transformations of at most two distinct sets.

one of the transformations 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  $\mathcal{F}$  introduced in Section 3.3; that is, in the metric of  $\mathbb{H}(\mathbb{H}(\mathbb{X}))$ ,*

$$(5.8) \quad \lim_{V \rightarrow \infty} \mathfrak{H}_{V,1} = \mathfrak{H}.$$

Moreover, if the probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  obey Equation (4.6), then in the metric of  $\mathbb{P}(\mathbb{H}(\mathbb{X}))$

$$\lim_{V \rightarrow \infty} \mathfrak{P}_{V,1} = \mathfrak{P},$$

where  $\mathfrak{P}$  is the stationary measure on random fractal sets associated with the superIFS  $\mathcal{F}$ .

*Proof.* We have, using the mapping  $\mathcal{F} : \Omega \rightarrow \mathbb{H}(\mathbb{X})$ ,

$$\begin{aligned} \lim_{V \rightarrow \infty} \mathfrak{H}_{V,1} &= \lim_{V \rightarrow \infty} \mathcal{F}(\Omega_{V,1}) \text{ (by Theorem 17)} \\ &= \mathcal{F}\left(\lim_{V \rightarrow \infty} \Omega_{V,1}\right) \text{ (since } \mathcal{F} : \Omega \rightarrow \mathbb{H}(\mathbb{X}) \text{ is continuous by Theorem 7)} \\ &= \mathcal{F}(\Omega) \text{ (by Theorem 12)} \end{aligned}$$

=  $\mathfrak{H}$  (by Equation (3.6)).

Similarly, using the mapping  $\mathcal{F} : \Omega \rightarrow \mathbb{P}(\mathbb{X})$ , have

$$\begin{aligned} & \lim_{V \rightarrow \infty} \mathfrak{P}_{V,1} = \lim_{V \rightarrow \infty} \mathcal{F}(\rho_V) \text{ (by Theorem 17)} \\ = & \mathcal{F}(\lim_{V \rightarrow \infty} \rho_V) \text{ (since } \mathcal{F} : \Omega \rightarrow \mathbb{P}(\mathbb{X}) \text{ is continuous by Theorem 7)} \\ & = \mathcal{F}(\rho) \text{ (by Theorem 12)} \\ & = \mathfrak{P} \text{ (by Equation (3.7)).} \end{aligned}$$

□

## 5.2. Contraction Mappings on $\mathbb{P}^V$ and the Superfractal Measures $\tilde{\mathfrak{H}}_{V,1}$ .

Recall that  $\mathbb{P} = \mathbb{P}(\mathbb{X})$ . Let  $\mathbb{P}^V = (\mathbb{P}(\mathbb{X}))^V$ . In this section we follow the same lines as in Section 5.1, constructing an IFS using the individual IFSSs of the superIFS  $\mathcal{F}$ , except that here the underlying space is  $\mathbb{P}^V$  instead of  $\mathbb{H}^V$ .

**Definition 11.** Let  $V \in \mathbb{N}$ , let  $\mathcal{A}$  be the index set introduced in Equation (4.1), let  $\mathcal{F}$  be given as in Equation (3.1), and let probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  be given as in Equation (4.5). Define

$$f^a : \mathbb{P}^V \rightarrow \mathbb{P}^V$$

by

$$(5.9) \quad f^a(\mu) = \left( \sum_{m=1}^M p_m^{n_1} f_m^{n_1}(\mu_{v_{1,m}}), \sum_{m=1}^M p_m^{n_2} f_m^{n_2}(\mu_{v_{2,m}}), \dots, \sum_{m=1}^M p_m^{n_V} f_m^{n_V}(\mu_{v_{V,m}}) \right)$$

$\mu = (\mu_1, \mu_2, \dots, \mu_V) \in \mathbb{P}^V$ . Let

$$(5.10) \quad \tilde{\mathcal{F}}_V := \{\mathbb{P}^V; f^a, \mathcal{P}^a, a \in \mathcal{A}\}.$$

**Theorem 20.**  $\tilde{\mathcal{F}}_V$  is an IFS with contractivity factor  $l$ .

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

$$\begin{aligned} & d_{\mathbb{P}}\left(\sum_{m=1}^M p_m^n f_m^n(\mu_m), \sum_{m=1}^M p_m^n f_m^n(\varphi_m)\right) \\ & \leq \sum_{m=1}^M d_{\mathbb{P}}(p_m^n f_m^n(\mu_m), p_m^n f_m^n(\varphi_m)) \\ & \leq l \cdot \max_m \{d_{\mathbb{P}}(\mu_m, \varphi_m)\} \\ & = l \cdot d_{\mathbb{P}^M}(\mu, \varphi). \end{aligned}$$

Hence,  $\forall (\mu_1, \mu_2, \dots, \mu_V), (\varphi_1, \varphi_2, \dots, \varphi_V) \in \mathbb{P}^V$ ,

$$\begin{aligned} & d_{\mathbb{P}^V}(f^a(\mu_1, \mu_2, \dots, \mu_V), f^a(\varphi_1, \varphi_2, \dots, \varphi_V)) \\ = & \max_v \{d_{\mathbb{P}}\left(\sum_{m=1}^M p_m^{n_v} f_m^{n_v}(\mu_{v,m}), \sum_{m=1}^M p_m^{n_v} f_m^{n_v}(\varphi_{v,m})\right)\} \\ \leq & \max_v \{l \cdot d_{\mathbb{P}^M}((\mu_{v,1}, \mu_{v,2}, \dots, \mu_{v,M}), \\ & (\varphi_{v,1}, \varphi_{v,2}, \dots, \varphi_{v,M}))\} \\ \leq & l \cdot d_{\mathbb{P}^V}((\mu_1, \mu_2, \dots, \mu_M), (\varphi_1, \varphi_2, \dots, \varphi_M)). \end{aligned}$$

□

The set attractor of the IFS  $\tilde{\mathcal{F}}_V$  is  $\tilde{\mathfrak{H}}_V \in \mathbb{H}(\mathbb{P}^V)$ , a subset of  $\mathbb{P}^V$ , a set of  $V$ -tuples of probability measures on  $\mathbb{X}$ . As we will see, each of these measures is supported on a  $V$ -variable fractal set belonging to the superfactal  $\tilde{\mathfrak{H}}_{V,1}$ . The measure attractor of the IFS  $\tilde{\mathcal{F}}_V$  is a probability measure  $\tilde{\mathfrak{P}}_V \in \mathbb{P}(\mathbb{P}^V)$ , 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  $\tilde{\mathcal{F}}_V$  may be used to approximate sequences of points in  $\tilde{\mathfrak{H}}_V$ , namely vectors of measures on  $\mathbb{X}$ , distributed according to the probability measure  $\tilde{\mathfrak{P}}_V$ .

As in Section 5.1, we define  $\tilde{\mathfrak{H}}_{V,v}$  to be the set of  $v^{\text{th}}$  components of sets in  $\tilde{\mathfrak{H}}_V$ , for  $v \in \{1, 2, \dots, V\}$ .

**Theorem 21.** *For all  $v \in \{1, 2, \dots, V\}$  we have*

$$\tilde{\mathfrak{H}}_{V,v} = \tilde{\mathfrak{H}}_{V,1}.$$

*When the probabilities in the IFS  $\tilde{\mathcal{F}}_V$  are given by Equation (4.6), then starting at any initial  $V$ -tuple of probability measures on  $\mathbb{X}$ , the probability measures  $\mu \in \mathbb{P}(\mathbb{X})$  that occur in the  $v^{\text{th}}$  component of points produced by the random iteration algorithm after  $n$  steps converge weakly to the marginal probability measure*

$$\tilde{\mathfrak{P}}_{V,1}(B) := \tilde{\mathfrak{P}}_V(B, \mathbb{P}, \mathbb{P}, \dots, \mathbb{P}) \forall B \in \mathbb{B}(\mathbb{P}),$$

*independently of  $v$ , almost always, as  $n \rightarrow \infty$ .*

*Proof.* The direct way to prove this theorem is to parallel the proof of Theorem 10, using the maps  $\{f^a : \mathbb{P}^V \rightarrow \mathbb{P}^V \mid a \in \mathcal{A}\}$  in place of the maps  $\{\eta^a : \Omega^V \rightarrow \Omega^V \mid a \in \mathcal{A}\}$ .

It is simpler however to lift Theorem 10 to the domain of the IFS  $\tilde{\mathcal{F}}_V$  using  $\tilde{\mathcal{F}} : \Omega^V \rightarrow \mathbb{P}^V$  which is defined in Theorem 22 with the aid of the mapping  $\tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P} = \mathbb{P}(\mathbb{X})$  introduced in Theorem 7. We omit the details as they are straightforward. □

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

**Example 6.** See Figure 23. This example corresponds to the same superIFS as in Example 4. The probabilities of the functions in the IFSs are  $p_1^1 = p_1^2 = 0.74$ , and  $p_2^1 = p_2^2 = 0.26$ . 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  $a \in \mathcal{A}$  and  $\eta^a = \Omega^V \rightarrow \Omega^V$  be defined as in Theorem 9. Let  $f^a : \mathbb{P}^V \rightarrow \mathbb{P}^V$  be defined as in Theorem 20. Let  $\tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P} = \mathbb{P}(\mathbb{X})$  be the mapping introduced in Theorem 7. Define  $\tilde{\mathcal{F}} : \Omega^V \rightarrow \mathbb{P}^V = (\mathbb{P}(\mathbb{X}))^V$  by*

$$\tilde{\mathcal{F}}(\omega_1, \omega_2, \dots, \omega_V) = (\tilde{\mathcal{F}}(\omega_1), \tilde{\mathcal{F}}(\omega_2), \dots, \tilde{\mathcal{F}}(\omega_V)),$$

*for all  $(\omega_1, \omega_2, \dots, \omega_V) \in \Omega^V$ . Then*

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



FIGURE 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 coloured either black or a shade of green, using a similar technique to the one used in Example 2. The intensity of the green of a pixel is a monotonic increasing function of the measure of the pixel.

Also

$$\tilde{\mathcal{F}}(\Omega_V) = \tilde{\mathfrak{H}}_V \text{ and } \tilde{\mathcal{F}}(\Omega_{V,1}) = \tilde{\mathfrak{H}}_{V,1},$$

where  $\Omega_{V,v}$  denotes the set of  $v^{\text{th}}$  components of members of  $\Omega_V$ . Similarly, when the probabilities in the IFS  $\tilde{\mathcal{F}}_V$  of Equation (5.10), are given by Equation (4.6) we have

$$\tilde{\mathcal{F}}(\mu_V) = \tilde{\mathfrak{P}}_V, \text{ and } \tilde{\mathcal{F}}(\rho_V) = \tilde{\mathfrak{P}}_{V,1},$$

where  $\rho_V$  is the marginal probability distribution given by Equation (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.  $\square$

**Definition 12.** The code tree  $\Phi(a_1a_2a_3\dots)$  is called a tree address of the  $V$ -variable fractal measure  $\tilde{\mathcal{F}}_V(a_1a_2a_3\dots)$ .

The mapping  $\tilde{\mathcal{F}} : \Omega_{V,1} \rightarrow \tilde{\mathfrak{H}}_{V,1}$  together with Theorem 11 allows us to characterize  $V$ -variable fractals as follows:

**Theorem 23.** Let  $\tilde{\mu} \in \tilde{\mathfrak{H}}_{V,1}$  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  $\mathbb{X}$ , and the diameter of the support of each of these transformed measures is at most  $\epsilon$ .

*Proof.* Choose  $n$  so that  $l^n < \epsilon$ . Note that

$$\tilde{\mathfrak{H}}_V = \bigcup_{a \in \mathcal{A}} f^a(\tilde{\mathfrak{H}}_V) = \bigcup_{(a_1, a_2, \dots, a_n) \in \mathcal{A}} f^{a_1} \circ f^{a_2} \circ \dots \circ f^{a_n}(\tilde{\mathfrak{H}}_V).$$

Hence, since  $(\tilde{\mu}, \tilde{\mu}, \dots, \tilde{\mu}) \in \mathfrak{H}_V$  it follows that there exists  $(\varpi_1, \varpi_2, \dots, \varpi_V) \in \mathfrak{H}_V$  such that

$$\tilde{\mu} \in \text{first component of } \bigcup_{(a_1, a_2, \dots, a_n) \in \mathcal{A}} f^{a_1} \circ f^{a_2} \circ \dots \circ f^{a_n}(\varpi_1, \varpi_2, \dots, \varpi_V).$$

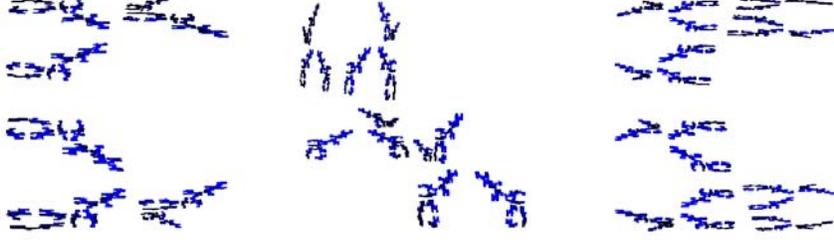


FIGURE 24. Three successive 2-variable fractal measures, in shades of blue. Illustrates the “shapes” and “forms” theorem. See Example 7.

Inspection of Equation (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 Figure 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 it 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 together with a scaling of brightness.

**Theorem 24.** *The set of  $V$ -variable fractal measures associated with the superIFS  $\tilde{\mathcal{F}}_V$  converges to the set of fractal measures introduced in Section 3.3; that is, in the metric of  $\mathbb{H}(\mathbb{P}(\mathbb{X}))$*

$$\lim_{V \rightarrow \infty} \tilde{\mathfrak{H}}_{V,1} = \tilde{\mathfrak{H}}.$$

If the probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  obey Equation (4.6), then in the metric of  $\mathbb{P}(\mathbb{P}(\mathbb{X}))$

$$\lim_{V \rightarrow \infty} \tilde{\mathfrak{P}}_{V,1} = \tilde{\mathfrak{P}},$$

where  $\tilde{\mathfrak{P}}$  is the stationary measure on fractal sets associated with the superIFS  $\mathcal{F}$ .

*Proof.* We have, using the mapping  $\tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P} = \mathbb{P}(\mathbb{X})$ ,

$$\begin{aligned} \lim_{V \rightarrow \infty} \tilde{\mathfrak{H}}_{V,1} &= \lim_{V \rightarrow \infty} \tilde{\mathcal{F}}(\Omega_{V,1}) \text{ (by Theorem 22)} \\ &= \tilde{\mathcal{F}}\left(\lim_{V \rightarrow \infty} \Omega_{V,1}\right) \text{ (since } \tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P} = \mathbb{P}(\mathbb{X}) \text{ is continuous by Theorem 7)} \\ &= \tilde{\mathcal{F}}(\Omega) \text{ (by Theorem 12)} \\ &= \tilde{\mathfrak{H}} \text{ (by Equation (3.6)).} \end{aligned}$$

We have, using the mapping  $\tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P}(\mathbb{P})$ ,

$$\begin{aligned} \lim_{V \rightarrow \infty} \tilde{\mathfrak{P}}_{V,1} &= \lim_{V \rightarrow \infty} \tilde{\mathfrak{P}}(\rho_V) \text{ (by Theorem 22)} \\ &= \tilde{\mathcal{F}}\left(\lim_{V \rightarrow \infty} \rho_V\right) \text{ (since } \tilde{\mathcal{F}} : \Omega \rightarrow \mathbb{P}(\mathbb{P}) \text{ is continuous by Theorem 7)} \end{aligned}$$

$$\begin{aligned}
&= \tilde{\mathcal{F}}(\rho) \text{ (by Theorem 12)} \\
&= \tilde{\mathfrak{P}} \text{ (by Equation (3.7)).}
\end{aligned}$$

□

**5.3. Fractal Dimensions.** Here we quantify and compare the Hausdorff dimensions of fractals corresponding to a (super) IFS of similitudes on  $\mathbb{R}^K$  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  $f_m^n(x) = s_m^n O_m^n x + t_m^n$  where  $O_m^n$  is an orthonormal transformation,  $s_m^n \in (0, 1)$ , and  $t_m^n \in \mathbb{R}^K$ , for all  $n \in \{1, 2, \dots, N\}$  and  $m \in \{1, 2, \dots, M\}$ .

**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

$$\sum_{m=1}^M (s_m^1)^D = 1.$$

**Example 8.** Suppose  $K \geq 2$ . Let the IFS  $F^1$  consists of three similitudes with  $s_1^1 = s_2^1 = s_3^1 = \frac{1}{2}$  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  $3 \frac{1}{2^{D_1}} = 1$  which implies  $D_1 = \frac{\ln 3}{\ln 2} = 1.585$ .

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

**5.3.2. Random Fractals.** By Theorem 8 the Hausdorff dimension  $D_R$  of  $\mathfrak{P}$ -almost all of the random fractals sets for the superIFS  $\mathcal{F}$  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  $\mathfrak{P}$ -almost all of the random fractals in the set is given by  $\frac{1}{2} 3 \frac{1}{2^{D_R}} + \frac{1}{2} 3 \frac{1}{3^{D_R}} = 1 \implies D_R = 1.262$ .

**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.** [15]. *Let the superIFS  $\mathcal{F}$  be as specified as in Theorem 8. Let  $V = 1$ . Then for  $\mathfrak{P}_{1,1}$  almost all  $A \in \mathfrak{H}_{1,1}$*

$$\dim_H A = D$$

where  $D$  is the unique solution of

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

**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  $\frac{1}{2} \log(3 \frac{1}{2^D}) + \frac{1}{2} \log(3 \frac{1}{3^D}) = 0$ ,  $\implies d = 2 \log 3 / (\log 2 + \log 3) = 1.226$ .

5.3.4. *V-Variable Fractals* ( $V \geq 1$ ). Let  $(a_1, a_2, \dots) \in \mathcal{A}^\infty$  denote an i.i.d. sequence of indices, with probabilities  $\{\mathcal{P}^a | a \in \mathcal{A}\}$  given in terms of the probabilities  $\{P_1, P_2, \dots, P_V\}$  according to Equation (4.6). Define, for  $\alpha \in [0, \infty)$  and  $a \in \mathcal{A}$ , the  $V \times V$  flow matrix

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

and let us write

$$M_{v,w}^k = M_{v,w}^k(\alpha) = M_{v,w}^{a_k}(\alpha).$$

We think of  $s_m^{n_v}$  as being the “flow” through the  $m^{\text{th}}$  channel from screen  $v$  to screen  $w$  where  $v_{v,m} = w$ . The sequence of random matrices  $M_{v,w}^1, M_{v,w}^2, \dots$  is i.i.d., again with probabilities induced from  $\{P_1, P_2, \dots, P_V\}$ . 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 [13],

$$\gamma(\alpha) := \lim_{k \rightarrow \infty} k^{-1} \log \|M^1(\alpha) \circ \dots \circ M^k(\alpha)\|$$

exists and has the same value with probability one. Provided that the superIFS obeys the open set condition, we have shown in [6] that the unique value of  $D \in [0, \infty)$  such that

$$\gamma(D) = 0$$

is the Hausdorff dimension of  $\mathfrak{P}_{V,1}$  almost all  $A \in \mathfrak{H}_{V,1}$ .

Kingman remarks that in general the calculation of  $\gamma$  “has pride of place among the unsolved problems of subadditive ergodic theory”, [20], p.897. However it is possible to estimate numerically. Namely, generate random copies of  $M^k$  and iteratively compute  $M^1, M^2, \dots, M^k$  and hence  $k^{-1} \log \|M^1(\alpha) \circ \dots \circ M^k(\alpha)\|$  for  $k = 1, 2, \dots$ . 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 geology: particle size distribution in soil, and landscape habitat diversity. In computer science: digital image compression and watermarking, compression of digital audio signals, image segmentation, and computer graphics. In engineering: wavelets, stochastic processes, rough surfaces, antennae and frequency selective surfaces, stochastic optimal control, signal processing, and fragmentation of thin plates.

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 gasses 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 [26], Chapter 9. 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  $F^1 = \{\square; f_1^1, f_2^1, f_3^1\}$  consists of affine maps whose actions we explain with the aid of the left-hand diagram in Figure 25.  $\square$  is the unit square in the diagram, while  $f_1^1(\square)$  is the lower left square,  $f_2^1(\square)$  is the upper left square, and  $f_3^1(\square)$  is the rectangle on the right. The transformations are chosen so that  $f_1^1(\overline{OC}) = \overline{OA}$ ,  $f_2^1(\overline{OC}) = \overline{AB}$ , and  $f_3^1(\overline{OC}) = \overline{BC}$ . Specifically  $f_1^1(x, y) = (\frac{1}{2}y, \frac{1}{2}x)$ ,  $f_2^1(x, y) = (-\frac{1}{2}y + \frac{1}{2}, -\frac{1}{2}x + 1)$ ,  $f_3^1(x, y) = (\frac{1}{2}x + \frac{1}{2}, -y + 1)$ .

The IFS  $F^2 = \{\square; f_1^2, f_2^2, f_3^2\}$  is explained with the aid of the right-hand diagram in Figure 25;  $f_1^2(\square)$  is the lower left rectangle,  $f_2^2(\square)$  is the upper left rectangle, and  $f_3^2(\square)$  is the rectangle on the right; such that  $f_1^2(\overline{OC}) = \overline{OA'}$ ,  $f_2^2(\overline{OC}) = \overline{A'B'}$ , and  $f_3^2(\overline{OC}) = \overline{B'C}$ . Specifically  $f_1^2(x, y) = (\frac{2}{3}y, \frac{1}{2}x)$ ,  $f_2^2(x, y) = (-\frac{2}{3}y + \frac{2}{3}, -\frac{1}{2}x + 1)$ ,  $f_3^2(x, y) = (\frac{1}{3}x + \frac{2}{3}, -y + 1)$ .

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  $\overline{OC}$  is chosen on both screens, and the random iteration algorithm is applied; typical images produced after five iterations are illustrated in Figure 26; an image produced after seven iterations is shown in Figure 27. Each of these images consists of line segments that have been assigned colours 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 [3].

**Example 12.** Here  $N=2, V=2, M=4$ . The two IFSs are given by

$$F^n = \{\square; f_1^n, f_2^n, f_3^n, f_4^n; \} \quad n \in \{1, 2\},$$

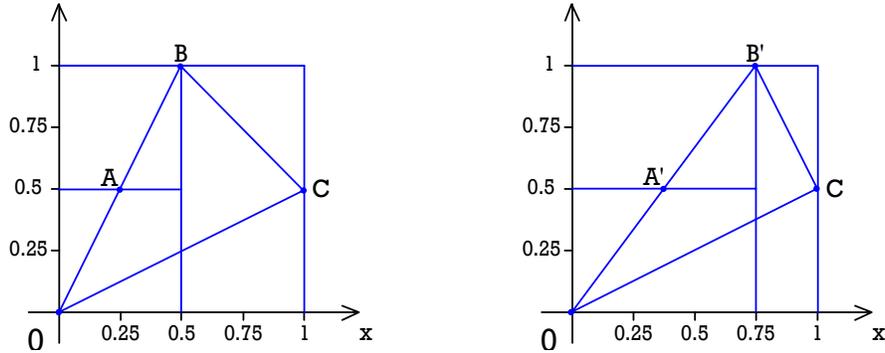


FIGURE 25. Transformations used for space-filling curves, see Example 11

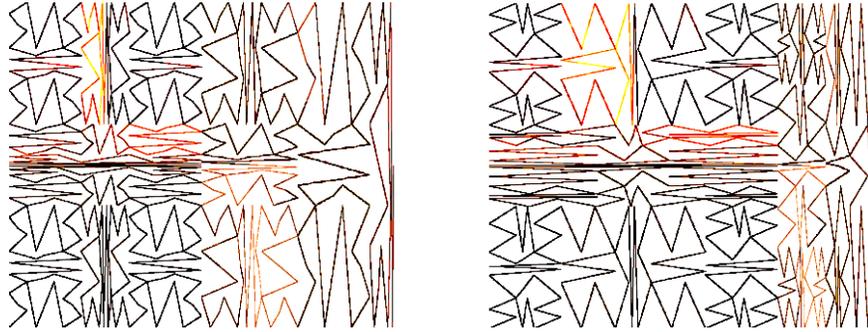


FIGURE 26. Low-order approximants to two 2-variable space filling curves, belonging to the same superfractal, see Example 11

where  $\square \subset \mathbb{R}^2$ , and each  $f_m^n : \square \rightarrow \square$  is a projective transformation. The colours were obtained as follows. A computer graphics rendering of the set attractor of  $F^1$  is shown in Figure 28, and of  $F^2$  in Figure 29.

The colouring of each of these two figures was obtained with the aid of an auxiliary IFS acting on the cube  $C := [0, 255]^3 \subset \mathbb{R}^3$  given by  $\mathcal{G} := \{C; g_1^n, g_2^n, g_3^n, g_4^n\}$  where each  $g_m$  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  $\mathcal{G}$  were calculated via the deterministic algorithm (Corollary 1); each pixel on the attractor of the IFS  $F^n$  was assigned the colour whose red, green, and blue components, each an integer from 0 to 255, were the three coordinates of the point on the attractor of  $\mathcal{G}$  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

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

with  $V = 2$ . Then Figures 30 and 31 show two examples, from among many different but similar ones, all equally visually complex, of computer graphics of 2-variable

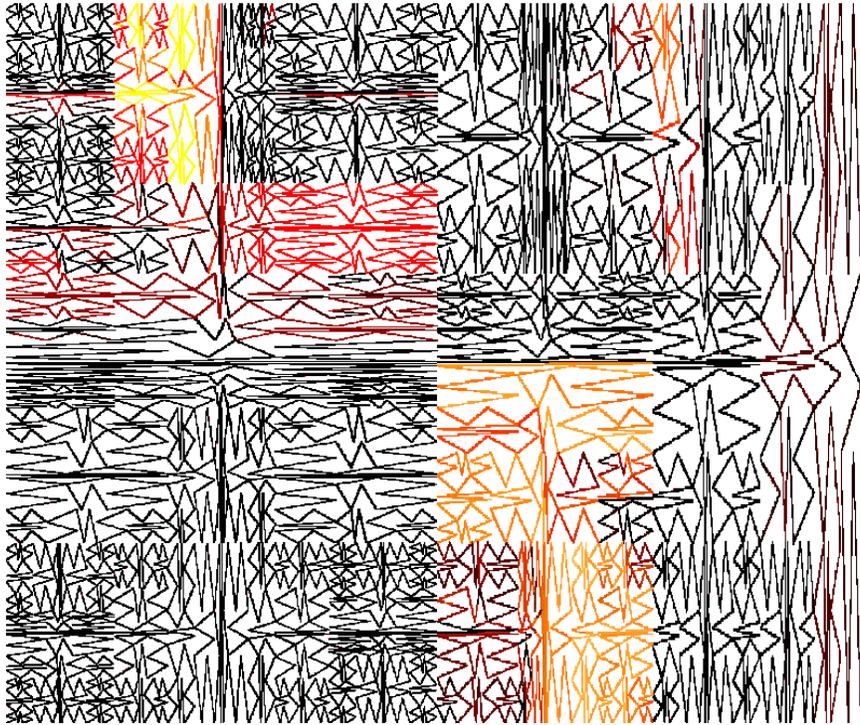


FIGURE 27. Finite-resolution approximation to a 2-variable space-filling curve. See Example 11.

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

$$\lim_{k \rightarrow \infty} f_2^1 \circ f_1^2 \circ f_2^2 \circ f_1^1 \circ f_2^2 \circ f_2^1 \circ \dots \circ f_{m_k}^{n_k}(x)$$

corresponds to the address  $212122\dots m_k\dots$ , in the obvious notation.

**6.3. V-variable Fractal Interpolation.** The technique of fractal interpolation has many applications including modelling 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  $\{(x_i, y_i) \in \mathbb{R}^2 | i = 0, 1, \dots, I\}$  be given. It is desired to find a continuous function  $f : [x_0, x_I] \rightarrow \mathbb{R}$  such that  $f(x_i) = y_i \forall i \in \{0, 1, \dots, M\}$ , 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

$$F = \{\mathbb{R}^2; f_1, f_2, \dots, f_M\}$$

with

$$f_m(x, y) = (a_m x + e_m, c_m x + d_m y + g_m),$$



FIGURE 28. The rendered set attractor of the IFS  $F^1$  in Example 12.



FIGURE 29. The rendered set attractor of the IFS  $F^2$  in Example 12.



FIGURE 30. A 2-variable fractal set for the superIFS  $\mathcal{F}$  in Example 12.

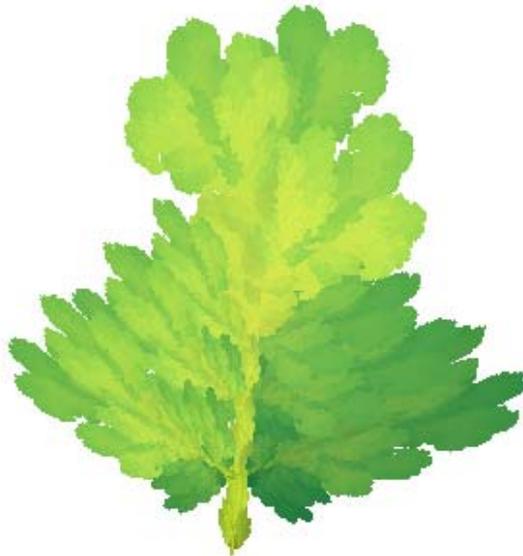


FIGURE 31. Another 2-variable fractal set for the superIFS  $\mathcal{F}$  in Example 12.

where the real coefficients  $a_m, e_m, c_m, d_m$  and  $e_m$  are chosen so that

$$f_m(x_0, y_0) = y_{m-1}, f_m(x_0, y_0) = y_m,$$

and  $d_m \in [0, 1)$ , for  $m \in \{1, 2, \dots, M\}$ . 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  $\{d_m : m = 1, 2, \dots, M\}$ .

Now let the superIFS  $\mathcal{F} = \{\square; F^1, F^2; P^1 = 0.5, P^2 = 0.5\}$  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.

## REFERENCES

- [1] M. A. Arbeiter *Random Recursive Constructions of Self-Similar Fractal Measures, the Non-compact Case*, Probab. Theory Related Fields, 88(1991), pp. 497-520.
- [2] M. T. Barlow and B. M. Hambly, Ann. Inst. H. Poincaré Probab. Statist. 33(1997), 531-537.
- [3] M. F. Barnsley, *Fractals Everywhere*, Academic Press, New York, NY, 1988.
- [4] M. F. Barnsley and S. Demko, *Iterated Function Systems and the Global Construction of Fractals*, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. 399(1985), pp. 243-275.
- [5] M. F. Barnsley, A. Deliu and R. Xie, *Stationary Stochastic Processes and Fractal Data Compression*, Int. J. of Bifurcation and Chaos, 7(1997), pp. 551-567.
- [6] M. F. Barnsley, J. E. Hutchinson and Ö. Stenflo, Supporting notes, and paper in preparation.
- [7] P. Billingsley, *Ergodic Theory and Information*, John Wiley, New York, NY, 1968.
- [8] P. Diaconis and D. Freedman, *Iterated Random Functions*, SIAM Rev., 41(1999), pp. 45-76.
- [9] R. M. Dudley, *Real Analysis and Probability*, Wadsworth, New York, NY, 1989.
- [10] J. Elton, *An Ergodic Theorem for Iterated Maps*, Ergodic Theory Dynam. Systems, 7(1987), pp. 481-488.
- [11] K. J. Falconer, *Random Fractals*, Math. Proc. Cambridge Philos. Soc., 100(1986), pp. 559-582.
- [12] K. Falconer, *Fractal Geometry - Mathematical Foundations and Applications*, John Wiley & Sons, Ltd., Chichester, England, 1990.
- [13] H. Furstenberf and H. Kesten, *Products of Random Matrices*, Ann. Math. Stat., 31(1960), pp. 457-469.
- [14] S. Graf, *Statistically Self-Similar Fractals*, Probab. Theory Related Fields, 74(1987), pp. 357-392.
- [15] B. M. Hambly, Ann. Probab. 25 (1997), 1059-1102.
- [16] J. E. Hutchinson, *Fractals and Self-Similarity*, Indiana. Univ. Math. J., 30 (1981), pp. 713-749
- [17] J. E. Hutchinson and L. Rüschemdorf, *Random Fractal Measures via the Contraction Method*, Indiana Univ. Math. J., 47(1998), pp. 471-487.
- [18] J. E. Hutchinson and L. Rüschemdorf, *Random Fractals and Probability Metrics*, Adv. in Appl. Probab., 32(2000), pp. 925-947.
- [19] Y. Kifer, *Fractals via Random Iterated Function Systems and Random Geometric Constructions*, in Fractal Geometry and Stochastics (Finsbergen, 1994), Progr. Probab. 37(1995), pp. 145-164.
- [20] J. F. C. Kingman, *Subadditive Ergodic Theory*, Ann. Probab. 1(1973), pp.883-909.

- [21] R. D. Mauldin and S. C. Williams, *Random Recursive Constructions; Asymptotic Geometrical and Topological Properties*, Trans. Amer. Math. Soc., 295(1986), pp. 325-346.
- [22] P. A. P. Moran, *Additive Functions of Intervals and Hausdorff Measure*, Proc. Cambridge Philos. Soc., 42(1946), pp.15-23.
- [23] L. Olsen, *Random Geometrically Graph Directed Self-Similar Multifractals*, Pitman Research Notes, 307(1994), Longman, Harlow.
- [24] N. Patzschke and U. Zähle, *Self-Similar Random Measures. IV. The Recursive Construction Model of Falconer, Graf, and Mauldin and Williams*, Math. Nachr. 149(1990), pp. 285-302.
- [25] M. Peruggia, *Discrete Iterated Function Systems*, A. K. Peters, Wellesley, MA, 1993.
- [26] H. Sagan, *Space-Filling Curves*, Springer-Verlag, New York, N.Y., 1994.
- [27] Ö. Stenflo, *Markov Chains in Random Environments and Random Iterated Function Systems*, Trans. Amer. Math. Soc. 353(2001), 3547-3562.
- [28] U. Zähle, *Self-Similar Random Measures. I. Notion, Carrying Hausdorff Dimension, and Hyperbolic Distribution*, Probab. Theory Related Fields, 80(1988), pp. 79-100.

335 PENNBROOKE TRACE, DULUTH,, GA 30097, USA  
E-mail address: Mbarnsley@aol.com

DEPARTMENT OF MATHEMATICS, AUSTRALIAN NATIONAL UNIVERSITY  
E-mail address: john.hutchinson@anu.edu.au

DEPARTMENT OF MATHEMATICS, STOCKHOLM UNIVERSITY, SE-10691 STOCKHOLM, SWEDEN  
E-mail address: stenflo@math.su.se