Self-Similar Structure in Hilbert's Space Filling Curve

Mark McClure

June 20, 2001

1 Introduction

Hilbert's space-filling curve is a continuous function that maps the unit interval onto the unit square. The construction of such curves in the 1890s surprised mathematicians of the time and led, in part, to the development of dimension theory. In this paper, we discuss how modern notions of self-similarity illuminate the structure of this curve. In particular, we show that Hilbert's curve has a basic self-similar structure and its coordinate function display a mixed self-affine structure.

The notions of self-similarity which we use are described in [1]. General information on space-filling curves may be found in [4].

2 Self-Similarity

The curve K shown in figure 1 is called the *Koch curve* and is an example of a *self-similar* set. Figure 1 shows how K is composed of 4 copies of itself, each scaled by a factor of $\frac{1}{3}$. The four copies of K have been separated slightly to illustrate this composition.

Any self-similar set may be described using an *iterated function system*, or IFS. If r is a positive real number, a *similarity* with ratio r is a function $f : \mathbb{R}^2 \to \mathbb{R}^2$ such that $|f(x) - f(y)| = r|x - y|$ for all $x, y \in \mathbb{R}^2$. If $r < 1$, the similarity is called *contractive*. An iterated function system is a finite collection of contractive similarities $\{f_i\}_{i=1}^m$. For any IFS, there is a unique

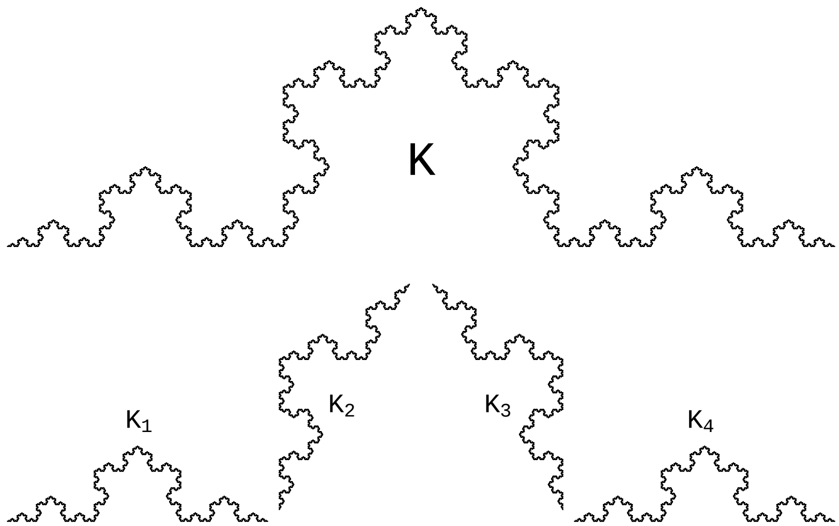

Figure 1: The Koch Curve

non-empty, closed, bounded subset E of $\mathbb{R}^2$ such that

$$E = \bigcup_{i=1}^{m} f_i(E).$$

The set E is called the *invariant set* of the IFS and sets constructed in this manner are also called self-similar.

Iterated function systems are easily described using matrix representations. Any contractive simililarity f_i may be expressed in the form $A\vec{x} + \vec{b}$, where A is a matrix and $\vec{b}$ is a translation vector. A rotation about the origin through angle θ may be represented using a matrix of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Dividing the matrix through by r achieves the desired contractivity factor. For example, the following list of functions defines the IFS for the Koch curve.

$$f_1(\vec{x}) \quad = \quad \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \vec{x}$$

$$f_2(\vec{x}) = \begin{pmatrix} \frac{1}{6} & -\frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} & \frac{1}{6} \end{pmatrix} \vec{x} + \begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix}$$

$$f_3(\vec{x}) = \begin{pmatrix} \frac{1}{6} & \frac{1}{2\sqrt{3}} \\ -\frac{1}{2\sqrt{3}} & \frac{1}{6} \end{pmatrix} \vec{x} + \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2\sqrt{3}} \end{pmatrix}$$

$$f_4(\vec{x}) = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \vec{x} + \begin{pmatrix} \frac{2}{3} \\ 0 \end{pmatrix}$$

Note that each f_i maps K onto K_i in figure 1.

3 Hilbert's Curve

Figure 2 depicts the action of an IFS with four transformations on the unit square. In Figure 2 (a), we see the unit square together with a path through four natural subsquares. In Figure 2 (b) we see the image of Figure 2 (a) under each of the four functions of the IFS. If we drop the arrows and connect the terminal point of one path to the initial point of the subsequent path, we obtain the bold path shown in Figure 2 (c). The ordering in which these paths are connected is determined by the initial path in Figure 2 (a). If we iterate this procedure two more times we obtain the fourth level approximation shown in Figure 2 (d). These paths represent approximations to a continuous function h mapping the unit interval onto the unit square. The function h is called *Hilbert's space filling curve.*

While the self-similar set determined by this iterated function system is simply the unit square, it is important to keep in mind that h is in fact a function. The unit square is the image of the unit interval under h. Given $t \in [0,1]$, we may estimate $h(t)$ as follows. If $\frac{i-1}{4^n} \leq t \leq \frac{i}{4^n}$, then $h(t)$ lies in the i^{th} closed subsquare determined by following the n^{th} level approximation. If $t = \frac{i}{4^n}$ for some i, then $h(t)$ lies on the border of two adjacent subsquares. For example, since $\frac{85}{4^4} < \frac{1}{3} < \frac{86}{4^4}$, $h(\frac{1}{3})$ lies in the 86^{th} subsquare determined by the fourth level approximation. This subsquare is shaded a light gray in the upper left corner of Figure 2 (d).

4 Mixed Self-Similar Sets

Consider the two curves A and B shown in Figure 3. The curve A is composed of one copy of itself, scaled by a factor of $\frac{1}{2}$, and two copies of B, rotated

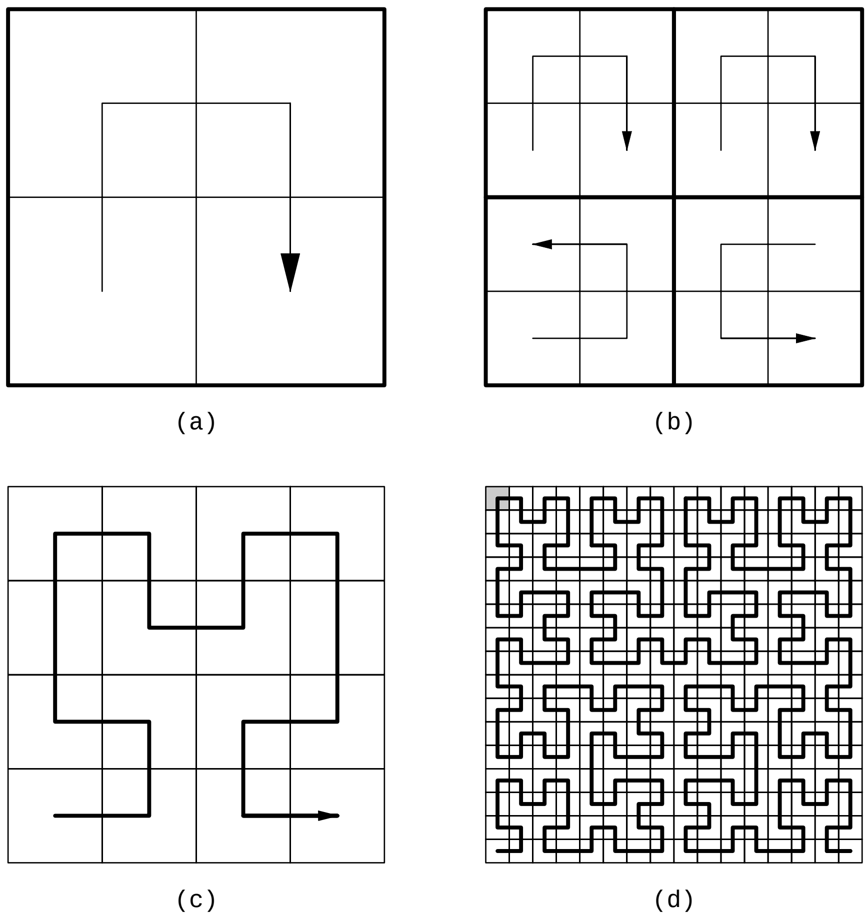

Figure 2: Approximations to Hilbert's Curve

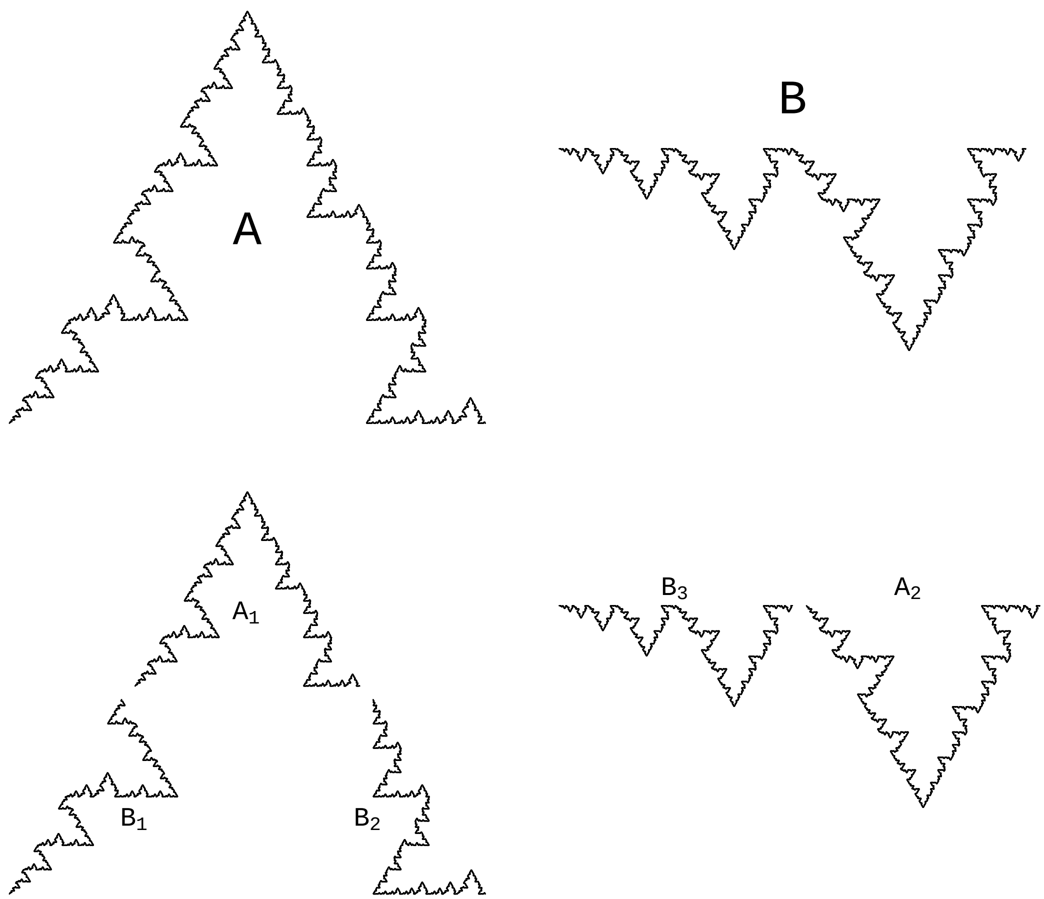

Figure 3: Mixed self-similar curves

and scaled by a factor a of $\frac{1}{2}$. The curve B is composed of one copy of itself, scaled by a factor of $\frac{1}{2}$ and one copy of A, reflected and scaled by a factor of $\frac{1}{2}$. The sets A and B form a pair of *mixed self-similar sets*.

Any collection of mixed self-similar sets can be described using a *directed-graph iterated function system*, or digraph IFS. A digraph IFS consists of a directed multigraph G together with a contractive similarity f_e from $\mathbb{R}^2$ to $\mathbb{R}^2$ associated with each edge of G. A directed multigraph consists of a finite set V of vertices and a finite set E of directed edges between vertices. Given two vertices, u and v, we denote the set of all edges from u to v by E_{uv}. Given a digraph IFS, there is a unique set of non-empty, closed, bounded sets K_v, one for each $v \in V$, such that for every $u \in V$

$$K_u = \bigcup_{v \in V,\, e \in E_{uv}} f_e(K_v).$$

The set $\{K_u : u \in V\}$ is called the *invariant list* of the digraph IFS and its members are the *invariant sets* of the digraph IFS.

The digraph IFS for the curves A and B is shown in Figure 4. The

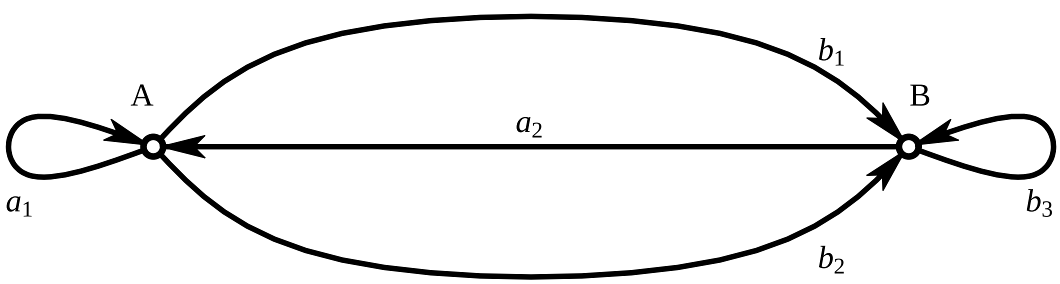

Figure 4: Digraph IFS for the curves

labels on the edges correspond to similarities mapping one set to part of another (perhaps the same) set. For example the label a_2 corresponds to the similarity mapping A to the portion of B labeled A_2 and is given by

$$a_2(\vec{x}) = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \vec{x} + \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}.$$

5 The Coordinate Functions of h

The matrix representation of a similarity mapping $\mathbb{R}^2$ to $\mathbb{R}^2$ suggests a further generalization. We may allow the linear part to be a contractive, but otherwise arbitrary, linear function. In particular, we may allow different contractivity factors in different directions. Sets generated using a digraph IFS with such functions are called *mixed self-affine sets*. If we write Hilbert's space filling curve in the form $h(t) = (x(t), y(t))$, then it turns out that the graphs of the coordinate functions $x(t)$ and $y(t)$ form such a pair of sets.

We now create a digraph IFS which defines two mixed self-affine sets X and Y. We then show that these sets coincide with the graphs of $x(t)$ and $y(t)$. Define A and B to be the following matrices.

$$A = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \qquad\qquad B = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$$

Note that the linear mappings defined by A and B both contract by a factor $\frac{1}{4}$ in the horizontal direction and by a factor $\frac{1}{2}$ in the vertical direction. B has the additional effect of reflecting about the horizontal axis. Now, let $\vec{x} \in \mathbb{R}^2$ denote a column vector and define affine functions as follows.

$$\begin{aligned}
a_{xx}(\vec{x}) &= A\vec{x} + \begin{pmatrix} 1/4 \\ 0 \end{pmatrix} & a_{yy}(\vec{x}) &= A\vec{x} + \begin{pmatrix} 1/4 \\ 1/2 \end{pmatrix} \\
b_{xx}(\vec{x}) &= A\vec{x} + \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} & b_{yy}(\vec{x}) &= A\vec{x} + \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \\
c_{xy}(\vec{x}) &= A\vec{x} & c_{yx}(\vec{x}) &= A\vec{x} \\
d_{xy}(\vec{x}) &= B\vec{x} + \begin{pmatrix} 3/4 \\ 1 \end{pmatrix} & d_{yx}(\vec{x}) &= B\vec{x} + \begin{pmatrix} 3/4 \\ 1/2 \end{pmatrix}
\end{aligned}$$

These are the affine functions used to define the digraph IFS shown in Figure 5. The sets X and Y which form the invariant list of this Digraph IFS are shown in Figure 6. The function c_{xy}, for example, maps the set Y onto the portion of X lying over the interval $[0, \frac{1}{4}]$.

We claim that the sets X and Y are the graphs of $x(t)$ and $y(t)$. To show this, we need only show that the graphs form an invariant list for the digraph IFS since such a list is unique. This may be deduced from the self-similar structure in Hilbert's construction. Note that each stage in the construction of Hilbert's curve (see Figure 2) may be obtained by piecing together four images similar to the previous stage scaled by a factor $\frac{1}{2}$. For example, the portion of Hilbert's curve in the upper left quadrant of the unit square is similar to the whole curve but scaled by the factor $\frac{1}{2}$. The affine functions a_{xx} and a_{yy} arise since h maps $[\frac{1}{4}, \frac{1}{2}]$ onto this subsquare. The other affinities may be derived in a similar manner. Note that the roles of x and y switch on

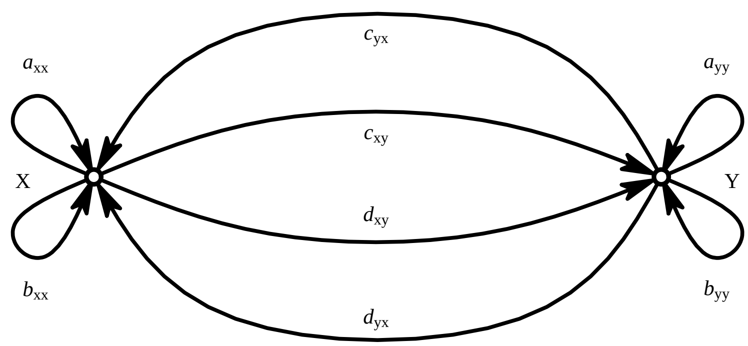

Figure 5: The digraph for X and Y

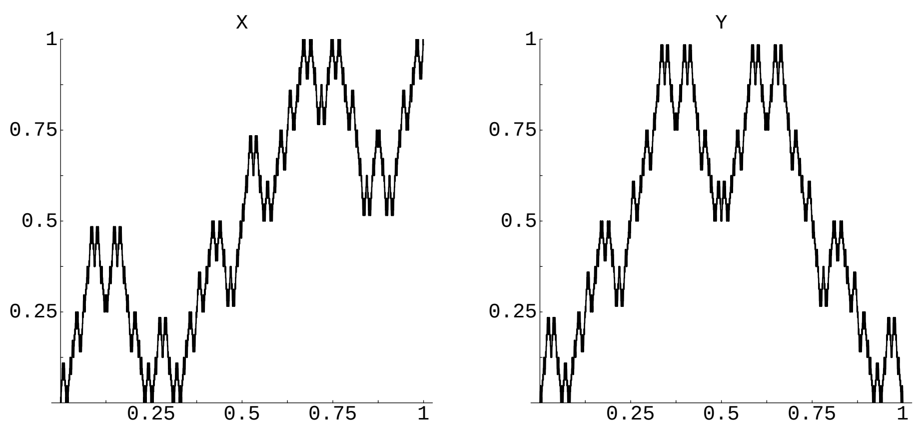

Figure 6: The sets X and Y

the lower left and lower right quadrants, due to the rotation in the similarities mapping the unit square onto those quadrants.

6 Box Counting Dimension

The digraph IFS scheme, is useful not only for generating images, but also for computing dimensions. The box counting dimension of the graph of a function $f : \mathbb{R} \to \mathbb{R}$ is a number between one and two and gives a measure of the "roughness" of the graph. An exposition of the box counting dimension may be found in [2]. In this section, we compute the box counting dimension of the sets X and Y. Another notion of dimension is called *Hausdorff dimension*. The computation of the Hausdorff dimension of X and Y may be found in [3].

To define the box counting dimension of a bounded set $S \subset \mathbb{R}^2$ we first consider covers of S by small squares. For $\varepsilon > 0$, the *ε-mesh* for $\mathbb{R}^2$ is the grid of squares of side length ε with one corner at the origin and sides parallel to the coordinate axes. For a bounded set $S \subset \mathbb{R}^2$, define

$$N_\varepsilon(S) = \text{number of } \varepsilon\text{-mesh squares that intersect } S.$$

As $\varepsilon \searrow 0$, $N_\varepsilon(S)$ usually grows larger. The rate at which $N_\varepsilon(S)$ grows reflects the dimension of S. For example, if $\mathcal{I}$ is the unit interval and $\mathcal{Q}$ is the unit square, then $N_\varepsilon(\mathcal{I})$ grows as $1/\varepsilon$ while $N_\varepsilon(\mathcal{Q})$ grows as $1/\varepsilon^2$. The exponent of ε indicates the dimension of the set. Thus we define the box counting dimension by

$$\dim_b(S) = \lim_{\varepsilon \to 0^+} \frac{\log(N_\varepsilon(S))}{\log(1/\varepsilon)},$$

provided this limit exists. An important property of $\dim_b$ (proved in [2] p. 41) is that the limit need only be taken along some sequence $\{c^n\}_{n=1}^{\infty}$ where $c \in (0,1)$ and we still obtain the same value. This simplifies its computation considerably.

To compute the box counting dimensions of X and Y, we use the rectangular covers generated by the digraph IFS which are illustrated in Figure 7. These covers are generated as follows. There are two versions of the unit square shown in Figure 7 (a), one labeled X and one labeled Y. We see how these parts fit together under the action of the digraph IFS after one iteration in Figure 7 (b) and after two iterations in Figure 7 (c). The underscores

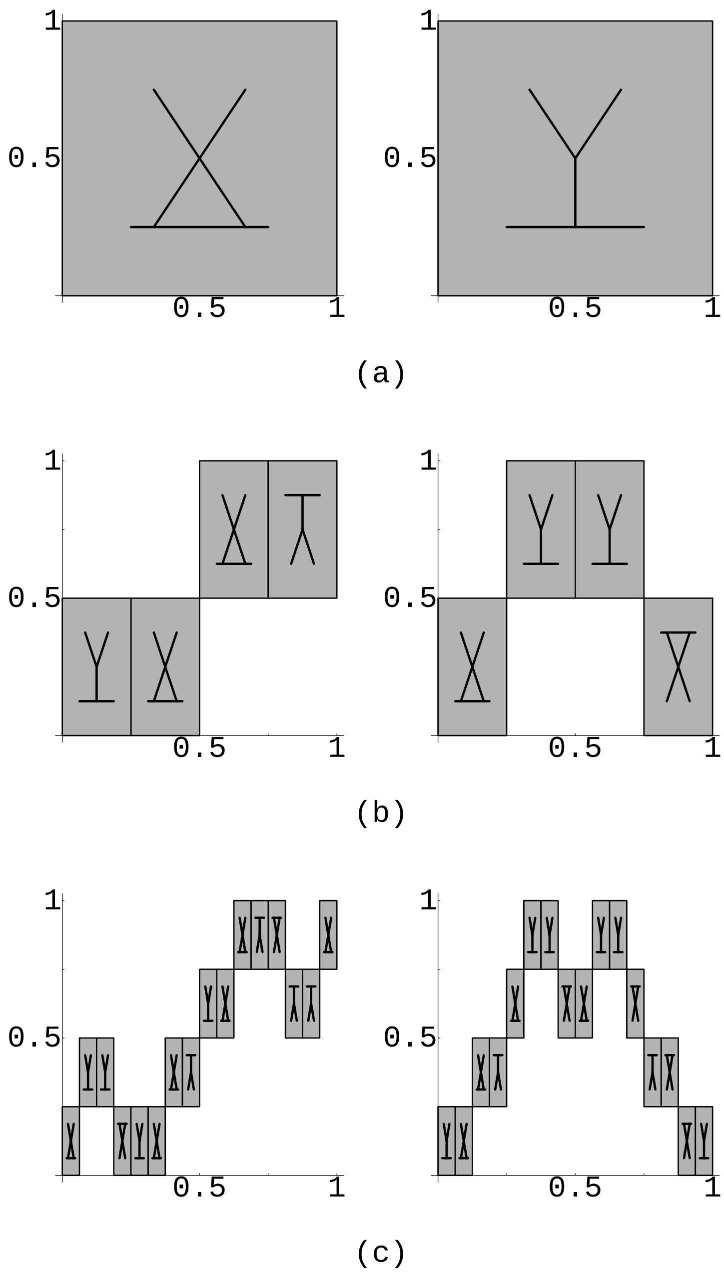

Figure 7: Approximations to X and Y using the digraph IFS

allow us to observe the orientation of the rectangles. For example, the lower left hand rectangle labeled Y in the approximation to X of Figure 7 (b) is the image of the unit square labeled Y under the affine function c_{xy}. This process is then iterated. A proof by induction shows that the rectangles of each level of the approximation are contained in the rectangles of the previous approximation. Thus these approximations in fact cover the invariant sets X and Y. It can also be proved by induction that after n iterations, each cover consists of 4^n rectangles of width 4^{-n} and height 2^{-n}. Furthermore, inside each of these rectangles is an affine image of either X or Y with height 2^{-n}. Each of the rectangles may be subdivided into a column of 2^n squares of side length 4^{-n}. Thus

$$N_{4^{-n}}(X) = N_{4^{-n}}(Y) = 2^n \cdot 4^n = 8^n$$

and

$$\dim_b(X) = \dim_b(Y) = \lim_{n\to\infty} \frac{\log(8^n)}{\log(1/4^{-n})} = \lim_{n\to\infty} \frac{\log(2^{3n})}{\log(2^{2n})} = \frac{3}{2}.$$

7 Questions

Many more space-filling curves are defined in [4]. It seems that many may be described using the techniques here. Carry out these descriptions. Better yet, identify some general principle at work. Use this principle to define some new space filling curves.

Are there any space filling curves which may not be characterized using a digraph IFS? Perhaps some modification of one of the curves described in [4] will work. Perhaps an entirely new construction will need to be found.

What are the possible dimensions of the graphs of the coordinate functions? Is there a space filling curve whose coordinate functions have box counting dimension 1? This question is related to the allowable degree of differentiability of space filling curves. Much has been written on this topic. See the references in [4].

References

[1] G. A. Edgar, *Measure, Topology, and Fractal Geometry.* Springer-Verlag, New York, NY, 1990.

[2] K. J. Falconer, *Fractal Geometry: Mathematical Foundations and Applications.* John Wiley and Sons, West Sussex, UK, 1990.

[3] Mark McClure, "The Hausdorff dimension of Hilbert's coordinate functions." *The Real Analysis Exchange.* **24** (1998/99) # 2, pp. 875-883.

[4] Hans Sagan, *Space-Filling Curves.* Springer-Verlag, New York, NY 1994.