Dimensions of the coordinate functions of space-filling curves

Pieter C. Allaart and Kiko Kawamura *

September 28, 2006

Abstract

The graphs of coordinate functions of space-filling curves such as those described by Peano, Hilbert, Pólya and others, are typical examples of self-affine sets, and their Hausdorff dimensions have been the subject of several articles in the mathematical literature. In the first half of this paper, we describe how the study of dimensions of self-affine sets was motivated, at least in part, by these coordinate functions and their natural generalizations, and review the relevant literature. In the second part, we present new results on the coordinate functions of Pólya's one-parameter family of space-filling curves. We give a lower bound for the Hausdorff dimension of their graphs which is fairly close to the box-counting dimension. Our techniques are largely probabilistic. The fact that the exact dimension remains elusive seems to indicate the need for further work in the area of self-affine sets.

AMS 2000 subject classification: 28A78, 28A80 (primary); 26A27 (secondary)

Key words and phrases: Hausdorff dimension, Box-counting dimension, Self-affine set, Space-filling curve

*Mathematics Department, P.O. Box 311430, Denton, TX 76203-1430; e-mail: allaart@unt.edu, kiko@unt.edu

1 Introduction

Until the late nineteenth century, mathematicians used the word 'dimension' in only a vague sense: we understand intuitively sets such as a line, a square and a cube, and have no hesitation to say that their dimensions are one, two and three respectively.

However, Peano's demonstration in 1890 [21] of a continuous map from the unit interval onto the unit square (now called *Peano's space-filling curve*) showed convincingly that this intuitive notion of dimension could not be satisfactory mathematically. Since the continuous image of a line segment could fill a square, it was no longer clear how to classify sets in the plane by their dimension. In particular, the old belief that the dimension of a space could be defined as the least number of continuous real parameters needed to describe the space became instantly obsolete.

In the following decades, further examples of space-filling curves were given by Hilbert [6], Lebesgue [16], Sierpiński [25], Pólya [22], and others. The resulting crisis led Urysohn, Menger and others in the 1920s to develop a rigorous definition of topological dimension. Furthermore, several other dimensions were introduced to measure general sets, including Hausdorff dimension and box-counting dimension. See Hurewicz and Wallman [7] for an account of dimension theory.

Almost a century after Peano's discovery, space-filling curves enjoyed a period of renewed interest, and again dimension played a central role. The triggering event was Mandelbrot's extensive work on fractals (e.g. [17]), and the subsequent development of a theory of self-similar and self-affine sets and self-affine functions. Coordinate functions of space-filling curves turned out to be prime examples of self-affine functions, and hence became models for more general classes of self-affine functions and self-affine sets. Computing the (usually fractional) dimensions of such sets became a central goal.

This paper reviews the development of the theory of self-affine sets from the point of view of their relationship to the coordinate functions of space-filling curves. The emphasis will be on the calculation of Hausdorff and box-counting dimensions. The second part of the paper presents new results on the dimensions of the graphs of the coordinate functions of Pólya's space-filling curve. These graphs, which do not fall in any of the general classes of self-affine sets described in the first half of the paper, make it clear that much work remains to be done in the study of self-affine sets.

The organization of this paper is as follows. Section 2 recalls the definitions of Hausdorff and box-counting dimensions. Section 3 describes McMullen's results on generalized Sierpinski carpets, considered the most fundamental work on self-affine sets. In section 4 we discuss the space-filling curves of Peano, Hilbert, and Lebesgue. We show that each of these inspired, perhaps indirectly or subconsciously, the creation of a general class of self-affine functions, which in turn gave rise to an appropriate generalization of McMullen's dimension formulas. In section 5, we consider the coordinate functions of Pólya's one-parameter family of space-filling curves, which map an interval onto a right triangle. We obtain the box-counting dimension as a function of the parameter, and give a fairly sharp lower bound function for the Hausdorff dimension. Unfortunately – except for the relatively easy case when the triangle is isosceles – the exact Hausdorff dimension remains elusive, and some radically new idea appears to be needed to analyze self-affine functions of the complexity of Pólya's coordinate functions.

This paper is not intended as a first introduction to space-filling curves; nor does it cover all the known space-filling curves. For an excellent general treatise on the subject, we refer to Sagan [24].

2 Hausdorff and box-counting dimensions

We briefly recall the definitions of Hausdorff and box-counting dimensions. For a broader introduction, however, see Falconer [4].

For a set F in $\mathbf{R}^n$, let $|F|$ denote the diameter of F. Let $s \geq 0$. For

$F \subset \mathbf{R}^n$ and $\delta > 0$, define

$$\mathcal{H}^s_\delta(F) := \inf \left\{ \sum_{i \in I} |U_i|^s : F \subseteq \bigcup_{i \in I} U_i \text{ and } |U_i| < \delta \text{ for every } i \in I \right\},$$

where I is understood to be countable. The s-dimensional Hausdorff measure of F is defined by

$$\mathcal{H}^s(F) := \lim_{\delta \to 0} \mathcal{H}^s_\delta(F),$$

and the *Hausdorff dimension* of F is the number

$$\dim_H F := \sup\{s \geq 0 : \mathcal{H}^s(F) = \infty\} = \inf\{s \geq 0 : \mathcal{H}^s(F) = 0\}.$$

Another common dimension to measure a fractal set F is its *box-counting dimension*, defined by

$$\dim_B F := \lim_{\delta \to 0} \frac{\log N(\delta)}{\log(1/\delta)}, \tag{1}$$

where $N(\delta)$ is the minimum number of δ-balls needed to cover F. If the limit does not exist, one considers upper and lower box-counting dimensions, denoted $\overline{\dim}_B F$ and $\underline{\dim}_B F$, and defined by taking lim sup and lim inf, respectively in (1). It is well known that $\dim_H F \leq \underline{\dim}_B F$ for any set $F \subset \mathbf{R}^n$.

3 McMullen carpets

A *self-affine set* in $\mathbf{R}^2$ is a nonempty compact set E which satisfies a set equation of the form

$$E = \psi_1(E) \cup \cdots \cup \psi_n(E), \tag{2}$$

where $\psi_1, \ldots, \psi_n$ are affine contractions of $\mathbf{R}^2$. If the ψ_i are similarities, E is said to be *self-similar*. Hutchinson [8] proved that (2) has a unique nonempty solution, and established a formula for $\dim_H E$ in case E is self-similar and satisfies the so-called open set condition.

In 1984, McMullen [19] generalized Hutchinson's result to a family of self-affine sets constructed as follows. Let S be the unit square, and choose integers $1 < m \leq n$. Draw $n-1$ vertical lines and $m-1$ horizontal lines to partition S into mn congruent rectangles, arranged in m 'rows' and n 'columns'. For $i = 0, \ldots, m-1$ and $j = 0, \ldots, n-1$, let $\psi_{i,j}$ be the orientation preserving affine contraction which maps S onto the rectangle in row i and column j. Thus,

$$\psi_{i,j}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} n^{-1} & 0 \\ 0 & m^{-1} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} j/n \\ i/m \end{pmatrix}. \tag{3}$$

Let R be a nonempty subset of $\{0, \ldots, m-1\} \times \{0, \ldots, n-1\}$. Then there exists a unique nonempty compact set $M \subset S$ satisfying the set equation

$$M = \bigcup_{(i,j)\in R} \psi_{i,j}(M).$$

Following Kenyon and Peres [11], we shall call M a *McMullen carpet*. Note that M can be approximated iteratively by putting $M_0 = S$, and for $k \geq 0$, $M_{k+1} = \bigcup_{(i,j)\in R} \psi_{i,j}(M_k)$. An example is shown in Figure 1.

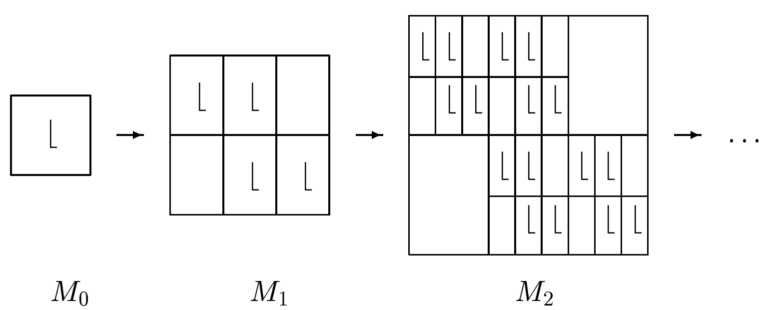

Figure 1: An example of a McMullen carpet with $m = 2, n = 3$.

McMullen's main result is that

$$\dim_H M = \log_m \left(\sum_{i=0}^{m-1} t_i^{\log_n m} \right), \tag{4}$$

where $t_i = \#\{(p,q) \in R : p = i\}$, the number of rectangles contained in row i of the generating pattern. McMullen also gave a formula for the box-counting dimension, and showed that $\dim_H M = \dim_B M$ if and only if each nonempty row in the generating pattern contains the same number of rectangles; that is, if there exists a constant l such that $t_i \in \{0, l\}$ for $i = 0, \dots, m-1$.

For example, the set determined by the pattern in Figure 1 has Hausdorff and box-counting dimension $1 + \log_3 2$. McMullen's work was extended in various directions by Falconer [3, 5], Lalley and Gatzouras [14], Kenyon and Peres [11] and Takahashi [26].

4 Peano's, Hilbert's and Lebesgue's curves

In this section, we review three famous space-filling curves and give the Hausdorff dimension of the graphs of their coordinate functions. Let $I = [0,1]$ denote the closed unit interval, and $S = [0,1] \times [0,1]$ the closed unit square.

4.1 Peano's coordinate functions

In 1890, Peano [21] constructed the first continuous mapping from the unit interval onto a square, which is now called Peano's space-filling curve. Ten years later, Moore [20] proved that the Peano curve is nowhere differentiable.

Divide S into nine congruent subsquares, and let ϕ_j $(j = 0, 1, \dots, 8)$ be the similar contractions which map S onto each subsquare in the order and with the orientations shown in Figure 2.

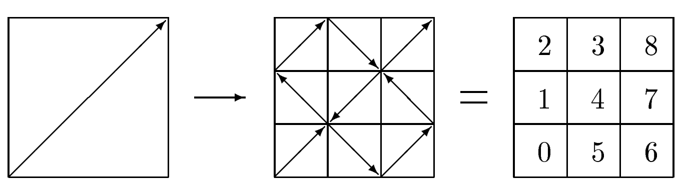

Figure 2: Construction of the Peano curve.

For $t \in I$, let $t = (0.t_1 t_2 t_3 \ldots)_9$ denote the nonary expansion of t. Since the ϕ_i are contractions and S is compact, the intersection

$$\bigcap_{n=1}^{\infty} \phi_{t_1} \circ \cdots \circ \phi_{t_n}(S)$$

consists of a single point, which we denote by $P(t)$. Brief reflection shows that each point of S can be obtained in this manner; hence the mapping P is surjective.

Note that some points t have two different nonary expansions. It is not difficult to convince oneself that both expansions yield the same point in S.

Let $x(t)$ and $y(t)$ denote the coordinate functions of $P(t)$, and denote their graphs by X and Y, respectively. It is known that $x(t)$ and $y(t)$ are continuous but nowhere differentiable and satisfy the functional equations

$$x\left(\frac{t+j}{9}\right) - x\left(\frac{j}{9}\right) = \begin{cases} \frac{x(t)}{3}, & \text{if } j = 0,2,3,5,6,8, \\ -\frac{x(t)}{3}, & \text{if } j = 1,4,7, \end{cases}$$

and

$$y\left(\frac{t+j}{9}\right) - y\left(\frac{j}{9}\right) = \begin{cases} \frac{y(t)}{3}, & \text{if } j = 0,1,2,6,7,8, \\ -\frac{y(t)}{3}, & \text{if } j = 3,4,5, \end{cases}$$

for $0 \le t \le 1$, with boundary values $x(0) = y(0) = 0$, and $x(1) = y(1) = 1$.

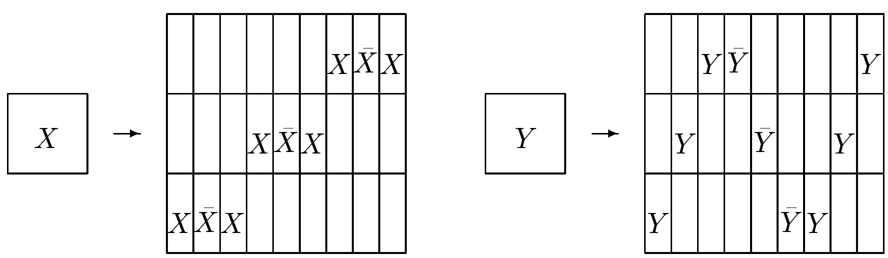

Figure 3: The structure of the graphs of Peano's coordinate functions. A bar indicates a top-to-bottom reflection.

See Figure 3. The graph of each coordinate function is a self-affine set constructed by affine maps from the unit square to rectangles of width 1/9 and height 1/3. However, since some of the affine maps involve vertical reflections, the sets X and Y are not quite McMullen carpets. Thus, McMullen's formula does not apply, or so it seems.

In 1986 Kono [12], presumably inspired by Peano's space-filling curve, introduced the following class of self-affine functions. Kono called a function $f : [0,1] \to \mathbf{R}$ *self-affine* if there exist positive integers $m, n > 1$ and constants $\varepsilon_j \in \{-1, 1\}$ $(j = 0, \dots, n-1)$ such that

$$f\left(\frac{t+j}{n}\right) - f\left(\frac{j}{n}\right) = \varepsilon_j \frac{f(t)}{m}$$

for $0 \le t \le 1$ and $j = 0, 1, \dots, n-1$.

Peano's coordinate functions clearly satisfy the above functional form. Kono studied the Hausdorff dimension of the graphs of self-affine functions under certain restrictions, proving in particular that the graphs of Peano's coordinate functions have Hausdorff dimension 3/2. Later, Urbanski [27] gave a general formula for the Hausdorff dimension of the graph of any continuous self-affine function in the sense of Kono with $f(0) = 0, f(1) = 1$. It is quite similar to McMullen's formula.

Note that, compared to McMullen's carpets, the graphs of self-affine functions in the sense of Kono are generated by affine mappings of the form

$$\psi_j \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} n^{-1} & 0 \\ 0 & \pm m^{-1} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} j/n \\ f(j/n) \end{pmatrix}, \qquad j = 0, \dots, n-1.$$

4.2 Hilbert's coordinate functions

Although it was Peano who discovered the first space-filling curve, Hilbert [6] was the first to outline a general geometrical procedure that allowed the construction of an entire class of space-filling curves. The simplest example is sketched below.

Divide the unit square S into four congruent subsquares, and let h_0, h_1, h_2 and h_3 be the similar contractions which map S onto each subsquare in the order and with the orientations shown in Figure 4.

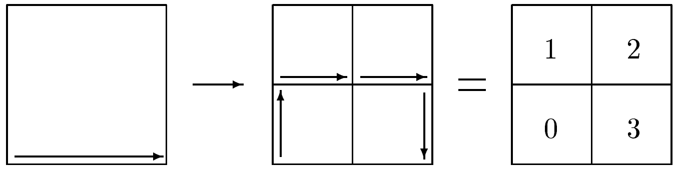

Figure 4: Construction of Hilbert's curve.

For $t \in I$, let $t = (0.t_1 t_2 t_3 \ldots)_4$ denote the quaternary expansion of t. As before, the intersection

$$\bigcap_{n=1}^{\infty} h_{t_1} \circ \cdots \circ h_{t_n}(S)$$

consists of a single point, which we denote by $H(t)$. As with Peano's curve, the image $H(t)$ is independent of the choice of quaternary expansion for t when t has two quaternary expansions.

Denote the coordinate functions of $H(t)$ by $x(t)$ and $y(t)$. They are continuous but nowhere differentiable, and satisfy the following functional equations:

$$x\left(\frac{t+j}{4}\right) - x\left(\frac{j}{4}\right) = \begin{cases} \frac{y(t)}{2}, & \text{if } j = 0, \\ \frac{x(t)}{2}, & \text{if } j = 1, \\ \frac{x(t)}{2}, & \text{if } j = 2, \\ -\frac{y(t)}{2}, & \text{if } j = 3, \end{cases} \tag{5}$$

and

$$y\left(\frac{t+j}{4}\right) - y\left(\frac{j}{4}\right) = \begin{cases} \frac{x(t)}{2}, & \text{if } j = 0, \\ \frac{y(t)}{2}, & \text{if } j = 1, \\ \frac{y(t)}{2}, & \text{if } j = 2, \\ -\frac{x(t)}{2}, & \text{if } j = 3, \end{cases} \tag{6}$$

for $0 \le t \le 1$. The boundary values are $x(0) = y(0) = y(1) = 0$, and $x(1) = 1$.

Let X denote the graph of $x(t)$ and Y the graph of $y(t)$. By (5) and (6), X and Y each consist of two affine contracted images of X, and two of Y, as shown in Figure 5. Thus, X and Y are not among the sets studied by McMullen or Kono.

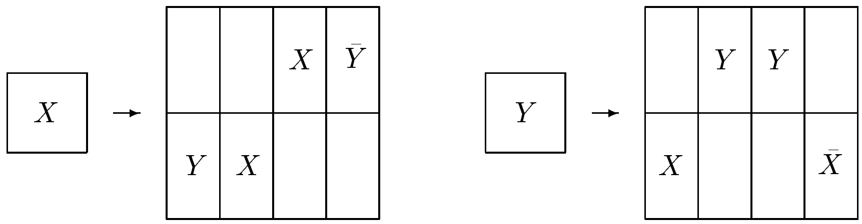

Figure 5: The structure of the graphs of Hilbert's coordinate functions. A bar indicates a top-to-bottom reflection.

However, Kono [13] pointed out that Hilbert's coordinate functions are self-affine in the sense of Kamae. A function $f : [0,1] \to \mathbf{R}$ is self-affine in the sense of Kamae [9] if the following conditions are satisfied:

1. There is a finite number of continuous functions $f_1, f_2, \ldots, f_N : [0,1] \to \mathbf{R}$ with $f_l(0) = 0$ for all $l = 1, 2, \ldots, N$, and $f_1 = f$.

2. There exist positive integers $m, n > 1$, and for each $l \in \{1, 2, \ldots, N\}$ and each $j \in \{0, 1, \ldots, n-1\}$, there exists $k \in \{1, 2, \ldots, N\}$ such that
$$f_l\left(\frac{t+j}{n}\right) - f_l\left(\frac{j}{n}\right) = \frac{f_k(t)}{m}, \qquad 0 \le t \le 1.$$

Motivated by Kamae's work, Kenyon and Peres [11] and Takahashi [26] introduced a class of generalized self-affine sets with N patterns which includes the graphs of Kamae's functions under some restrictions. Both papers give a formula to calculate the Hausdorff dimension of these sets, but for ease of presentation we follow Takahashi [26].

Let $1 < m \le n$ be integers, and divide the unit square into m rows and n columns as in Section 3. Let $\psi_{i,j}$ be the affine maps of (3). For

$l = 1, 2, \ldots, N$, let g_l be a map from $\{0, \ldots, m-1\} \times \{0, \ldots, n-1\}$ to $\{0, 1, \ldots, N\}$.

Next, let $X_0 = \emptyset$, and let $\{X_1, X_2, \ldots, X_N\}$ be a family of non-empty compact sets which satisfies

$$X_l = \bigcup_{i,j} \psi_{i,j}(X_{g_l(i,j)}), \qquad l = 1, \ldots, N.$$

Put $X := X_1$, and assume that $\{X_1, X_2, \ldots, X_N\}$ is *irreducible*: for each pair of indices (l, l'), X_l contains an affine contracted image of $X_{l'}$. For integers $0 \le y_1, \ldots, y_k < m$, write $y_1 \ldots y_k = \sum_{i=1}^k y_i m^{k-i}$, and define

$$N(y_1 \ldots y_k) = \#\{q : \left(\frac{q}{n^k}, \frac{q+1}{n^k}\right) \times \left(\frac{y_1 \ldots y_k}{m^k}, \frac{y_1 \ldots y_k + 1}{m^k}\right) \cap X \neq \emptyset\}.$$

That is, $N(y_1 \ldots y_k)$ is the number of affine contracted images of $X_1, \ldots, X_N$ contained in the 'row' $[0,1] \times [\frac{y_1 \ldots y_k}{m^k}, \frac{y_1 \ldots y_k + 1}{m^k}]$ of X. Then

$$\dim_H X = \lim_{k \to \infty} \frac{1}{k} \log_m \left[\sum_{y_1 \ldots y_k} N(y_1 \ldots y_k)^{\log_n m} \right]. \tag{7}$$

(A formula for the box-counting dimension of X is known as well: see Kenyon and Peres [11].)

The graphs of Hilbert's coordinate functions are an example of the above set-up: take $n = 4$, $m = 2$, $X_1 = X$, $X_2 = Y$, $X_3 = \bar{X}$, and $X_4 = \bar{Y}$. Figure 5 shows that the system is irreducible, and (7) yields that these graphs have Hausdorff dimension 3/2. This was shown independently by McClure [18] using a different method. In fact, McClure obtained the stronger result that the $\mathcal{H}^{3/2}$-measure of X and Y is strictly positive and finite.

In general, calculating the limit in (7) can be difficult. Kenyon and Peres [11] explain how spectral theory can sometimes be used to do this. Moreover, they point out how (7) reduces to Urbanski's formula in the case of graphs of Kono's self-affine functions.

4.3 Lebesgue's space-filling curve

So far, the coordinate functions we have encountered all had Hausdorff dimension 3/2. A natural question is: Which numbers can be the Hausdorff dimensions of coordinate functions of a space-filling curve? To answer this question, we start with a very different type of space-filling curve, proposed in 1904 by Lebesgue [16] and based on the middle-third Cantor set C. For a point $t \in C$, express t by its ternary expansion: $t = (0.t_1 t_2 t_3 \dots)_3$, where $t_i \in \{0, 2\}$, $i \in \mathbf{N}$. Define

$$x(t) = \frac{1}{2}\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j t_{2j-1}, \qquad y(t) = \frac{1}{2}\sum_{j=1}^{\infty}\left(\frac{1}{2}\right)^j t_{2j}. \tag{8}$$

Then it is almost obvious that the function $L(t) = (x(t), y(t))$ maps C onto the unit square S. To extend L to all of I, Lebesgue used linear interpolations in the intervals that were removed in the construction of C. It is intuitively plausible that the resulting function $L : I \to S$ is continuous, but the precise proof requires some technicalities; see Theorem 5.4.1 of [24].

Clearly, $L(t)$ is differentiable at every point of $I \backslash C$. On the other hand, it can be shown that $L(t)$ is not differentiable at any point of C. The main impact of Lebesgue's curve was that it put an end to the belief, based on Peano's and Hilbert's examples, that space-filling curves must necessarily be nowhere differentiable.

We will show that the graphs of $x(t)$ and $y(t)$ have Hausdorff and box-counting dimension $1 + \log_9 2 = 1.315 \cdots$. Define the planar sets

$$X = \{(t, x(t)) : t \in I\}, \qquad Y = \{(t, y(t)) : t \in I\},$$
$$\hat{X} = \{(t, x(t)) : t \in C\}, \qquad \hat{Y} = \{(t, y(t)) : t \in C\}.$$

The equations (8) imply that $\hat{X}$ and $\hat{Y}$ are McMullen carpets with $n = 9$ and $m = 2$, as illustrated in Figure 6.

McMullen's formula (4) yields

$$\dim_H \hat{X} = \log_2\left(2^{\log_9 2} + 2^{\log_9 2}\right) = 1 + \log_9 2,$$

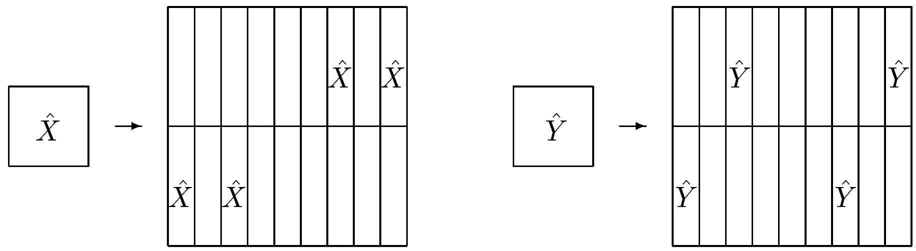

Figure 6: The McMullen carpets representing the coordinate functions of Lebesgue's space-filling curve, restricted to the Cantor set C.

with the same value for the box-counting dimension since each 'row' in the diagram in Figure 6 contains two affine images of $\hat{X}$. Since $X \backslash \hat{X}$ consists of a countable union of open line segments, its box-counting dimension is 1. Hence,

$$\dim_H X = \dim_B X = 1 + \log_9 2.$$

The same argument applies to Y.

Generally, if C is replaced by a Cantor set of dimension δ $(0 < \delta < 1)$, the resulting graphs of $x(t)$ and $y(t)$ have (box and Hausdorff) dimension $1 + \frac{\delta}{2}$. This shows that a wide range of dimensions is possible for the graphs of coordinate functions of space-filling curves. In fact, any two numbers d_1 and d_2 in the interval $[1, 2]$ can be the dimensions of the two coordinate functions of a space-filling curve. To see this, let C_0 be a 'thin' Cantor set of dimension zero. In the same way as above, C_0 can be mapped continuously onto the unit square, with coordinate functions $x(t)$ and $y(t)$. Now consider a particular interval, say $[a, b]$, that was removed in the construction of C_0. Choose continuous functions $\phi : [a, b] \to [0, 1]$ and $\psi : [a, b] \to [0, 1]$ whose graphs have Hausdorff dimensions d_1 and d_2, respectively and such that $\phi(a) = x(a)$, $\phi(b) = x(b)$, $\psi(a) = y(a)$ and $\psi(b) = y(b)$. Now extend the definitions of $x(t)$ and $y(t)$ to (a, b) by putting $x(t) = \phi(t)$, $y(t) = \psi(t)$. On the other removed intervals, simply define $x(t)$ and $y(t)$ by linear interpolation, as before. Since the restrictions of the graphs of $x(t)$ and

$y(t)$ to C_0 have dimension one, the full graphs have dimensions d_1 and d_2, respectively.

5 Pólya's space-filling curves

In 1912, Sierpiński [25] proposed a continuous mapping from the unit interval onto an isosceles right triangle. Sierpiński's space-filling curve was studied further by Knopp in 1917, and it is now known as the Sierpiński-Knopp curve. In 1913, Pólya [22] generalized Sierpiński's example to obtain a one-parameter family of space-filling curves, which we now describe in detail. Let Δ be any right triangle, and divide it into two subtriangles Δ_0 and Δ_1, each similar to Δ, as shown in Figure 7.

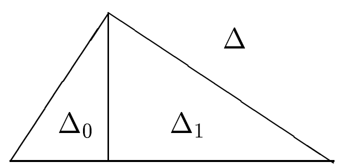

Figure 7: The right triangle Δ, and its similar subtriangles Δ_0 and Δ_1.

Let ϕ_0 and ϕ_1 be the affine transformations which map Δ onto Δ_0 and Δ_1, respectively. For $t \in I$, let $t = (0.t_1t_2t_3\ldots)_2$ denote the binary expansion of t. Since ϕ_0 and ϕ_1 are contractions and Δ is compact, the intersection

$$\bigcap_{n=1}^{\infty} \phi_{t_1} \circ \cdots \circ \phi_{t_n}(\Delta)$$

consists of a single point; denote it by $\Pi(t)$. It is easy to see that the function Π thus defined maps I onto Δ, and that, for those numbers t having two binary expansions, $\Pi(t)$ does not depend on the choice of expansion. With some additional effort, it may be seen that Π is continuous (see [22]).

The main novelty of Pólya's curve was that, unlike the curves of Peano and Hilbert which have quadruple points, Pólya's curve has at most triple points whenever the ratio of the length of the shorter side over the length of the hypotenuse is transcendental. (See [22] or [24, Th. 4.6].)

Much later, another surprising difference was discovered: Pólya's curve has a much more subtle differentiability structure. Denote by θ the smallest of the two acute angles of Δ. Lax [15] and Bumby [2] showed that $\Pi(t)$ is nowhere differentiable if $30° \leq \theta \leq 45°$; that it is non-differentiable almost everywhere and has derivative zero on an uncountable set if $15° \leq \theta < 30°$; and that it has derivative zero almost everywhere if $\theta < 15°$. Prachar and Sagan [23] proved that the same is true for the coordinate functions of $\Pi(t)$.

Clearly, there is more than one natural choice for a coordinate system here, with different choices yielding different pairs of coordinate functions for $\Pi(t)$. We choose a frame with the hypotenuse of Δ lying along the positive x-axis, as in Figure 8. However, the results presented below are valid for any rectangular coordinate frame.

Fix the length of the hypotenuse to be 1, and let a denote the abscissa of the altitude separating Δ_0 and Δ_1; see Figure 8. Let $x(t)$ and $y(t)$ denote the corresponding coordinates of $\Pi(t)$. Figure 9 shows their graphs for selected values of a. As can be seen, the graphs become more and more irregular as a decreases.

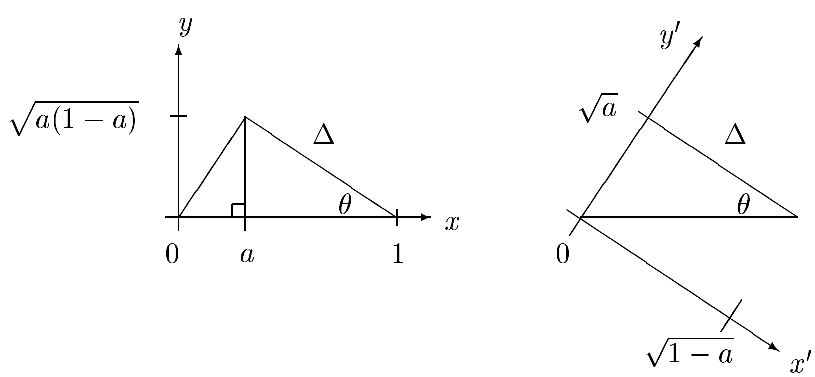

Figure 8: Two coordinate systems for Pólya's space-filling curve.

It will be convenient to consider a second coordinate system (x', y'), with

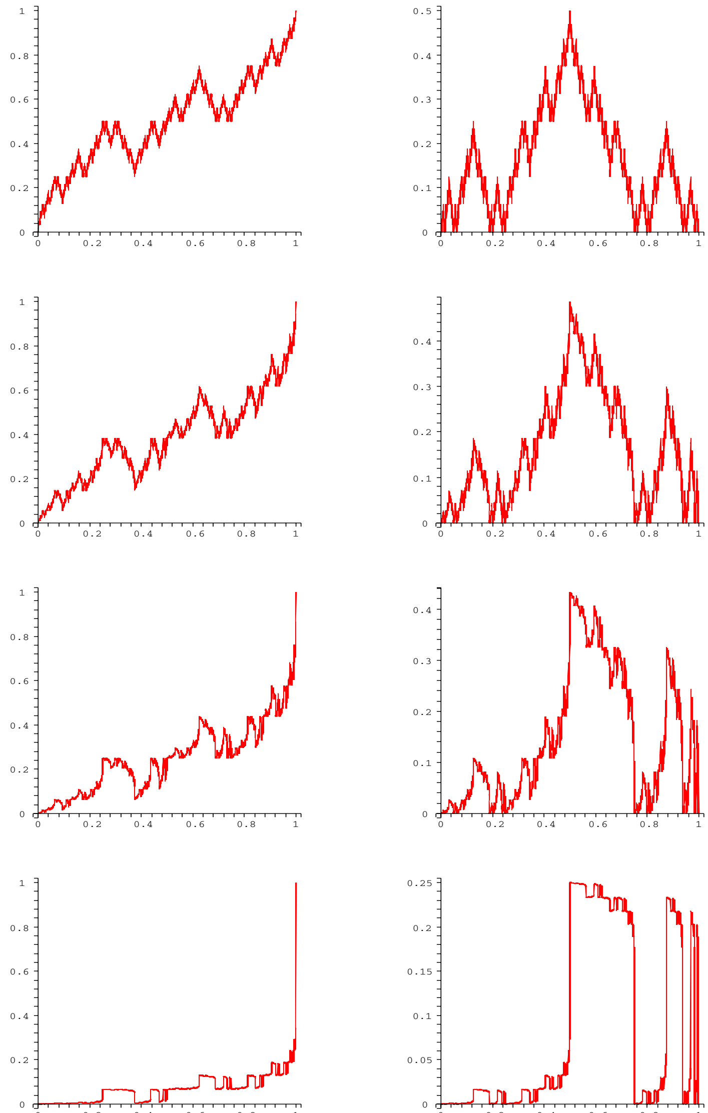

Figure 9: Graphs of $x(t)$ (left column) and $y(t)$ (right column, with variable vertical scale). *Top*: $a = 0.5$, when $\theta = 45^\circ$ (the Sierpiński-Knopp curve). *Second from top*: $a = (3-\sqrt{5})/2 \approx 0.3819$, when $a/(1-a)$ is the golden mean. *Second from bottom*: $a = 0.25$, when $\theta = 30^\circ$. *Bottom*: $a = (2-\sqrt{3})/4 \approx 0.067$, when $\theta = 15^\circ$.

axes parallel to the two sides of Δ abutting the right angle; see Figure 8. Note that the (x', y')-frame is obtained by rotating the (x, y)-frame clockwise by θ. Let $x'(t)$ and $y'(t)$ denote the corresponding coordinate functions.

By continuity of $\Pi(t)$, the graphs of $x(t)$, $y(t)$, $x'(t)$ and $y'(t)$ are compact connected sets in the plane. Denote them by X, Y, X' and Y', respectively.

The following two theorems are our main results.

Theorem 5.1 *The box-counting dimension d_B of X, Y, X' and Y' is given by*

$$d_B = 1 + \log_2(\sqrt{a} + \sqrt{1-a}) = 1 + \log_2(\sin\theta + \cos\theta).$$

Theorem 5.2 *The Hausdorff dimension d_H of X, Y, X' and Y' satisfies the inequalities*

$$1 + [-a\log_4 a - (1-a)\log_4(1-a)] \leq d_H \leq 1 + \log_2(\sqrt{a} + \sqrt{1-a}).$$

For the case $a = 1/2$ (the Sierpiński-Knopp curve), Theorem 5.2 gives the exact Hausdorff dimension $d_H = 3/2$, but for other values of a we must be content with estimates. As Figure 10 shows, the bounds are quite close when a is not too far from $1/2$. The largest distance between the bounds is about 0.12, and occurs around $a \approx 0.05$. It would be of interest to know whether the calculation of d_H breaks into different cases along the same lines as the differentiability of the curve $\Pi(t)$. In fact, we do not even know whether d_H varies continuously with a.

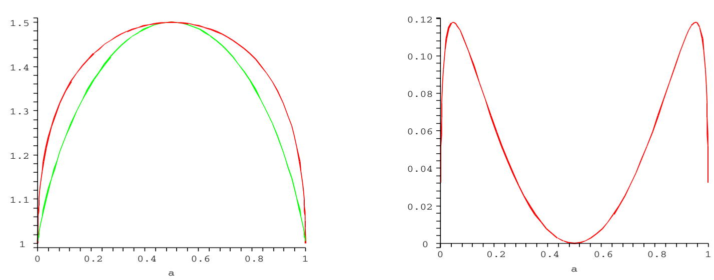

Figure 10: The bounds of Theorem 5.2 (left), and their difference (right).

The rest of the paper is devoted to the proofs of Theorems 5.1 and 5.2. A first useful observation is that the graphs X, Y, X' and Y' each consist of affine contracted images of two of the other three graphs. To see this, note that the left and right halves of each graph trace the corresponding coordinate function as $\Pi(t)$ traverses the smaller subtriangle Δ_0 and the larger subtriangle Δ_1, respectively. Recall that $\Delta_0 = \phi_0(\Delta)$ and $\Delta_1 = \phi_1(\Delta)$. Since the x-axis is the image of the y'-axis under ϕ_0 and is parallel to the image of the x'-axis under ϕ_1, we see that the left half of X is a contracted image of Y', and the right half is a contracted image of X'. This is shown in the top left diagram in Figure 11. Similarly, the y-axis is the image of the x'-axis under ϕ_0 and is parallel to the image of the y'-axis under ϕ_1, with the orientation reversed. Hence Y consists of a contracted image of X' and one of Y', the latter reflected vertically; see the top right diagram in Figure 11. The remaining two diagrams can be understood similarly. (Observe that the point where the altitude meets the base of Δ has (x, y)-coordinates $(a, 0)$, and (x', y')-coordinates $(a\sqrt{a}, a\sqrt{1-a})$.)

Note that the rectangles containing the affine images of X, Y, X' and Y' do not "line up" in horizontal strips, as was the case in sections 3 and 4. This severely complicates the determination of the Hausdorff dimension.

To simplify notation, write

$$f_1(t) = x(t), \qquad f_2(t) = y(t), \qquad f_3(t) = x'(t), \qquad f_4(t) = y'(t),$$

and let F_i denote the graph of f_i, for $i = 1, \dots, 4$. Let $\mathcal{V} = \{1, 2, 3, 4\}$. We can represent the relationships in Figure 11 by a tuple $(\mathcal{G}, r)$ where $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is the directed graph with vertex set $\mathcal{V}$ and set of edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ shown in Figure 12, and r is a labeling on $\mathcal{E}$. A directed edge $e = (u, v)$ from u to v $(u, v \in \mathcal{V})$ indicates that F_u contains an affine image of F_v. The corresponding affine mapping is denoted by T_e or $T_{u,v}$, and the label $r(e)$ of e is the contraction ratio of T_e in the vertical direction. Formally, we have

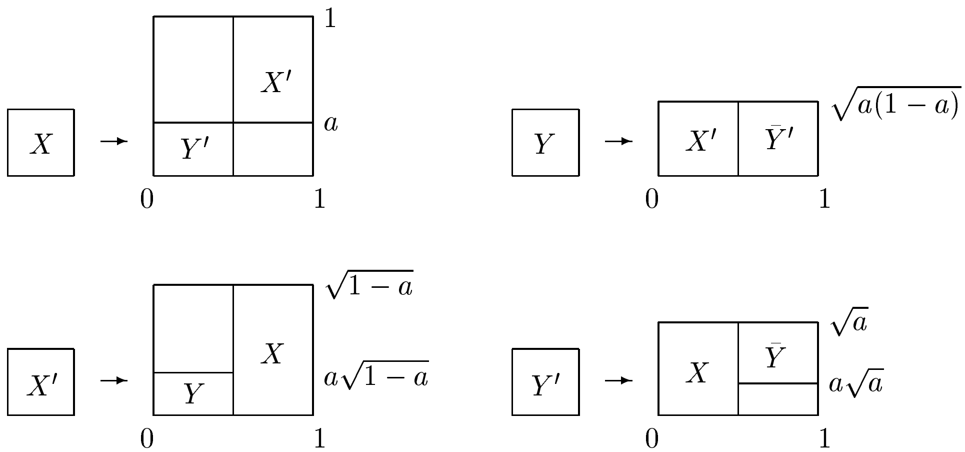

Figure 11: The relationships between X, Y, X' and Y'. A bar indicates a top-to-bottom reflection.

the relationship

$$F_u = \bigcup_{v:(u,v)\in\mathcal{E}} T_{u,v}(F_v), \qquad u \in \mathcal{V}.$$

Observe that the similarity maps ϕ_0 and ϕ_1 have contraction ratios $\sqrt{a}$ and $\sqrt{1-a}$, respectively. Thus, from each vertex $v \in \mathcal{V}$ emanate exactly two edges, say e and e', such that $r(e) = \sqrt{a}$ and $r(e') = \sqrt{1-a}$.

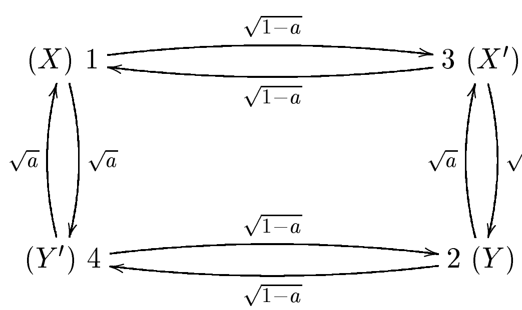

Figure 12: The digraph $\mathcal{G}$, labeled with the vertical contraction ratios of Figure 11.

Proof of Theorem 5.1. Since the digraph $\mathcal{G}$ is homogeneous and the affine mappings T_e do not involve "shears", the box-counting dimension of F_v follows from an easy modification of Example 11.4 in [4]. □

The upper bound in Theorem 5.2 is now an immediate consequence of Theorem 5.1, since the Hausdorff dimension of a set is never greater than its box-counting dimension. For the lower bound, we use the "mass distribution principle", which says that if μ is a finite measure on a set F and c and ε are positive numbers such that

$$\mu(U) \le c|U|^s$$

for all sets U with $|U| \le \varepsilon$, then $\dim_H F \ge s$. (See Falconer [4, Ch. 4].)

Before constructing an appropriate measure μ_v on F_v, some additional notation is needed. For $v \in \mathcal{V}$, denote by J_v the smallest closed rectangle with edges parallel to the coordinate axes which contains F_v. Let Γ_v^n be the set of all directed paths of length n in $\mathcal{G}$ with initial vertex v, and let Γ_v^∞ be the set of all such paths having infinite length. For a path $\gamma \in \Gamma_v^\infty$, let $\gamma|_n$ denote the finite subpath of γ consisting of the first n edges of γ. For a path $\gamma = (e_1, \dots, e_n)$ in Γ_v^n, let $\tau(\gamma) \in \mathcal{V}$ denote the terminal vertex of γ, let T_γ denote the composition $T_{e_1} \circ \cdots \circ T_{e_n}$, and let $R(\gamma) := T_\gamma(J_{\tau(\gamma)})$. Observe that

$$F_v \subset \bigcup \{R(\gamma) : \gamma \in \Gamma_v^n\}, \qquad n = 1, 2, \dots, \quad v \in \mathcal{V}.$$

Finally, for a rectangle R with edges parallel to the coordinate axes, let $\ell(R)$ denote the height of R.

We are now ready to construct the measures μ_v. Fix $v \in \mathcal{V}$. Define

$$p(e) = r(e)^2, \qquad e \in \mathcal{E}, \tag{9}$$

and extend p to Γ_v^n by putting $p(\gamma) = p(e_1) \cdots p(e_n)$ if $\gamma = (e_1, \dots, e_n)$. From the labeling shown in Figure 12 it is clear that p defines a probability measure on Γ_v^n. By Kolmogorov's consistency theorem, there is a unique probability measure $\tilde{\mu}_v$ on the Borel sets of Γ_v^∞ whose restriction to the cylinder sets $\{\gamma \in \Gamma_v^\infty : \gamma|_n = (e_1, \dots, e_n)\}$ agrees with p.

Define the mapping $t : \Gamma_v^\infty \to [0,1]$ by $t(\gamma) = (0.t_1 t_2 t_3 \dots)_2$, where $t_j = 0$ if $r(e_j) = \sqrt{a}$, and $t_j = 1$ if $r(e_j) = \sqrt{1-a}$, $\gamma = (e_1, e_2, \dots)$. In other words, $t(\gamma)$ is the number in $[0,1]$ whose jth binary digit is 0 if the jth edge of γ is labeled $\sqrt{a}$, and 1 otherwise. Define mappings $\pi^{\Delta} : \Gamma_v^{\infty} \to \Delta$ and $\pi_v : \Gamma_v^{\infty} \to F_v$ by

$$\pi^{\Delta}(\gamma) := \Pi(t(\gamma)), \qquad \pi_v(\gamma) := (t(\gamma), f_v(t(\gamma))).$$

Through these mappings, $\tilde{\mu}_v$ induces probability measures μ^{Δ} on Δ and μ_v on F_v defined by

$$\mu^{\Delta} := \tilde{\mu}_v \circ (\pi^{\Delta})^{-1}, \qquad \mu_v := \tilde{\mu}_v \circ \pi_v^{-1}.$$

For convenience, we consider μ_v to be defined on all of $\mathbf{R}^2$. Let $\hat{\mu}_v$ be the projection of μ_v onto the vertical axis. That is, $\hat{\mu}_v(A) := \mu_v([0,1] \times A)$, $A \subset \mathbf{R}$. Note that alternatively, $\hat{\mu}_v$ can be thought of as the projection of μ^{Δ} onto the appropriate coordinate axis in Figure 8.

Key observation: μ^{Δ} is the uniform distribution on Δ. Hence, $\hat{\mu}_v$ is absolutely continuous with respect to Lebesgue measure on $\mathbf{R}$.

This follows since the subtriangles Δ_0 and Δ_1 have areas in a ratio of $a : (1-a) = p(e) : p(e')$, where e and e' are any two edges in $\mathcal{G}$ emanating from the same vertex and labeled $\sqrt{a}$ and $\sqrt{1-a}$, respectively. Hence picking an edge at random in $\mathcal{G}$ according to the probability distribution p corresponds to picking a subtriangle at random with probability proportional to its area.

Proof of the lower bound in Theorem 5.2. If $\gamma = (e_1, e_2, \dots) \in \Gamma_v^{\infty}$, the contraction ratios

$$r(e_1), r(e_2), \dots$$

are independent, identically distributed random variables with respect to $\tilde{\mu}_v$. Thus, the Strong Law of Large Numbers implies that

$$(r(e_1) \cdots r(e_n))^{1/n} \to L := (\sqrt{a})^a (\sqrt{1-a})^{1-a}, \qquad \tilde{\mu}_v - a.e., \tag{10}$$

so

$$[\ell(R(\gamma|_n))]^{1/n} \to L, \qquad \tilde{\mu}_v - a.e. \tag{11}$$

Fix $\varepsilon > 0$. By (11), there exist a positive integer N and a subset $\tilde{F}$ of Γ_v^∞ of positive $\tilde{\mu}_v$-measure such that $\gamma \in \tilde{F}$ implies

$$(1-\varepsilon)^n L^n < \ell(R(\gamma|_n)) < (1+\varepsilon)^n L^n, \qquad n \geq N. \tag{12}$$

Put $F := \pi_v(\tilde{F})$.

Now choose $\delta > 0$ such that $\delta < (1/2)^N$, and imagine a covering of F by squares of sides less than or equal to δ. Let U be such a square. Without loss of generality, we may assume that the side of U is exactly $(1/2)^n$ for some $n \geq N$, and that U lies completely inside some rectangle $R := R(\gamma|_n)$, where $\gamma \in \tilde{F}$. Write $\tau = \tau(\gamma|_n)$, and $T = T_{\gamma|_n}$. Note the factorization

$$\mu_v(U) = \mu_v(R)\mu_\tau(T^{-1}(U)). \tag{13}$$

The first factor in (13) is estimated by

$$\mu_v(R) = p(\gamma|_n) \leq const \cdot [\ell(R)]^2,$$

in view of (9). Since $T^{-1}(U) = [0,1] \times [\alpha, \beta]$ for certain numbers α and β with $(\beta - \alpha)/\ell(U) = \ell(J_\tau)/\ell(R)$, the second factor may be estimated by

$$\mu_\tau(T^{-1}(U)) = \hat{\mu}_\tau([\alpha,\beta]) \leq const \cdot \frac{\ell(U)}{\ell(R)},$$

using the absolute continuity of $\hat{\mu}_\tau$.

Combining these estimates and using (12) yields

$$\mu_v(U) \leq const \cdot (1+\varepsilon)^n L^n \ell(U) = const \cdot |U|^{s(\varepsilon)},$$

where

$$s(\varepsilon) = 1 + \frac{\log(1+\varepsilon) + \log L}{-\log 2}.$$

It follows by the mass distribution principle that

$$\dim_H F_v \geq \dim_H F \geq s(\varepsilon).$$

Letting $\varepsilon \searrow 0$ and recalling the definition of L from (10) completes the proof. □

Remark 5.3 The difficulty in obtaining the exact dimension of the F_v appears to lie in the fact that for fixed n, the rectangles $R(\gamma)$ $(\gamma \in \Gamma_v^n)$ are of varying sizes, and do not line up neatly in horizontal strips. This provides a sharp contrast with the settings of McMullen, Kono, and Takahashi described in sections 3 and 4.

Observe that the proof of the lower bound depends in a a crucial way on the absolute continuity of the measures $\hat{\mu}_v$. In order to try to obtain a better bound, one might want to replace the probability measure p defined by (9) with a measure that makes $p(e)$ proportional to $r(e)$, and define measures $\tilde{\mu}_v$, μ_v and $\hat{\mu}_v$ correspondingly. If one could prove that the projection $\hat{\mu}_v$ has dimension 1, it would follow as in Bedford and Urbanski [1] that $\dim_H F_v = \dim_B F_v$. However, checking whether $\hat{\mu}_v$ has dimension 1 appears to be hard, as $\hat{\mu}_v$ cannot be expected to be absolutely continuous. Other measures can of course be tried, but one keeps running into the problem, caused by the irregular arrangement of the rectangles, of finding the dimension of $\hat{\mu}_v$.

Acknowledgements

We express our gratitude to Prof. Peter Lax and Prof. Dan Mauldin for directing our interest in this subject.

References

[1] T. Bedford and M. Urbanski, 'The box and Hausdorff dimension of self-affine sets', *Ergodic Theory Dynam. Systems* (4) 10 (1990) 627-644.

[2] R. T. Bumby, 'The differentiability of Pólya's Function', *Adv. Math.* 18 (1975) 243-244.

[3] K. J. Falconer, 'The Hausdorff dimension of self-affine fractals', *Math. Proc. Camb. Phil. Soc.* (2) 103 (1988) 339–350.

[4] K. J. Falconer, *Fractal Geometry. Mathematical Foundations and Applications* (Wiley, 1990).

[5] K. J. Falconer, 'The dimension of self-affine fractals II', *Math. Proc. Camb. Phil. Soc.* (1) 111 (1992) 169–179.

[6] D. Hilbert, 'Über die stetige Abbildung einer Linie auf ein Flächenstück', *Math. Annln.* 38 (1891) 459-460.

[7] W. Hurewicz and H. Wallman, *Dimension Theory* (Princeton Univ. Press, 1948).

[8] J. E. Hutchinson, 'Fractals and self-similarity', *Indiana Univ. Math. J.* (5) 30 (1981) 713–747.

[9] T. Kamae, 'A characterization of self-affine functions', *Japan J. Appl. Math.* 3 (1986) 271-280.

[10] T. Kamae and S. Takahashi, *Ergodic Theory and Fractals* (in Japanese) (Springer-Verlag, Tokyo, 1993).

[11] R. Kenyon and Y. Peres, 'Hausdorff dimensions of sofic affine-invariant sets', *Israel J. Math.* 94 (1996) 157-178.

[12] N. Kono, 'On self-affine functions', *Japan J. Appl. Math.* 3 (1986) 259-269.

[13] N. Kono, 'On self-affine functions II', *Japan J. Appl. Math.* 5 (1988) 441-454.

[14] S. Lalley and D. Gatzouras, 'Hausdorff and box dimensions of certain self-affine fractals', *Indiana Univ. Math. J.* 41 (1992) 533–568.

[15] P. D. Lax, 'The differentiability of Pólya's function', *Adv. Math.* 10 (1973) 456-464.

[16] H. Lebesgue, *Lecons sur l'integration et la recherche des fonctions primitives.* pp. 44-45 (Gauthier-Villars, Paris, 1904).

[17] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).

[18] M. McClure, 'The Hausdorff dimension of Hilbert's coordinate functions', *Real Anal. Exchange* 24 (1998/9) 875–884.

[19] C. McMullen, 'The Hausdorff dimension of general Sierpiński carpets', *Nagoya Math. J.* 96 (1984) 1-9.

[20] E. H. Moore, 'On certain crinkly curves', *Trans. Amer. Math. Soc.* 1 (1900) 72-90.

[21] G. Peano, 'Sur une courbe qui remplit toute une aire plane', *Math. Annln.* 36 (1890) 157-160.

[22] G. Pólya, 'Über eine Peanosche Kurve', *Bull. Acad. Sci. Cracovie* A (1913) 305-313.

[23] K. Prachar and H. Sagan, 'On the differentiability of the coordinate functions of Pólya's space-filling curve', *Monatsh. Math.* 121 (1996) 125-138.

[24] H. Sagan, *Space-Filling Curves* (Springer-Verlag, New York, 1994).

[25] W. Sierpiński, 'Sur une nouvelle courbe continue qui remplit toute une aire plane', *Bull. Acad. Sci. Cracovie* A (1912) 462-478.

[26] S. Takahashi, 'Dimension spectra of self-affine sets', *Israel J. Math.* 127 (2002) 1-18.

[27] M. Urbanski, 'The Hausdorff dimension of the graphs of continuous self-affine functions', *Proc. Amer. Math. Soc.* 108 (1990) 921-930.