Fractal transformations of harmonic functions

Michael F. Barnsley, Uta Freiberg
Department of Mathematics, Australian National University, Canberra, ACT, Australia

28 December 2006

ABSTRACT

The theory of fractal homeomorphisms is applied to transform a Sierpinski triangle into what we call a Kigami triangle. The latter is such that the corresponding harmonic functions and the corresponding Laplacian $\Delta$ take a relatively simple form. This provides an alternative approach to recent results of Teplyaev. Using a second fractal homeomorphism we prove that the outer boundary of the Kigami triangle possesses a continuous first derivative at every point. This paper shows that IFS theory and the chaos game algorithm provide important tools for analysis on fractals.

1 Introduction

We begin this paper with a brief review of the recently developed theory of fractal tops and fractal transformations. We emphasize that fractal transformations may be computed readily by means of the chaos game algorithm. Then we develop a beautiful application: we show how the theory may be applied to transform the Sierpinski triangle so that the corresponding harmonic functions and the corresponding Laplacian $\Delta$ take a relatively simple form. This provides an alternative approach to recent results of Teplyaev. In particular, Kigami appears to have written the first papers regarding representation of Sierpinski triangle in harmonic coordinates and Meyers may have been to present the geometrical interpretation of the harmonic representation. Here we prove, by use the fractal homeomorphism theorem a second time and a third time, that the basic curves, from which the Kigami triangle may be constructed, are continuously differentiable.

Other relevant references of which we are aware are Berger, Kusuoka, and also Strichartz, top of page 194, where there is mention of three-by-three row-stochastic transformations in equations (5) and their relation to the corresponding two-by-two matrices in equations (8), with mention of the eigenvectors and eigenvalues of the latter. See also Barnsley concerning the impedance functions and spectra of renormalizable electro-mechanical systems and their relation to Julia sets.

2 Hyperbolic IFS and Birkhoff’s ergodic theorem

Definition 1. Let $(\mathcal{X}, d_\mathcal{X})$ be a complete metric space. Let $\{f_1, f_2, \ldots, f_N\}$ be a finite sequence of strictly
contractive transformations, \( f_n : \mathbb{X} \to \mathbb{X} \), for \( n = 1, 2, \ldots, N \). Then
\[
\mathcal{F} := \{ \mathbb{X}; f_1, f_2, \ldots, f_N \}
\]
is called a hyperbolic iterated function system or hyperbolic IFS.

A transformation \( f_n : \mathbb{X} \to \mathbb{X} \) is strictly contractive iff there exists a number \( l_n \in [0, 1) \) such that \( d(f_n(x), f_n(y)) \leq l_n d(x, y) \) for all \( x, y \in \mathbb{X} \).

Let \( \Omega \) denote the set of all infinite sequences \( \{\sigma_k\}_{k=1}^{\infty} \) of symbols belonging to the alphabet \( \{1, \ldots, N\} \). We write \( \sigma = \sigma_1\sigma_2\sigma_3\ldots \in \Omega \) to denote a typical element of \( \Omega \), and we write \( \omega_k \) to denote the \( k^{th} \) element of \( \omega \in \Omega \). Then \((\Omega, d_{\Omega})\) is a compact metric space, where the metric \( d_{\Omega} \) is defined by \( d_{\Omega}(\sigma, \omega) = 0 \) when \( \sigma = \omega \) and \( d_{\Omega}(\sigma, \omega) = 2^{-k} \) when \( k \) is the least index for which \( \sigma_k \neq \omega_k \). We call \( \Omega \) the code space associated with the IFS \( \mathcal{F} \).

Let \( \sigma \in \Omega \) and \( x \in \mathbb{X} \). Then, using the contractivity of \( \mathcal{F} \), it is straightforward to prove that
\[
\phi_{\mathcal{F}}(\sigma) := \lim_{k \to \infty} f_{\sigma_1} \circ f_{\sigma_2} \circ \ldots \circ f_{\sigma_k}(x)
\]
exists, uniformly for \( x \) in any fixed compact subset of \( \mathbb{X} \), and depends continuously on \( \sigma \). See for example\(^1 \), Theorem 3. Let
\[
A_{\mathcal{F}} = \{ \phi_{\mathcal{F}}(\sigma) : \sigma \in \Omega \}.
\]
Then \( A_{\mathcal{F}} \subset \mathbb{X} \) is called the attractor of \( \mathcal{F} \). The continuous function
\[
\phi_{\mathcal{F}} : \Omega \to A_{\mathcal{F}}
\]
is called the address function of \( \mathcal{F} \). We call \( \phi_{\mathcal{F}}^{-1}(\{x\}) := \{ \sigma \in \Omega : \phi_{\mathcal{F}}(\sigma) = x \} \) the set of addresses of the point \( x \in A_{\mathcal{F}} \).

Clearly \( A_{\mathcal{F}} \) is compact, nonempty, and has the property
\[
A_{\mathcal{F}} = f_1(A_{\mathcal{F}}) \cup f_2(A_{\mathcal{F}}) \cup \ldots \cup f_N(A_{\mathcal{F}}).
\]
Indeed, if we define \( \mathbb{H}(\mathbb{X}) \) to be the set of nonempty compact subsets of \( \mathbb{X} \), and we define \( \mathcal{F} : \mathbb{H}(\mathbb{X}) \to \mathbb{H}(\mathbb{X}) \) by
\[
\mathcal{F}(S) = f_1(S) \cup f_2(S) \cup \ldots \cup f_N(S),
\]
for all \( S \in \mathbb{H}(\mathbb{X}) \), then \( A_{\mathcal{F}} \) can be characterized as the unique fixed point of \( \mathcal{F} \), see Hutchinson\(^{14} \), section 3.2, and Williams\(^{32} \).

IFSs may be used to represent diverse subsets of \( \mathbb{R}^2 \). An IFS whose attractor is the Sierpinski triangle with vertices at \( A = (0, 0) \), \( B = (1, 0) \), and \( C = (0.5, \sqrt{0.75}) \) is
\[
\mathcal{F} = \{ \mathbb{R}^2; f_1, f_2, f_3 \}
\]
where \( f_1(x) = \frac{1}{2}(x + A) \), \( f_2(x) = \frac{1}{2}(x + B) \), and \( f_3(x) = \frac{1}{2}(x + C) \) with \( x = (x, y) \). We may represent \( f_1 \) using matrix notation for affine transformations, namely,
\[
\begin{pmatrix}
\frac{1}{2} & 0 \\
0 & \frac{1}{2}
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}
+ \begin{pmatrix}
A_x \\
A_y
\end{pmatrix},
\]
and similarly for \( f_2, f_3 \). We will write \( ABC \) to denote the triangle whose vertices are \( A, B, \) and \( C \). Notice that \( ABC \) is an equilateral triangle.

**Definition 2.** An IFS with probabilities is a hyperbolic IFS \( \{\mathbb{X}; f_1, f_2, \ldots, f_N\} \) together with a set of real numbers \( p_n > 0 \) for \( n = 1, 2, \ldots, N \) with \( p_1 + p_2 + \ldots + p_N = 1 \).
It is a basic result of IFS theory that given any IFS $F$ with probabilities there exists a unique Borel measure $\mu$ of norm one, namely a probability measure, such that

$$\mu(B) = \sum_{n=1}^{N} p_n \mu(f_n^{-1}(B))$$

for all Borel subsets $B$ of $X$. The support of this measure is the attractor $A_F$ of the IFS. This measure may be referred to as the measure-attractor of the IFS. The standard measure on the Sierpinski triangle corresponds to $p_1 = p_2 = p_3 = \frac{1}{3}$.

This is a normalized $\frac{\ln 3}{\ln 2}$-dimensional Hausdorff measure.

Both the measure-attractor and the attractor of an IFS can be described in terms of the chaos game algorithm. We define a random orbit of $x_0 \in X$ to be $\{x_k\}_{k=0}^\infty$ where

$$x_{k+1} = f_{\sigma_k}(x_k),$$

$\sigma_k \in \{1, 2, ..., N\}$ with $\sigma_k = n$ with probability $p_n$ independently of all other choices. (That is, the sequence $\sigma_1 \sigma_2 \sigma_3...$ is i.i.d.) Then, with probability one, the sequence of measures

$$\mu_k := \frac{1}{k} \sum_{j=1}^{k} \delta_{x_j}$$

converges weakly to $\mu$ and, also with probability one, the intersection of the nested (decreasing) sequence of compact sets

$$A_k = \overline{\{x_j\}_{j=k}^\infty}$$

equals $A_F$. The bar denotes closure.

The proof of these results relies on Birkhoff’s ergodic theorem, see for example Vrscay. The scholarly history of the chaos game is discussed by Kaijer and Stenflo, and appears to begin in 1935 with the work of Onicescu and Mihok. Mandelbrot, pp.196-199, used a version of it to help compute pictures of certain Julia sets; it was introduced to IFS theory and developed by one of the authors and his coworkers, see for example Barnsley and Demko, Barnsley, Berger, and Elton, where the relevant theorems and much discussion can be found.

### 3 Homeomorphisms between fractals

We order the elements of $\Omega$ according to

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

where $k$ is the least index for which $\sigma_k \neq \omega_k$. This is a linear ordering, sometimes called the lexicographic ordering.

Notice that all elements of $\Omega$ are less than or equal to $\bar{T} = 11111...$ and greater than or equal to $\bar{N} = NNNNN...$. Also, any pair of distinct elements of $\Omega$ is such that one member of the pair is strictly greater than the other. In particular, the set of addresses of a point $x \in A_F$ is both closed and bounded above by $\bar{T}$. It follows that $\phi_F^{-1}(\{x\})$ possesses a unique largest element. We denote this element by $\tau_F(x)$.

**Definition 1.** Let $F$ be a hyperbolic IFS with attractor $A_F$ and address function $\phi_F : \Omega \rightarrow A_F$. Let

$$\tau_F(x) = \max \{\sigma \in \Omega : \phi_F(\sigma) = x\} \text{ for all } x \in A_F.$$
Then
\[ \Omega_F := \{ \tau_F(x) : x \in A_F \} \]
is called the tops code space and
\[ \tau_F : A_F \to \Omega_F \]
is called the tops function, for the IFS \( F \).

We remark that \( \Omega_F \) is a shift invariant subspace \( \Omega \), see Barnsley\(^7\). Consequently the rich theory of symbolic dynamics, see Lind\(^2\), can be brought to bear on the analysis of tops code spaces and can provide much information about the underlying fractal structures.

But we will need the following. Let \( G \) denote a hyperbolic IFS that also consists of \( N \) functions. Then \( \phi_G \circ \tau_F : A_F \to A_G \) is a mapping from the attractor of \( F \) into the attractor of \( G \). We refer to \( \phi_G \circ \tau_F \) as a fractal transformation.

**Definition 2.** The address structure of \( F \) is defined to be the set of sets
\[ C_F = \{ \phi_F^{-1}(\{x\}) \cap \Pi_F : x \in A_F \}. \]

The address structure of an IFS is a certain partition of \( \Pi_F \). Let \( C_G \) denote the address structure of \( G \). Let us write \( C_F \prec C_G \) to mean that for each \( S \in C_F \) there is \( T \in C_G \) such that \( S \subset T \). Notice that if \( C_F = C_G \) then \( \Omega_F = \Omega_G \). Some examples of address structures are given by Barnsley\(^8\).

**Theorem 3.** Let \( F \) and \( G \) be two hyperbolic IFSs such that \( C_F \prec C_G \). Then the fractal transformation \( \phi_G \circ \tau_F : A_F \to A_G \) is continuous. If \( C_F = C_G \) then \( \phi_G \circ \tau_F \) is a homeomorphism.

The proof is given by Barnsley\(^7\)\(^8\).

Given the two hyperbolic IFSs \( F \) and \( G \), such that \( C_F = C_G \), we note that the attractor of the IFS \( F \times G \) defined by
\[ F \times G = \{ X \times X; w_n = (f_n, g_n), n = 1, 2, \ldots, N \}. \]
contains the graph \( G \) of the homeomorphism \( \phi_G \circ \tau_F \). Loosely speaking, the "top" of the attractor of \( F \times G \) equals \( G \). Precisely, if \( A_F \) is totally disconnected or finitely ramified then \( A_{F \times G} = G \); in general it is straightforward to extract \( G \) from \( A_{F \times G} \). But the point we want to emphasize here is that \( G \) can be computed by means of the chaos game algorithm applied to the IFS \( F \times G \), see Barnsley\(^7\).

### 4 Energy and harmonic functions on a sierpinski triangle

Let \( F = \{ \mathbb{R}^2; f_1, f_2, f_3 \} \) denote the IFS associated with the Sierpinski triangle \( A_F \) with vertices at \( A, B, \) and \( C \). Let \( V_0 \) denote the set of vertices of the triangle \( ABC \). Namely, \( V_0 = \{ A, B, C \} \).

Let \( \Omega' \) denote the set of finite sequences of symbols from the alphabet \( \{ 1, 2, \ldots, N \} \), including the empty string. That is,
\[ \Omega' = \bigcup_{k=1}^{\infty} \{ 1, 2, 3 \}^k \cup \{ \emptyset \}. \]
Let \( \sigma \in \Omega' \). We write \(|\sigma|\) to denote the length of \( \sigma \). Then \( \sigma \) is the empty string if and only if \( |\sigma| = 0 \). We write \( \sigma = \sigma_1 \sigma_2 \sigma_3 \ldots \sigma_{|\sigma|} \) to denote the components of \( \sigma \) when \( |\sigma| \neq 0 \).

We write \( f_{\emptyset} \) to denote the identity map on \( A_F \) and, for \( \sigma \in \Omega' \) with \( |\sigma| \neq 0 \), we define
\[ f_{\sigma} = f_{\sigma_1} \circ f_{\sigma_2} \ldots \circ f_{|\sigma|}. \]
We define, for all $m = 0, 1, 2, \ldots$,

$$V_m = \bigcup_{\{\sigma \in \Omega^m : |\sigma| = m\}} f_\sigma(\{A, B, C\}).$$

We also define $V = \bigcup_{m=0}^{\infty} V_m$. That is, $V$ is the set of all vertices of the Sierpinski triangle $A_F$.

We now follow a standard construction due to Kigami, see Strichartz$^{27}$. Let $u : V \to \mathbb{R}$. Then we define, for all $m = 0, 1, 2, \ldots$,

$$E_m(u) = \sum_{\{x, y \in V_m : |x-y| = 2^{-m}\}} (u(x) - u(y))^2.$$

Then if

$$E(u) = \lim_{m \to \infty} \frac{5}{3} m E_m(u)$$

exists, we say that "$u$ has finite energy $E(u)$". We also say that "$u$ is in the domain of (the closure of) the Laplacian", see Strichartz$^{27}$. The value of the renormalization constant $\frac{5}{3}$ follows from Equation (6) below.

Suppose we are given the values

$$u(A) = h_A, u(B) = h_B, u(C) = h_C. \quad (3)$$

Then the corresponding harmonic function $h : V \to \mathbb{R}$ is defined to be the function $u$ which minimizes $E_m(u)$ for each $m$, subject to the constraints in Equation (3).

We are going to construct explicitly $h$. It is useful, for generalizations, to understand where the various special numbers and matrices come from. Let $a$ denote the midpoint of the line segment $BC$, $b$ denote the midpoint of the line segment $CA$, and $c$ denote the midpoint of $AB$. Then $V_1 = \{A, B, C, a, b, c\}$. We have

$$E_0(u) = |h_A - h_B|^2 + |h_B - h_C|^2 + |h_C - h_A|^2$$

and

$$E_1(u) = |h_A - u(c)|^2 + |u(c) - h_B|^2 + |h_B - u(a)|^2 + |u(a) - h_C|^2 + |h_C - u(b)|^2 + |u(b) - h_A|^2 + |u(a) - u(b) + (u(b) - u(c)) + (u(c) - u(a))|^2.$$

We minimize $E_1(u)$ with respect to the values $u(a)$, $u(b)$, $u(c)$. We find that at the minimum, where $u(a) = h(a)$, $u(b) = h(b)$, and $u(c) = h(c)$,

$$\frac{\partial E_1(u)}{\partial u(a)} = (h(a) - h_B) + (h(a) - h_C) + (h(a) - h(c)) + (h(b) - h(b)) = 0,$$

that is,

$$4h(a) - h(b) - h(c) = h_B + h_C$$

and two other similar equations which may be obtained by cyclic permutation. It follows that

$$\begin{pmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{pmatrix} \begin{pmatrix} h(a) \\ h(b) \\ h(c) \end{pmatrix} = \begin{pmatrix} h_B + h_C \\ h_C + h_A \\ h_A + h_B \end{pmatrix}.$$

Inverting, we find

$$\begin{pmatrix} h(a) \\ h(b) \\ h(c) \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} h_A \\ h_B \\ h_C \end{pmatrix}.$$
It follows that
\[ E_1(h) = \left\{ -\frac{6}{5} h_A h_C - \frac{6}{5} h_A h_B + \frac{6}{5} h_A^2 + \frac{6}{5} h_B^2 - \frac{6}{5} h_B h_C + \frac{6}{5} h_C^2 \right\} = \frac{3}{5} E_0(h). \]

Now observe that \( A = f_1(A), B = f_2(B), C = f_3(C), a = f_3(B) = f_2(C), b = f_1(C) = f_3(A), \) and \( c = f_2(A) = f_1(B). \) Using this, we deduce from (4) that
\[
\begin{bmatrix}
  h(f_1(A)) \\
  h(f_2(B)) \\
  h(f_3(C))
\end{bmatrix} = A_i \begin{bmatrix}
  h_A \\
  h_B \\
  h_C
\end{bmatrix},
\]
where
\[
A_1 = \begin{pmatrix}
  1 & 0 & 0 \\
  \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\
  \frac{2}{5} & \frac{2}{5} & \frac{2}{5}
\end{pmatrix}, \quad A_2 = \begin{pmatrix}
  \frac{2}{5} & 2 & 0 \\
  \frac{2}{5} & \frac{2}{5} & 0 \\
  \frac{2}{5} & \frac{2}{5} & \frac{2}{5}
\end{pmatrix}, \quad A_3 = \begin{pmatrix}
  \frac{2}{5} & \frac{2}{5} & 2 \\
  \frac{2}{5} & \frac{2}{5} & 0 \\
  \frac{2}{5} & \frac{2}{5} & \frac{2}{5}
\end{pmatrix}.
\]

By iterating the above steps, we readily deduce that
\[ E_m(h) = \frac{3}{5} E_{m-1}(h) = \left( \frac{3}{5} \right)^m E_0(h), \]
and
\[
\begin{bmatrix}
  h(f_{\sigma_1 \sigma_2 \ldots \sigma_m}(A)) \\
  h(f_{\sigma_1 \sigma_2 \ldots \sigma_m}(B)) \\
  h(f_{\sigma_1 \sigma_2 \ldots \sigma_m}(C))
\end{bmatrix} = A_{\sigma_m} \begin{bmatrix}
  h(f_{\sigma_1 \sigma_2 \ldots \sigma_{m-1}}(A)) \\
  h(f_{\sigma_1 \sigma_2 \ldots \sigma_{m-1}}(B)) \\
  h(f_{\sigma_1 \sigma_2 \ldots \sigma_{m-1}}(C))
\end{bmatrix} = A_{\sigma_m} A_{\sigma_{m-1}} \ldots A_{\sigma_1} \begin{bmatrix}
  h_A \\
  h_B \\
  h_C
\end{bmatrix},
\]
for all \( \sigma_1 \sigma_2 \ldots \sigma_m \in \Omega' \) with \( m \geq 1. \) Strichartz\textsuperscript{27}, p.17, says that it should be possible to obtain any desired information about harmonic functions from this last expression but that in practice it may require a lot of work. We are going to show that it is quite easy to obtain a satisfactory description of \( h, \) both theoretical and practical, starting from this expression.

## 5 Harmonic functions and the Kigami triangle

### 5.1 Values of harmonic functions via the chaos game

We depart from the standard construction. We let \( m \) tend to infinity, use the continuity of \( h \) on \( \mathcal{A}_\sigma, \) and exploit the uniform convergence in (1) to deduce that that \( A^m: = \lim_{m \to \infty} A_{\sigma_m} A_{\sigma_{m-1}} \ldots A_{\sigma_1} \) exists for all \( \sigma \in \Omega \) and
\[
\begin{bmatrix}
  h(\phi_{\sigma}(\sigma)) \\
  h(\phi_{\sigma}(\sigma)) \\
  h(\phi_{\sigma}(\sigma))
\end{bmatrix} = A_{\sigma} \begin{bmatrix}
  h_A \\
  h_B \\
  h_C
\end{bmatrix} \text{ for all } \sigma \in \Omega.
\]

We rewrite this as
\[
h(\phi_{\sigma}(\sigma)) = A_{\sigma} h_0
\]
where \( h(\phi_{\sigma}(\sigma)) = (h(\phi_{\sigma}(\sigma)), h(\phi_{\sigma}(\sigma)), h(\phi_{\sigma}(\sigma)))^T \) and \( h_0 = (h_A, h_B, h_C)^T. \) This provides the value of \( h \) at the point \( \phi_{\sigma}(\sigma). \) In order to obtain global information let \( v^T = (v_1, v_2, v_3) \in \mathbb{R}^3 \) be a test vector and consider
\[
v. h(\phi_{\sigma}(\sigma)) = v. A_{\sigma} h_0 = A_{\sigma}^T v. h_0
\]
It follows that
\[
A_i^T A_{\sigma}^T v. h_0 = A_{\sigma}^T v. h_0 = v. h(\phi_{\sigma}(i\sigma)) \text{ for } i = 1, 2, 3.
\]
Hence we can apply algorithms of the chaos game type to evaluate sequences of values \( \{ v \cdot h(x_k, y_k) \}_{k=0}^{\infty} \) where the sequence of points \( \{ (x_k, y_k) \}_{k=0}^{\infty} \) are distributed ergodically on the Sierpinski triangle according, say, to the standard measure. For simplicity suppose we start from a point \((x_0, y_0) = \phi_F(\sigma)\) and vector \(v_0 = A_0^\top v\). Then at the \(k\)th iterative step we choose \( i \in \{1,2,3\} \) randomly, independently of all other choices, and set

\[
(x_k, y_k) = f_i(x_{k-1}, y_{k-1}), \quad v_k = A_i^\top v_{k-1}, \quad v_0 \cdot h(x_k, y_k) = v_k \cdot h_0 \quad \text{for} \quad k = 1, 2, 3, \ldots
\]

Notice that the elements of the columns of the matrices \(A_i^\top\) sum to one. Hence each of the transformations represented by the \(A_i^\top\) maps the plane

\[
\Pi_\delta := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = \delta\}
\]

into itself for each \(\delta \in \mathbb{R}\). It follows that if \(v_0 \in \Pi_1\) then \(v_k \in \Pi_1\) for all \(k\), and then the last equation in (7) yields

\[
h(x_k, y_k) = (v_{k,1} h_A + v_{k,2} h_B + v_{k,3} h_C)
\]

for \(k = 1, 2, 3, \ldots\) Thus we have a simple algorithm which allows us to compute, rapidly, values of the harmonic function \(h\).

### 5.2 The Kigami triangle

We can obtain much finer information by looking at the set of points \(T_{v_0} = \{ A_{\sigma}^\top v_0 : \sigma \in \Omega \} \subset \mathbb{R}^3\). Clearly \(T_{v_0}\) obeys the self-referential equation

\[
T_{v_0} = A_1^\top (T_{v_0}) \cup A_2^\top (T_{v_0}) \cup A_3^\top (T_{v_0}).
\]

Notice that \(v_0 \in \Pi_3\) implies \(T_{v_0} \subset \Pi_3\). Without loss of generality we choose \(C = 1\) and we represent the points of \(\theta : \Pi_1 \to \mathbb{R}^2\) by means of the transformation \(\theta((\alpha, \beta, \gamma)) = (\beta, \gamma)\) with inverse \(\theta^{-1}(\beta, \gamma) = (1 - \beta - \gamma, \beta, \gamma)\). We readily find, in matrix representation,

\[
\theta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \quad \text{and} \quad \theta^{-1} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}
\]

Then the new transformations \(g_i = \theta A_i^\top \theta^{-1} : \mathbb{R}^2 \to \mathbb{R}^2\) are defined explicitly by

\[
g_1 \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} \\
g_2 \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix} \\
g_3 \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \beta \\ \gamma \end{pmatrix} + \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}
\]

Notice that each of these affine transformations is strictly contractive. Hence

\[
G = \{ \mathbb{R}^2 : g_1, g_2, g_3 \}
\]

is a hyperbolic IFS. Let \(\tilde{T}_{v_0} = \theta(T_{v_0})\). Then \(\tilde{T}_{v_0} = g_1(\tilde{T}_{v_0}) \cup g_2(\tilde{T}_{v_0}) \cup g_3(\tilde{T}_{v_0})\). This set is compact and nonempty. Hence it must be the unique attractor of the IFS \(G\), that is

\[
A_G = \tilde{T}_{v_0}.
\]

It is easy to make sketches of \(A_G\) with the aid of the chaos game algorithm. See Figure 1. We refer to \(A_G\) as a Kigami triangle because he first noted the simplicity of the Sierpinski triangle when it is represented using
harmonic coordinates. Alexander Teplyaev\textsuperscript{31} \textsuperscript{28} \textsuperscript{29} \textsuperscript{30} appears to be the first to publish pictures of a Kigami triangle and some of its relatives,

We now prove that the Sierpinski triangle $A_F$ and the Kigami triangle $A_G$ are homeomorphic. Let $\Delta$ denote the triangle with vertices $A = (0, 0), B = (1, 0),$ and $C = (0, 1)$. Let $\tilde{a} = (\frac{2}{5}, \frac{3}{5}), \tilde{b} = (\frac{1}{5}, \frac{2}{5})$ and $\tilde{c} = (\frac{3}{5}, \frac{1}{5})$. Then we note that

\[ g_1(\Delta) \cap g_2(\Delta) = \tilde{c}, g_2(\Delta) \cap g_3(\Delta) = \tilde{a}, g_3(\Delta) \cap g_1(\Delta) = \tilde{b}. \]

It follows that

\[ g_1(A_G) \cap g_2(A_G) = \tilde{c}, g_2(A_G) \cap g_3(A_G) = \tilde{a}, g_3(A_G) \cap g_1(A_G) = \tilde{b}. \]

This implies that $A_F$ has the same code structure as $A_G$. Hence, by Theorem 3, the Sierpinski triangle $A_F$ and the Kigami triangle $A_G$ are homeomorphic. This is a much shorter proof, for those familiar with IFS theory, than the one given by Teplyaev\textsuperscript{31}; in essence the proofs are the same, but the isolation of the tools used is much clearer in the IFS framework. We will use these same tools to prove that the basic curves, from which the Kigami triangle may be constructed, are continuously differentiable.

Let $H : A_F \to A_G$ denote the homeomorphism between the attractors. This continuous invertible transformation, with continuous inverse $H^{-1} : A_G \to A_F$, relates each point on one attractor to the corresponding point on the other attractor with the same set of addresses. It readily follows that we can write, consistently with our earlier usage

\[ (\beta, \gamma) = H(x, y). \]

The harmonic function $h : A_F \to \mathbb{R}$ becomes, on $A_G$, $\tilde{h} : A_G \to \mathbb{R}$ given by

\[ \tilde{h} = h \circ H^{-1} \text{ and } \tilde{h}(\beta, \gamma) = h(x, y). \]

and we have

\[ \tilde{h}(\beta, \gamma) = h(x, y). \]

Finally note that

\[ \tilde{h}(\beta, \gamma) = h_A(1 - \beta - \gamma) + h_B\beta + h_C\gamma = h(x, y). \]

In the new coordinates, the harmonic function may be described using the plane $\Pi$ which passes through the points $(h_A, 0, 0)$, $(0, h_B, 0)$ and $(0, 0, h_C)$. Treat the Kigami triangle as being located on the plane $\alpha = 0$, in three-dimensional space described using rectangular coordinates $(\alpha, \beta, \gamma)$ so that a typical point on the Kigami triangle is represented by $(0, \beta, \gamma)$. If the unique corresponding point on $\Pi$, with the same $\beta$, and $\gamma$ coordinates, is $(\alpha, \beta, \gamma)$ then $\tilde{h}(\beta, \gamma) = \alpha$. See Figure 1.

To obtain a symmetrical version of the Kigami triangle we make the change of variable

\[ T = \left( \begin{array}{cc} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{array} \right), \quad T^{-1} = \left( \begin{array}{cc} 1 & -\frac{1}{2}\sqrt{3} \\ 0 & \frac{\sqrt{3}}{2} \end{array} \right) \]

which corresponds to mapping the vertices $(0, 0), (1, 0)$ and $(0, 1)$ to $(0, 0), (1, 0)$ and $(0, \sqrt{3}/2)$, respectively, which lie of an equilateral triangle. Then the linear transformations become self-adjoint, specifically

\[ \tilde{g}_1(\beta) = R_{\frac{\pi}{6}} DR_{-2\pi}(\beta), \quad \tilde{g}_2(\beta) = R_{-2\pi} DR_{2\pi}(\beta) + t_2, \quad \tilde{g}_3(\beta) = D(\beta) + t_3 \]

where

\[ \beta = \left( \begin{array}{c} \beta \\ \gamma \end{array} \right), \quad D = \left( \begin{array}{cc} 1 & 0 \\ 0 & \frac{\sqrt{3}}{2} \end{array} \right), \quad \beta = \left( \begin{array}{c} \beta \\ \gamma \end{array} \right), \quad t_2 = \left( \begin{array}{c} \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} \end{array} \right), \quad t_3 = \left( \begin{array}{c} \frac{2}{\sqrt{3}} \\ \frac{2}{\sqrt{3}} \end{array} \right) \]

and $R_{\theta}$ denotes clockwise rotation about the origin through angle $\theta$. See Figure 2.
Figure 1: The Sierpinski triangle and the Teplyaev triangle. They are homeomorphic because they have the same code structure. Unlike the Sierpiskis triangle, the Tephaev triangle has a unique "direction" at every point of its outer boundary and at every point on the internal closed curves.

Figure 2: An equilateral Kigami triangle.
We introduce the vectors
\[
\begin{align*}
    \mathbf{w}_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\
    \mathbf{w}_2 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
    \mathbf{w}_3 &= \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \\
    \mathbf{w}_4 &= \begin{pmatrix} \frac{3}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \\
    \mathbf{w}_5 &= \begin{pmatrix} \frac{2}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}, \\
    \mathbf{w}_6 &= \begin{pmatrix} \frac{1}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}.
\end{align*}
\]

Then we note that
\[
\begin{align*}
    \tilde{g}_1(\mathbf{w}_1) &= \mathbf{w}_1, \quad \tilde{g}_1(\mathbf{w}_2) = \mathbf{w}_6, \\
    \tilde{g}_1(\mathbf{w}_3) &= \mathbf{w}_5, \\
    \tilde{g}_2(\mathbf{w}_2) &= \mathbf{w}_2, \quad \tilde{g}_2(\mathbf{w}_1) = \mathbf{w}_6, \\
    \tilde{g}_2(\mathbf{w}_3) &= \mathbf{w}_4, \quad \tilde{g}_2(\mathbf{w}_4) = \mathbf{w}_5, \\
    \tilde{g}_3(\mathbf{w}_3) &= \mathbf{w}_3, \quad \tilde{g}_3(\mathbf{w}_2) = \mathbf{w}_4, \quad \tilde{g}_3(\mathbf{w}_1) = \mathbf{w}_5.
\end{align*}
\]

A Kigami triangle can thought of as a type of fractal interpolation, see Barnsley\(^2\), taking the topological form of a Sierpinski triangle in place a line segment, and passing through, or interpolating, the points \(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4, \mathbf{w}_5, \) and \(\mathbf{w}_6\).

### 5.3 The nature of the Kigami triangle

By zooming in and looking closely at pictures of \(\tilde{T}_{\mathbf{w}_6}\), we are led to the impression that there is a well-defined tangent direction at every point on each of the curves which define the outer boundary and also on all of the internal loops, and that this tangent direction depends continuously on location. This is indeed the case. Thus IFS theory provides geometrical information which makes more precise the known fact, proved using Osceledet’s theorem on products of random matrices, that there is a well-defined direction at almost every point, with respect to the standard measure. The proof is quite beautiful.

**Theorem 1.** The attractor \(\tilde{G}\) of the IFS
\[
\tilde{G} = \{\mathbb{R}^2; \tilde{g}_1, \tilde{g}_2\}
\]
is the graph of a continuously differentiable function \(b : [0, 1] \to [0, 1]\).

**Proof.** It follows from Equations (9) that
\[
\begin{align*}
    \tilde{g}_1(\mathbf{w}_1) &= \mathbf{w}_1, \quad \tilde{g}_1(\mathbf{w}_2) = \mathbf{w}_6, \\
    \tilde{g}_2(\mathbf{w}_2) &= \mathbf{w}_2, \quad \tilde{g}_2(\mathbf{w}_1) = \mathbf{w}_6, \\
    \tilde{g}_2(\mathbf{w}_3) &= \mathbf{w}_4, \quad \tilde{g}_2(\mathbf{w}_4) = \mathbf{w}_5, \\
    \tilde{g}_3(\mathbf{w}_3) &= \mathbf{w}_3, \quad \tilde{g}_3(\mathbf{w}_2) = \mathbf{w}_4, \quad \tilde{g}_3(\mathbf{w}_1) = \mathbf{w}_5.
\end{align*}
\]
Furthermore
\[
\begin{align*}
    \tilde{g}_1(\mathbf{w}_1 \mathbf{w}_2 \mathbf{w}_3) &= \mathbf{w}_1 \mathbf{w}_6 \mathbf{w}_5, \\
    \tilde{g}_2(\mathbf{w}_1 \mathbf{w}_2 \mathbf{w}_3) &= \mathbf{w}_6 \mathbf{w}_2 \mathbf{w}_4, \\
    \tilde{g}_3(\mathbf{w}_1 \mathbf{w}_2 \mathbf{w}_3) &= \mathbf{w}_4 \mathbf{w}_3 \mathbf{w}_6,
\end{align*}
\]
and the triangles \(\mathbf{w}_1 \mathbf{w}_6 \mathbf{w}_5\) and \(\mathbf{w}_6 \mathbf{w}_2 \mathbf{w}_4\) are both contained in the triangle \(\mathbf{w}_1 \mathbf{w}_2 \mathbf{w}_3\). It follows that that the code structure of the attractor of the IFS \(\{\mathbb{R}^2; \tilde{g}_1, \tilde{g}_2\}\) is the same as the code structure of the attractor of the IFS
\[
\{\mathbb{R}; f_1, f_2\}
\]
whose attractor is the line segment \([0, 1] \subset \mathbb{R}\). It follows, via the fractal homeomorphism theorem, that \(\tilde{G}\) is the graph of a continuous function \(b : [0, 1] \to [0, 1]\).

Finally we outline the proof that \(b\) is differentiable and that its derivative is continuous. Consideration of the manner in which the two functions which comprise \(\tilde{G}\) map ellipses into ellipses leads to the conclusion that the slope of \(\tilde{G}\) exists at every point if and only if it is defined by the attractor of the IFS
\[
\mathcal{H} = \left\{ \begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix} \right\}; \quad w_1(x) = \frac{\sqrt{3} - 3x}{\sqrt{3}x - 5}, \quad w_2(x) = \frac{\sqrt{3} + 3x}{\sqrt{3}x + 5}.
\]
That is, we need to show that $\mathcal{H}$ is indeed a hyperbolic IFS; then the slope of $\tilde{G}$ at the point whose address is $\sigma$ is the same as $\phi_{\mathcal{H}}(\sigma)$ provided the address structures the two IFSs $\tilde{G}$ and $\mathcal{H}$ are the same. It is readily verified that $\mathcal{H}$ is indeed a hyperbolic IFS. Moreover

$$w_1([-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]) = [-\frac{1}{\sqrt{3}}, 0] \text{ and } w_2([-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]) = [0, \frac{1}{\sqrt{3}}].$$

It follows that the attractor of $\mathcal{H}$ is the interval $[-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]$ and that the two IFSs $\tilde{G}$ and $\mathcal{H}$ do indeed have the same address structures. Hence, via the fractal homeomorphism theorem, $\tilde{G}$ is not only differentiable but also it is homeomorphic to the set of its slopes. Hence it is continuously differentiable. \[\Box\]

6 REFERENCES


