Theory and Applications of Fractal Tops

Michael Barnsley

Australian National University, Canberra
`mbarnsley@aol.com`

Summary. We consider an iterated function system (IFS) of one-to-one contractive maps on a compact metric space. We define the **top** of an IFS; define an associated symbolic dynamical system; present and explain a fast algorithm for computing the top; describe an example in one dimension with a rich history going back to work of A.Rényi [*Representations for Real Numbers and Their Ergodic Properties*, Acta Math. Acad. Sci. Hung.,**8** (1957), pp. 477-493]; and we show how tops may be used to help to model and render synthetic pictures in applications in computer graphics.

1 Introduction

It is well-known that an iterated function system (IFS) of 1-1 contractive maps, mapping a compact metric space into itself, possesses a set attractor and various invariant measures. But it also possesses another type of invariant object which we call a *top*. One application of tops is to modelling and rendering new families of synthetic pictures in computer graphics. Another application is to information theory and data compression. Tops are mathematically fascinating because they have a rich symbolic dynamics structure, they support intricate Markov chains, and they provide examples of IFS with place-dependent probabilities in a regime where not much research has taken place.

In this paper we define the top of an IFS; define an associated symbolic dynamical system; present and explain a fast algorithm for computing the top; describe an example in one dimension with a rich history going back to work of A.Rényi [11]; and we show how tops may be used to help to model and render synthetic pictures in applications in computer graphics.

This is a short version of a paper, [5], which includes proofs and more detail. This work was supported by the Australian Research Council.

The author thanks John Hutchinson for many useful discussions and much help with this work. The author thanks Louisa Barnsley for editorial help and for producing the graphics. The author thanks a referee for helpful comments.

2 The Top of an IFS

Let an iterated function system (IFS) be denoted

$$\mathcal{W} := \{\mathbb{X}; w_0, ..., w_{N-1}\}. \tag{1}$$

This consists of a finite of sequence of one-to-one contraction mappings

$$w_n : \mathbb{X} \to \mathbb{X}, n = 0, 2, ..., N - 1 \tag{2}$$

acting on the compact metric space

$$(\mathbb{X}, d) \tag{3}$$

with metric d so that for some

$$0 \leq l < 1 \tag{4}$$

we have

$$d(w_n(x), w_n(y)) \leq l \cdot d(x, y) \tag{5}$$

for all $x, y \in \mathbb{X}$.

Let A denote the attractor of the IFS, that is $A \subset \mathbb{X}$ is the unique non-empty compact set such that

$$A = \bigcup_n w_n(A).$$

Let the associated code space be denoted by $\Omega = \Omega_{\{0,1,...,N-1\}}$. This is the space of infinite sequences of symbols $\{\sigma_i\}_{i=1}^{\infty}$ belonging to the alphabet $\{0, 1, ..., N - 1\}$ with the discrete product topology. We will also write $\sigma = \sigma_1\sigma_2\sigma_3... \in \Omega$ to denote a typical element of Ω, and we will write ω_k to denote the kth element of the sequence $\omega \in \Omega$. We order the elements of Ω according to

$$\sigma < \omega \text{ iff } \sigma_k < \omega_k$$

where k is the least index for which $\sigma_k \neq \omega_k$.

Let

$$\phi : \Omega \to A$$

denote the associated *continuous* addressing map *from* code space *onto* the attractor of the IFS. We note that the set of addresses of a point $x \in A$, defined to be $\phi^{-1}(x)$, is compact and so possesses a unique largest element. We denote this value by $\tau(x)$. That is, $\tau : A \to \Omega$ is defined by

$$\tau(x) = \max\{\sigma \in \Omega : \phi(\sigma) = x\}.$$

We call τ the *tops function* of the IFS. We also call $G_\tau := \{(x, \tau(x)) : x \in A\}$ the graph of the top of the IFS or simply *the top* of the IFS.

The top of an IFS may be described as follows: consider the *lifted* IFS

$$\{\mathbb{X} \times \Omega : W_0, W_1, ..., W_{N-1}\}$$

where

$$W_n(x, \sigma) = (w_n(x), n\sigma)$$

where, for the avoidance of any doubt, $n\sigma := \omega$ where $\omega_1 = n$ and $\omega_{n+1} = \sigma_n$ for $n = 0, 1, ..., N - 1$. (The metric d_Ω on Ω is defined, for all $\omega, \sigma \in \Omega$, by $d_\Omega(\omega, \sigma) = 0$ when $\omega = \sigma$, and otherwise $d_\Omega(\omega, \sigma) = \frac{1}{2^k}$ where k is the least integer for which $\sigma_k \neq \omega_k$.) Let the unique attractor of this IFS be denoted by $\widehat{A}$. Then the projection of $\widehat{A}$ on the $\mathbb{X}$-direction is $\mathbb{X}$, and in the Ω-direction it is Ω. The top of the original IFS is related to $\widehat{A}$ according to:

$$G_\tau = \{(x, \sigma) \in \widehat{A} : (x, \omega) \in \widehat{A} \Longrightarrow \omega \leq \sigma\}.$$

This latter formulation is useful because we can use the chaos game algorithm (also called a Markov Chain Monte Carlo (MCMC) algorithm or a random iteration algorithm) to compute approximations to $\widehat{A}$ and hence to G_τ. According to this method we select a sequence of symbols

$$\sigma_1 \sigma_2 \sigma_3 ... \in \{1, 2, ..., N\}^\infty$$

with probability $p_n > 0$ for the choice $\sigma_k = n$, independent of all of the other choices. We also select $X_0 \in \mathbb{X}$ and let

$$X_{n+1} = W_{\sigma_{n+1}}(X_n) \text{ for } n = 0, 1, 2, ...$$

Then, almost always,

$$\lim_{k \to \infty} \overline{\{X_n : n = k, k + 1, ...\}} = \widehat{A}$$

where $\overline{S}$ denotes the closure of the set S. (The value of the limit is intersection of the decreasing sequence of compact sets whose limit is being taken.) This algorithm provides in many cases a simple efficient fast method to compute approximations to the attractor of an IFS, for example when $\mathbb{X} = \square$, a compact subset of $\mathbb{R}^2$. By keeping track of points which, for each approximate value of $x \in \mathbb{X}$, have the greatest code space value, we can compute approximations to G_τ. We illustrate this approach in the following example which we continue in Section 6.

Example 1. Consider the IFS

$$\{[0, 1] \subset \mathbb{R}; w_0(x) = \alpha x, w_2(x) = \alpha x + (1 - \alpha)\} \tag{6}$$

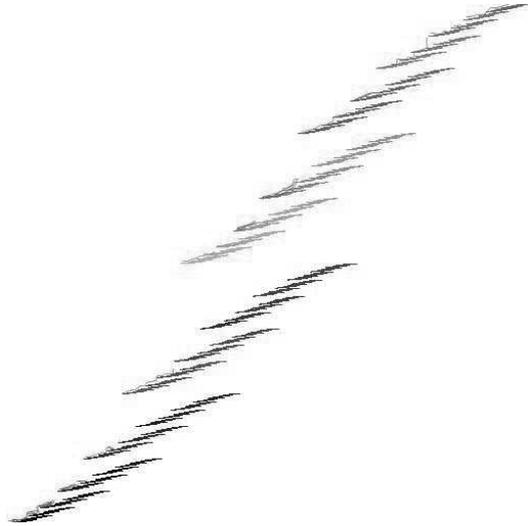We are interested in the case where

$$\frac{1}{2} < \alpha < 1,$$

Fig. 1. The attractor of the IFS in Equation 7. This represents the attractor $\widehat{A}$ of the lifted IFS corresponding to Equation 6. The top of the IFS is indicated in red. The visible part of the "x-axis" represents the real interval $[0, 1]$ and the visible part of the "y-axis" represents the code space Ω between the points 000000000.... and 111111111.....

which we refer to as "overlapping" because $w_0([0,1]) \cap w_1([0,1])$ contains a non-empty open set. In Figure 1 we show the attractor $\widehat{A}$ of the associated lifted IFS, and upon this attractor we have indicated the top with some red squiggles. Figure 1 was computed using random iteration: we have represented points in code space by their binary expansions which are interpreted as points in $[0, 1]$. Since the invariant measure of both the IFS and the lifted IFS contain no atoms, the information lost by this representation is irrelevant to pictures. Accordingly, the actual IFS used to compute Figure 1 is

$$\{[0,1]\times[0,1] \subset \mathbb{R}; W_0(x,y) = (\alpha x, \frac{1}{2}y), W_2(x,y) = (\alpha x+(1-\alpha), \frac{1}{2}y+\frac{1}{2})\} \quad (7)$$

with $\alpha = \frac{2}{3}$.

3 Application of Tops to Computer Graphics

Here we introduce the application of tops to computer graphics. There is a great deal more to say about this, but to serve as motivation as well as to provide an excellent method for graphing fractal tops, we explain the basic idea here.

A **picture function** is a mapping $\mathfrak{P} : D_{\mathfrak{P}} \subset \mathbb{R}^2 \rightarrow \mathfrak{C}$ where $\mathfrak{C}$ is a colour space, for example $\mathfrak{C} = [0, 255]^3 \subset \mathbb{R}^3$. The domain $D_{\mathfrak{P}}$ is typically a rectangular subset of $\mathbb{R}^2$: we often take

$$D_{\mathfrak{P}} = \square := \{(x, y) \in \mathbb{R}^2 : 0 \leq x, y \leq 1\}.$$

The domain of a picture function is an important part of its definition; for example a segment of a picture may be used to define a picture function. A picture in the usual sense may then be thought of as the graph of a picture function. But we will use the concepts of picture, picture function, and graph of a picture function interchangeably. We do not discuss here the important questions of how such functions arise in image science, for example, nor about the relationship between such abstract objects and real world pictures. Here we assume that given a picture function, we have some process by which we can render it to make pictures which may be printed, viewed on computer screens, etc. This is far from a simple matter in general.

Let two IFS's

$$\mathcal{W} := \{\square; w_0, ..., w_{N-1}\} \text{ and } \widetilde{\mathcal{W}} := \{\square; \widetilde{w}_0, ..., \widetilde{w}_{N-1}\}$$

and a picture function

$$\widetilde{\mathfrak{P}} : \square \rightarrow \mathfrak{C}$$

be given. Let A denote the attractor of the IFS $\mathcal{W}$ and let $\widetilde{A}$ denote the attractor of the IFS $\widetilde{\mathcal{W}}$. Let

$$\tau : A \rightarrow \Omega$$

denote the tops function for $\mathcal{W}$. Let

$$\widetilde{\phi} : \Omega \rightarrow \widetilde{A} \subset \square$$

denote the addressing function for the IFS $\widetilde{\mathcal{W}}$. Then we define a new picture function

$$\mathfrak{P} : A \rightarrow \mathfrak{C}$$

by

$$\mathfrak{P} = \widetilde{\mathfrak{P}} \circ \widetilde{\phi} \circ \tau.$$

This is the **unique picture function** defined by the IFS's $\mathcal{W}$, $\widetilde{\mathcal{W}}$, and the picture $\widetilde{\mathfrak{P}}$. We say that it has been **produced by tops + colour stealing**. We think in this way: colours are "stolen" from the picture $\widetilde{\mathfrak{P}}$ to "paint" code space; that is, we make a code space picture, that is the function $\widetilde{\mathfrak{P}} \circ \widetilde{\phi} : \Omega \rightarrow \mathfrak{C}$, which we then use together with top of $\mathcal{W}$ to paint the attractor A.

Notice the following points. (i) Picture functions have properties that are determined by their source; digital pictures of natural scenes such as clouds and sky, fields of flowers and grasses, seascapes, thick foliage, etc. all have their own distinctive palettes, relationships between colour and position, "continuity" and "discontinuity" properties, and so on. (ii) Addressing functions are

Fig. 2. Colours were stolen from this picture to produce Figure 4 and the right-hand image in Figure 3.

continuous. (iii) Tops functions have their own special properties; for example they are continuous when the associated IFS is totally disconnected, and they contain the geometry of the underlying IFS attractor A plus much more, and so may have certain self-similarities and, assuming the IFS are built from low information content transformations such as similitudes, possess their own harmonies. Thus, the picture functions produced by tops plus colour steal- ing may define pictures which are interesting to look at, carrying a natural palette, possessing certain continuities and discontinuities, and also certain self-similarities. There is much more one could say here.

The stolen picture $\widetilde{\mathfrak{P}} \circ \widetilde{\phi} \circ \tau$ may be computed by random iteration, by coupling the lifted IFS associated with $\mathcal{W}$ to the IFS $\widetilde{\mathcal{W}}$. This is the method used until recently, and has been described in [3] and in [4]. Recently we have discovered a much faster algorithm for computing the entire stolen picture at a given resolution. It is based on a symbolic dynamical system associated with the top of $\mathcal{W}$. We describe this new method in Section 5.

In Figure 3 we illustrate the attractor, and invariant measure, and a picture defined by tops + colour stealing, all for the IFS of projective transformations

$$\mathcal{W} = \{\Box; w_n(x,y) = (\frac{a_n x + b_n y + c_n}{g_n x + h_n y + j_n}, \frac{d_n x + e_n y + f_n}{g_n x + h_n y + j_n}), n = 0, 1, 2, 3\} \quad (8)$$

where the coefficients are given by

n	a_n	b_n	c_n	d_n	e_n	f_n	g_n	h_n	j_n
0	1.901	-0.072	0.186	0.015	1.69	0.028	0.563	-0.201	2.005
1	0.002	-0.044	0.075	0.003	-0.044	0.104	0.002	-0.088	0.154
2	0.965	-0.352	0.058	1.314	-0.065	-0.191	1.348	-0.307	0.075
3	-0.325	-0.0581	-0.029	-1.229	-0.001	0.199	-1.281	0.243	-0.058

The picture from which the colours were stolen is shown in Figure 2.

A close-up on the tops picture is illustrated in Figure 4.

4 The Tops Dynamical System

We show how the IFS leads to a natural dynamical system $T : G_\tau \to G_\tau$. The notation is the same as above. We also need the shift mapping

Fig. 3. From left to right, the attractor, an invariant measure, and a picture of the top made by colour stealing, for the IFS in Equation 8. Figure 4 shows a zoom on the picture of the top.

Fig. 4. Close-up on the fractal top in Figure 3. Details of the structure are revealed by colour-stealing.

$$S : \Omega \to \Omega$$

defined by

$$S\sigma_1\sigma_2\sigma_3... = \sigma_2\sigma_3...$$

for all $\sigma = \sigma_1\sigma_2\sigma_3.. \in \Omega$.

Lemma 1. *Let $(x,\sigma) \in G_\tau$. Then $(w_{\sigma_1}^{-1}(x), S\sigma) \in G_\tau$.*

Proof. See [5].

Lemma 2. *Let $(x,\sigma) \in G_\tau$. Then there is $(y,\omega) \in G_\tau$ such that $(w_{\sigma_1}^{-1}(y), S\omega) = (x,\sigma) \in G_\tau$.*

Proof. See [5].

It follows that the mapping

$$T : G_\tau \to G_\tau \text{ defined by } T(x,\sigma) = (w_{\sigma_1}^{-1}(x), S\sigma)$$

is well-defined, and onto. It can be treated as a dynamical system which we refer to as $\{G_\tau, T\}$. As such we may explore its invariant sets, invariant measures, other types of invariants such as entropies and information dimensions, and its ergodic properties, using "standard" terminology and machinery.

We can we project $\{G_\tau, T\}$ onto the Ω-direction, as follows: let

$$\Omega_\gamma := \{\sigma \in \Omega : (x,\sigma) \in G_\tau \text{ for some } x \in \mathbb{X}\}$$

Then Ω_γ is a shift invariant subspace of Ω, that is $S : \Omega_\gamma \to \Omega_\gamma$ with

$$S(\Omega_\gamma) = \Omega_\gamma,$$

and we see that $\{\Omega_\gamma, S\}$ is a **symbolic dynamical system**, see for example [10].

Indeed, $\{\Omega_\gamma, S\}$ is the symbolic dynamical system corresponding to a partition of the domain of yet a third dynamical system $\{A, \widetilde{T}\}$ corresponding to a mapping $\widetilde{T} : A \to A$ which is obtained by projecting $\{G_\tau, T\}$ onto A. This system is defined by

$$\widetilde{T}(x) = \begin{cases} w_{N-1}^{-1}(x) \ if & x \in D_{N-1} := w_{N-1}(A), \\ w_{N-2}^{-1}(x) \ if \ x \in D_{N-2} := w_{N-2}(A)\backslash w_{N-1}(A) \\ \quad . \qquad\qquad . \qquad\qquad\qquad\qquad . \\ w_0^{-1}(x) \ \ if & x \in D_0 := w_0(A)\backslash \bigcup_{n=1}^{N-1} w_n(A) \end{cases} \tag{9}$$

for all $x \in A$ and we have

$$\widetilde{T}(A) = A.$$

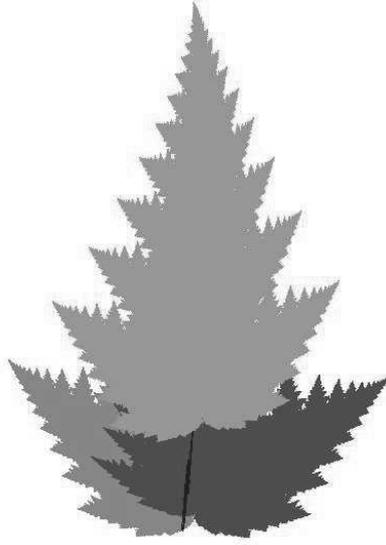We call $\{A, \widetilde{T}\}$ the **tops dynamical system** (TDS) associated with the IFS.

Fig. 5. Illustrates the domains D_0, D_1, D_2, D_3 for the tops dynamical system associated with the IFS in Equation 8. This was the IFS used in Figures 3 and 4. Once this "picture" has been computed it is easy to compute the tops function. Just follow orbits of the tops dynamical system!

$\{\Omega_\gamma, S\}$ is the symbolic dynamical system obtained by starting from the tops dynamical system $\{A, T\}$ and partitioning A into the disjoint sets $D_0, D_1, ..., D_{N-1}$ defined in Equation 9, where

$$A = \bigcup_{n=0}^{N-1} D_n \text{ and } D_i \cap D_j = \varnothing \text{ for } i \neq j.$$

An example of such a partition is illustrated in Figure 5. We refer to $\{\Omega_\gamma, S\}$ as the **symbolic tops dynamical system** associated with the original dynamical system.

Theorem 1. *The tops dynamical system $\{A, \widetilde{T}\}$ and the symbolic dynamical system $\{\Omega_\gamma, S\}$ are conjugate. The identification between them is provided by the tops function $\tau : A \to \Omega_\gamma$. That is,*

$$\widetilde{T}(x) = \phi \circ S \circ \tau(x)$$

for all $x \in A$, and μ is an invariant measure for $\{A, \widetilde{T}\}$ iff $\tau \circ \mu$ is an invariant measure for $\{\Omega_\gamma, S\}$.

Proof. Follows directly from everything we have said above.

5 Symbolic Dynamics Algorithm for Computing Tops

Corollary 1. *The value of $\tau(x)$ may be computed by following the orbit of $x \in A$ as follows. Let $x_1 = x$ and $x_{n+1} = \widetilde{T}(x_n)$ for $n = 1, 2, ...,$ so that the orbit of x is $\{x_n\}_{n=1}^{\infty}$. Then $\tau(x) = \sigma$ where $\sigma_n \in \{0, 1, ..., N-1\}$ is the unique index such that $x_n \in D_n$ for all $n = 1, 2, ...$*

Corollary 1 provides us with a delightful algorithm for computing approximations to the top of an IFS in cases in which we are particularly interested, namely when the IFS acts in $\mathbb{R}^2$ as in Equation 8.

Here we describe briefly one of many possible variants of the algorithm, with concentration on the key idea.

(i) Fix a resolution $L \times M$. Set up a rectangular array of little rectangular boxes, "pixels", corresponding to a rectangular region in $\mathbb{R}^2$ which contains the attractor of the IFS. Each box "contains" either a null value or a floating point number $x_{l,m} \in \mathbb{R}^2$ and a finite string of indexes $\omega_{l,m} = \sigma_{l,m}^1 \sigma_{l,m}^2 ... \sigma_{l,m}^{L_{m,n}}$ where $L_{l,m}$ denotes the length of the string and each $\sigma_{l,m}^k \in \{0, 1, ..., N-1\}$. The little boxes are initialized to null values.

(ii) Run the standard random iteration algorithm applied to the IFS with appropriate choices of probabilities, for sufficiently many steps to ensure that it has "settled" on the attractor, then record in all of those boxes, which taken together correspond to a discretized version of the attractor, a representative floating point value of $x \in A$ and the highest value, so far encountered, of the map index corresponding to that x-value. This requires that one runs the algorithm for sufficiently many steps that each box is visited many times and that, each time a map with a higher index value than the one recorded in a visited little box, the index value at the box is replaced by the higher index. [The result will be a discretized "picture" of the attractor, defined by the boxes with non null entries, partitioned into the domains $D_0, D_1, .., D_{N-1}$ with a high resolution value of $x \in A$ for each pixel. We will use these high resolution values to correct at each step the approximate orbits of $\widetilde{T} : A \to A$ which otherwise would rapidly loose precision and "leave" the attractor.]

(iii) Choose a little rectangular box, indexed by say l_1, m_1, which does not have a null value. (If the value of $\tau(x_{l_1,m_1})$, namely the string ω_{l_1,m_1}, is already recorded in the little rectangular box to sufficient precision, that is $L_{l_1,m_1} = L$, say go to another little rectangular box until one is found for which $L_{l_1,m_1} = 1$.) Keep track of $l_1, m_1, \sigma_{l_1,m_1}$. Compute $w_{\sigma_{l_1,m_1}}(x_{l_1,m_1})$ then discretize and identify the little box l_2, m_2 to which it belongs. If $L_{l_2,m_2} = L$ then set

$$\omega_{l_1,m_1} = \sigma_{l_1,m_1}^1 \sigma_{l_2,m_2}^1 \sigma_{l_2,m_2}^2 ... \sigma_{l_2,m_2}^{L-1}$$

and go to (iv). If $l_2, m_2 = l_1, m_1$ set

$$\omega_{l_1,m_1} = \sigma_{l_1,m_1}^1 \sigma_{l_1,m_1}^1 \sigma_{l_1,m_1}^1 ... \sigma_{l_1,m_1}^1$$

and go to (iv). Otherwise, keep track of $l_1, m_1, \sigma_{l_1,m_1}; l_2, m_2, \sigma_{l_2,m_2}$ and repeat the iterative step now starting at l_2, m_2 and computing $w_{\sigma_{l_2,m_2}}(x_{l_2,m_2})$.

Continue in this manner until *either* one lands in a box for which the string value is already of length L, in which case one back-tracks along the orbit one has been following, filling in all the string values up to length L, *or* until the sequence of visited boxes first includes the address of one box twice; i.e. a discretized periodic orbit is encountered. The strings of all of the points on the periodic orbit can now be filled-out to length L, and then the strings of all of the points leading to the periodic cycle can be deduced and entered into their boxes.

(iv) Select systematically a new little rectangular box and repeat step (iii), and continue until all the strings $\omega_{l,m}$ have length L. Our final approximation is

$$\tau(x_{l,m}) = \omega_{l,m}.$$

This algorithm includes "pixel-chaining" as described in [9] and is very efficient because only one point lands in each pixel during stages (iii) and (iv).

6 Analysis of Example 1

6.1 Invariant Measures and Random Iteration on Fractal Tops

We are interested in developing algorithms which are able to compute directly the graph G_τ of the tops function by some form of random iteration in which, at every step, the new points remain on G_τ. For while the just-described algorithm is delightful for computing approximations to the whole of G_τ it appears to be cumbersome and to have large memory requirements if very high resolution approximations to a part of G_τ are required, as when one "zooms in on" a fractal. But the standard chaos game algorithm has huge benefits in this regard and we would like to be able to repeat the process here. (The variant of the random iteration algorithm mentioned earlier, where one works on the whole of $\widehat{A}$ and keeps track of "highest values" is clearly inefficient in overlapping cases where the measure of the overlapping region is not negligible. Orbits of points may spend time wandering around deep within $\widehat{A}$ rarely visiting the top!)

But this is not the only motivation for studying stochastic processes and invariant measures on tops. Such processes have very interesting connections to information theory and data compression. In the end one would like to come back, full-circle, to obtain insights into image compression by understanding these processes. This topic in turn relates to IFS's with place-dependent probabilities and to studies concerning when such IFS's possess unique invariant measures.

Here we extend our discussion of Example 1 in some detail, to show the sort of thing we mean. We show that this example is related to a dynamical system studied by A. Renyi [11] connected to information theory. The IFS in

this example also shows up in the context of Bernoulli convolutions, where the absolute continuity or otherwise of its invariant measures is discussed. We present a theorem concerning this example which is, hopefully, new.

Much of what we say may be generalized extensively.

6.2 The TIFS and Markov Process for Example 1

We continue Example 1. The tops IFS (TIFS) associated with the IFS in Equation 6 is the "IFS" made from the following two functions, one of which has as its domain a set that is not all of $[0, 1]$:

$$\widetilde{w}_0(x) = \alpha x \text{ for all } x \in [0, \frac{1 - \alpha}{\alpha}); \tag{10}$$
$$\widetilde{w}_1(x) = \alpha x + (1 - \alpha) \text{ for all } x \in [0, 1].$$

We are interested in invariant measures for the following type of Markov process. This relates to a "chaos game" with place-dependent probabilities. Define a Markov transition probability by

$$\widetilde{P}(x, B) = \widetilde{p}_0(x)\chi_B(\widetilde{w}_0(x)) + \widetilde{p}_1(x)\chi_B(\widetilde{w}_1(x)). \tag{11}$$

Here

$$\widetilde{p}_0(x) = \begin{cases} p_0 \ for \ 0 \leq x < \frac{1}{\alpha} - 1 \\ 0 \ for \ \frac{1}{\alpha} - 1 < x \leq 1 \end{cases} \tag{12}$$

and

$$\widetilde{p}_1(x) = \begin{cases} p_1 \ for \ 0 \leq x < \frac{1}{\alpha} - 1 \\ 1 \ for \ \frac{1}{\alpha} - 1 < x \leq 1 \end{cases} \tag{13}$$

where $p_0, p_1 > 0$ and $p_0 + p_1 = 1$. $\widetilde{P}(x, B)$ is the probability of transfer from $x \in [0, 1]$ into the Borel set B. Intuitively, pick a number $i \in \{0, 1\}$ according to the distribution $\widetilde{p}_i(x)$ and transfer from x to $\widetilde{w}_i(x)$. What types of measures may be generated by such random orbits? When are such measures invariant for the Markov process? A Borel measure $\widetilde{\mu}$ on $[0, 1]$ is said to be invariant for the Markov process iff

$$\widetilde{\mu}(B) = \int \widetilde{P}(x, B) d\widetilde{\mu}(x)$$

for all Borel sets $B \subset [0, 1]$.

6.3 The TDS and Trapping Region for Example 1

The tops dynamical system (TDS) associated with this TIFS in Equation 6 is obtained by "inverting" the TIFS and is readily found to be defined as follows:

$$D_0 = [0, 1 - \alpha), D_1 = [1 - \alpha, 1]$$

and

$$\widetilde{T}(x) = \begin{cases} \frac{1}{\alpha}x & for\ x \in D_0 \\ \frac{1}{\alpha}x - (1 - \frac{1}{\alpha}) & for\ x \in D_1 \end{cases}. \tag{14}$$

Notice that orbits under $\widetilde{T}$ of points in $(\frac{1-\alpha}{\alpha}, 1)$ eventually arrive in the interval $R_{trapping} = [0, \frac{1-\alpha}{\alpha}]$ and that once there, they never leave again. We refer to $R_{trapping}$ as the trapping region. Notice too that $\widetilde{T}(1) = 1$ and that $x = 1$ is a repulsive fixed point of the dynamical system. That is, the tops dynamical system $\widetilde{T} : [0, 1] \to [0, 1]$ admits the invariant measure $\widetilde{\mu} = \delta_1(x)$. Note that this atomic measure is also invariant for the Markov process with transition probability $\widetilde{P}(x, B)$. But the following theorem tells us that there are no other possible atoms than one at $x = 1$ for invariant probability measures for the process.

6.4 Non-atomic invariant measures for Example 1

The proof of the following result typifies an argument which may be used in many cases to establish the existence of non-atomic invariant measures for TIFS.

Theorem 2. *Let $\widetilde{\mu}$ be an invariant probability measure for the Markov process described by Equation 11. Then $\widetilde{\mu}(\{a\}) = 0$ for all $a \in [0, 1)$.*

Proof. See [5].

6.5 The TIFS and TDS for Example 1 in the trapping region

From this point forward in this section we concentrate on the behaviour of the TDS and the TIFS in the trapping region. We study invariant probability measures and ergodic properties for the TIFS and TDS restricted to the trapping region. In view of Theorem 2, we restrict attention to invariant probability measures which do not contain atoms. This will allow us to modify the "trapped" TDS/TIFS on any countable sets of points without altering potential invariant probability measures off a set of measure zero.

We make the change of variable $x' = \frac{1-\alpha}{\alpha}x = g(x)$ to rescale the behaviour of the TDS restricted to $R_{trapping}$ to produce an equivalent dynamical system acting on $[0, 1]$: that is $\widetilde{\widetilde{T}} : [0, 1] \to [0, 1]$ is defined by $\widetilde{\widetilde{T}} = g \circ \widetilde{T} \circ g^{-1}$. The result is

$$\widetilde{\widetilde{D}}_0 = [0, \alpha), \widetilde{\widetilde{D}}_1 = [\alpha, 1]$$

and

$$\widetilde{\widetilde{T}}(x) = \begin{cases} \frac{1}{\alpha}x & for\ x \in \widetilde{\widetilde{D}}_0 \\ \frac{1}{\alpha}x - 1 & for\ x \in \widetilde{\widetilde{D}}_1 \end{cases}. \tag{15}$$

Notice that

$$\widetilde{\widetilde{T}}(x) = (\beta x) \tag{16}$$

where (y) denotes the fractional part of the real number y and

$$\beta = \frac{1}{\alpha}.$$

By using the symbol β we connect our problem to standard notation for a much studied dynamical system, Equation 16, see for example [11], [10]. The focus of all of the work, with which we are familiar, connected to the dynamical system in Equation 15, concerns invariant measures which are absolutely continuous with respect to Lebesgue measure and which maximize topological entropy. This is not our focus. We are interested in the existence and computation of invariant measures for the associated IFS with place-dependent, typically piecewise constant, probabilities.

Next we invert the trapped TDS to produce a corresponding restricted TIFS. This is defined with the aid of the two functions

$$\widetilde{\widetilde{w}}_0(x) = \alpha x \text{ for all } x \in [0, 1); \tag{17}$$

$$\widetilde{\widetilde{w}}_1(x) = \alpha x + \alpha \text{ for all } x \in [0, \frac{1 - \alpha}{\alpha}].$$

The Markov process in the trapped region corresponds to the transition probability

$$\widetilde{\widetilde{P}}(x, B) = \widetilde{\widetilde{p}}_0(x)\chi_B(\widetilde{\widetilde{w}}_0(x)) + \widetilde{\widetilde{p}}_1(x)\chi_B(\widetilde{\widetilde{w}}_1(x)). \tag{18}$$

with the place-dependent probabilities

$$\widetilde{\widetilde{p}}_0(x) = \begin{cases} p_0 \ for \ 0 \le x \le \frac{1}{\alpha} - 1 \\ 1 \ for \ \frac{1}{\alpha} - 1 < x \le 1 \end{cases} \tag{19}$$

and

$$\widetilde{\widetilde{p}}_1(x) = \begin{cases} p_1 \ for \ 0 \le x \le \frac{1}{\alpha} - 1 \\ 0 \ for \ \frac{1}{\alpha} - 1 < x \le 1 \end{cases} \tag{20}$$

where $p_0, p_1 > 0$ and $p_0 + p_1 = 1$. The reader may find it interesting to compare the trapped TIFS in Equations 17, 19 and 20 (*system II*) to the original TIFS in Equations 10, 12 and 13 (*system I*). The two systems look very similar. But the trapped system admits no invariant probability measure which contains an atom.

Notice that if $\widetilde{\widetilde{\mu}}$ is an invariant probability measure for the trapped system (*system II*) then

$$\widetilde{\mu} = g \circ \widetilde{\widetilde{\mu}}.$$

is an invariant measure for the original system (*system I)*.

6.6 Existence and Unicity of Invariant Measures for the trapped IFS

The following theorem asserts that there are two basically different situations which can occur.

Theorem 3. *Let $\frac{1}{2} < \alpha < 1$. Let $\widetilde{w}_0 : [0,1] \to [0,1]$ be defined by $\widetilde{w}_0(x) = \alpha x$. Let $\widetilde{w}_1 : [0, \frac{1}{\alpha} - 1] \to [0,1]$ be defined by $\widetilde{w}_1(x) = \alpha x + \alpha$. Let a Markov transition probability be given by Equation 18, with probabilities given by Equations 19 and 20. Then the Markov process possesses at least one and at most two linearly independent invariant probability measures. In particular, there exists at most two invariant probability measures which are also ergodic invariant measures for the tops dynamical system $\widetilde{\widetilde{T}} : [0,1] \to [0,1]$ defined by*

$$\widetilde{\widetilde{T}}(x) = (\frac{1}{\alpha}x) \text{ for all } x \in [0,1],$$

where (y) denotes the fractional part of the real number y. If α possesses certain arithmetic properties, namely that $\beta = \frac{1}{\alpha}$ is a "β-number", then the Markov process possesses a unique invariant probability measure which is ergodic for $\widetilde{\widetilde{T}}$.

Proof. See [5]. This relies on analysis of a corresponding code space system which we describe next.

6.7 Corresponding Code Space Systems

We can convert the system in Theorem 3 to an equivalent symbolic dynamical system with the aid of the maps

$$\widehat{w}_0(x, \sigma) = (\alpha x, 0\sigma) \text{ for all } (x, \sigma) \in [0, 1) \times \Omega \tag{21}$$

$$\widehat{w}_1(x, \sigma) = (\alpha x + \alpha, 1\sigma) \text{ for all } (x, \sigma) \in [0, \frac{1 - \alpha}{\alpha}] \times \Omega$$

We call the system described by Equation 21 "the lifted trapped TIFS". This system possesses a maximal invariant set $G_{\widetilde{\tau}} \subset [0, 1) \times \Omega$ which obeys

$$G_{\widetilde{\tau}} = \widehat{w}_0(G_{\widetilde{\tau}}) \cup \widehat{w}_1(G_{\widetilde{\tau}})$$

$G_{\widetilde{\tau}}$ is the graph of a one-to-one function $\widetilde{\tau} : [0, 1] \to \Omega$. The projection of this system onto the $[0, 1]$-direction gives us back the original system, the one in Theorem 3.

Projection in the Ω-direction provides us with a symbolic dynamical system $S : \Omega_\gamma \to \Omega_\gamma$ where $\widetilde{\tau}([0, 1]) = \Omega_\gamma$, and $S(\Omega_\gamma) = \Omega_\gamma$. $\{\Omega, S\}$ is the usual shift dynamical system on code space and the symbolic dynamical system $\{\Omega_\gamma, S\}$ is obtained by restricting the domain of S to Ω_γ. We could choose to write $S : \Omega_\gamma \to \Omega_\gamma$ as $S|_{\Omega_\gamma} : \Omega_\gamma \to \Omega_\gamma$ and $\{\Omega_\gamma, S|_{\Omega_\gamma}\} = \{\Omega_\gamma, S\}$, but do not. Note that we can compute $\widetilde{\tau}(x)$ by following the orbit of $x \in [0, 1]$ under the trapped TDS (Equation 15).

Thus we obtain the *symbolic IFS with place-dependent probabilities*

$$\{\Omega_\gamma; s_0|_{\Omega_\gamma}, s_1|_{\Omega_\gamma}; p_0(\sigma), p_1(\sigma)\}$$

where

$$s_0|_{\Omega_\gamma}(\sigma) = 0\sigma \text{ for all } \sigma \in \Omega_\gamma \tag{22}$$
$$s_1|_{\Omega_\gamma}(\sigma) = 1\sigma \text{ for all } \sigma \in \Omega_\gamma,$$

and the probabilities are given by $p_i(\sigma) = \widetilde{\widetilde{p}}_i(\widetilde{\tau}^{-1}(\sigma))$, that is

$$p_0(\sigma) = \begin{cases} p_0 \text{ for all } \sigma \in \Omega_\gamma \text{ with } 0 \le \sigma \le \gamma \\ 1 \text{ for all } \sigma \in \Omega_\gamma \text{ with } \gamma < \sigma \le 1 \end{cases} \tag{23}$$

and

$$p_1(\sigma) = \begin{cases} p_1 \text{ for all } \sigma \in \Omega_\gamma \text{ with } 0 \le \sigma \le \gamma \\ 0 \text{ for all } \sigma \in \Omega_\gamma \text{ with } \gamma < \sigma \le 1 \end{cases} \tag{24}$$

where $\gamma = \widetilde{\tau}(\frac{1-\alpha}{\alpha})$. This system is equivalent to the one in Theorem 3. It is the restriction to Ω_γ of the IFS $\{\Omega; s_0, s_1; p_0(\sigma), p_1(\sigma)\}$ where

$$s_0(\sigma) = 0\sigma \text{ for all } \sigma \in \Omega \tag{25}$$
$$s_1(\sigma) = 1\sigma \text{ for all } \sigma \in \Omega$$

and the probabilities $p_i : \Omega \to [0, 1]$ are defined in the same way as the $p_i : \Omega_\gamma \to [0, 1]$ in Equations 23 and 24 with Ω_γ replaced by Ω.

The system described by Equation 25 is an extension of the system described by Equation 22. Notice the following relationship between these two systems: Ω_γ *is the unique set attractor of the IFS* $\{\Omega; s_0, s_1|_{[0,\gamma]}\}$; that is Ω_γ is the unique nonempty compact subset of Ω such that

$$\Omega_\gamma = s_0(\Omega) \cup s_1|_{[0,\gamma]}(\Omega) = s_0(\Omega) \cup s_1(\Omega \cap [0, \gamma]).$$

By ignoring a countable set of points we can embed the code space in Ω in $[0, 1]$ to make pictures of, and simplify the visualization of, the *symbolic IFS's* in Equations 22 and 25. For example the symbolic system in Equation 22 may be represented by the following pretty IFS with place-dependent probabilities

$$\{[0, 1]; s_0 : [0, 1] \to [0, 1], s_1 : [0, \gamma] \to [0, 1]; p_0(x), p_1(x)\}$$

where

$$s_0(x) = \frac{1}{2}x \text{ for all } x \in [0, 1] \tag{26}$$
$$s_1(x) = \frac{1}{2}x + \frac{1}{2} \text{ for all } x \in [0, 1]$$

where

$$p_0(x) = \begin{cases} p_0 \text{ for all } 0 \le x \le \gamma \\ 1 \text{ for all } \gamma < x \le 1 \end{cases}$$

and

$$p_1(x) = \begin{cases} p_1 & \text{for all } 0 \leq x \leq \gamma \\ 0 & \text{for all } \gamma < x \leq 1 \end{cases}$$

Note that we use the symbols $s_0(\cdot)$ and $s_1(\cdot)$ to denote the maps and $p_0(\cdot)$ to denote the probabilities for embedded system.

Our Markov process, represented on Ω_γ, involves applying the maps $s_0|_{\Omega_\gamma}, s_1|_{\Omega_\gamma}$ with probabilities p_0, p_1 respectively when $\sigma \in \Omega_\gamma$ with $\sigma \leq \gamma$, and applying the map $s_0|_{\Omega_\gamma}$ with probability one when $\sigma \in \Omega_\gamma$ with $\sigma > \gamma$. This process equivalent to the Markov process on $[0,1]$ corresponding to the transition probability $\widetilde{\widetilde{P}}(x, B)$ in Equation 18 in Theorem 3.

But we can also consider the corresponding process on Ω which involves applying the maps s_0, s_1 with probabilities p_0, p_1 respectively when $\sigma \in \Omega$ with $\sigma \leq \gamma$, and applying the map s_0 with probability one when $\sigma \in \Omega$ with $\sigma > \gamma$. This process extends, to all of Ω, the domain of the original symbolic process. We will find it very useful to consider this latter process. This is because, when we change α, the maps and the space upon which they act remain unaltered; only the set of values of $\sigma \in \Omega$ for which $p_1(\sigma) = 0$ changes. In the original system, in Theorem 3, the slopes of the maps, the location where a probability becomes zero, and the set of allowed codes, all change with α.

We find that the structure of the system in Theorem 3 depends fundamentally on whether or not γ, the address of the point $\frac{1-\alpha}{\alpha}$, which may be computed by following the orbit of $\frac{1-\alpha}{\alpha}$ under the trapped tops dynamical system $\widetilde{\widetilde{T}} : [0,1] \rightarrow [0,1]$, terminates in an endless string of zeros or not. That is, on whether or not

$$\frac{r}{2^{k-1}} = \frac{1-\alpha}{\alpha} = \beta - 1$$

for some pair of positive integers k and r.

We work with the closure of the symbolic system. We are interested in those invariant measures which correspond to orbits of the following type of random iteration. When the current point lies in the region $0000... \leq \sigma \leq \gamma$ there is a nonzero probability that s_0 may be applied to the current point and there is a nonzero probability that s_1 may be applied to the current point. When the current point lies in the region $\gamma < \sigma < 1$ the map s_0 is applied with probability one. Theorem 2 implies that this Markov process possesses no invariant measures which include point masses.

Notice that if γ terminates in $\overline{0}$ or $\overline{1}$ then $S : \Omega_\gamma \rightarrow \Omega_\gamma$ is open, that is it maps open sets to open sets. This implies that the inverse branches of S are continuous in the product topology.

Another approach to the proof of Theorem 3 in the case $\gamma = x....xx1\overline{0}$ uses the formulation, and a theorem, of W. Parry [10] which involves intrinsic Markov chains. Our framework is broader since we work on all of Ω, which has a further implication: in the "recurrent" case, we obtain a stronger result than I. Werner, [12], regarding convergence of orbits produced by random iteration in the case of certain graph-directed IFS. This also has implications which are

entirely new, so far as we know, for the "not open" case, which corresponds for example to the case where γ is irrational.

References

1. M. F. Barnsley, *Fractals Everywhere*, Academic Press, New York, NY, 1988.
2. 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.
3. M. F. Barnsley and L. F. Barnsley, *Fractal Transformations*, in "The Colours of Infinity: The Beauty and Power of Fractals", by Ian Stewart et al., Published by Clear Books, London, (2004), pp.66-81.
4. M. F. Barnsley, *Ergodic Theory, Fractal Tops and Colour Stealing,* Lecture Notes*,* Australian National University, *2004.*
5. M. F. Barnsley, *Theory and Application of Fractal Tops (Full Version),* Preprint, Australian National University, 2005.
6. M.F.Barnsley, J.E.Hutchinson, O. Stenflo *A Fractal Valued Random Iteration Algorithm and Fractal Hierarchy* (2003) to appear in Fractals journal.
7. J. Elton, *An Ergodic Theorem for Iterated Maps,* Ergodic Theory Dynam. Systems, 7 (1987), pp. 481-488.
8. J. E. Hutchinson, *Fractals and Self-Similarity*, Indiana. Univ. Math. J., 30 (1981), pp. 713-749.
9. N. Lu, *Fractal Imaging*, Academic Press, (1997).
10. W. Parry, *Symbolic Dynamics and Transformations of the Unit Interval*, Trans. Amer. Math. Soc., **122** (1966),368-378.
11. A.Rényi, *Representations for Real Numbers and Their Ergodic Properties*, Acta Math. Acad. Sci. Hung., **8** (1957), pp. 477-493.
12. I. Werner, *Ergodic Theorem for Contractive Markov Systems,* Nonlinearity **17** (2004) 2303-2313