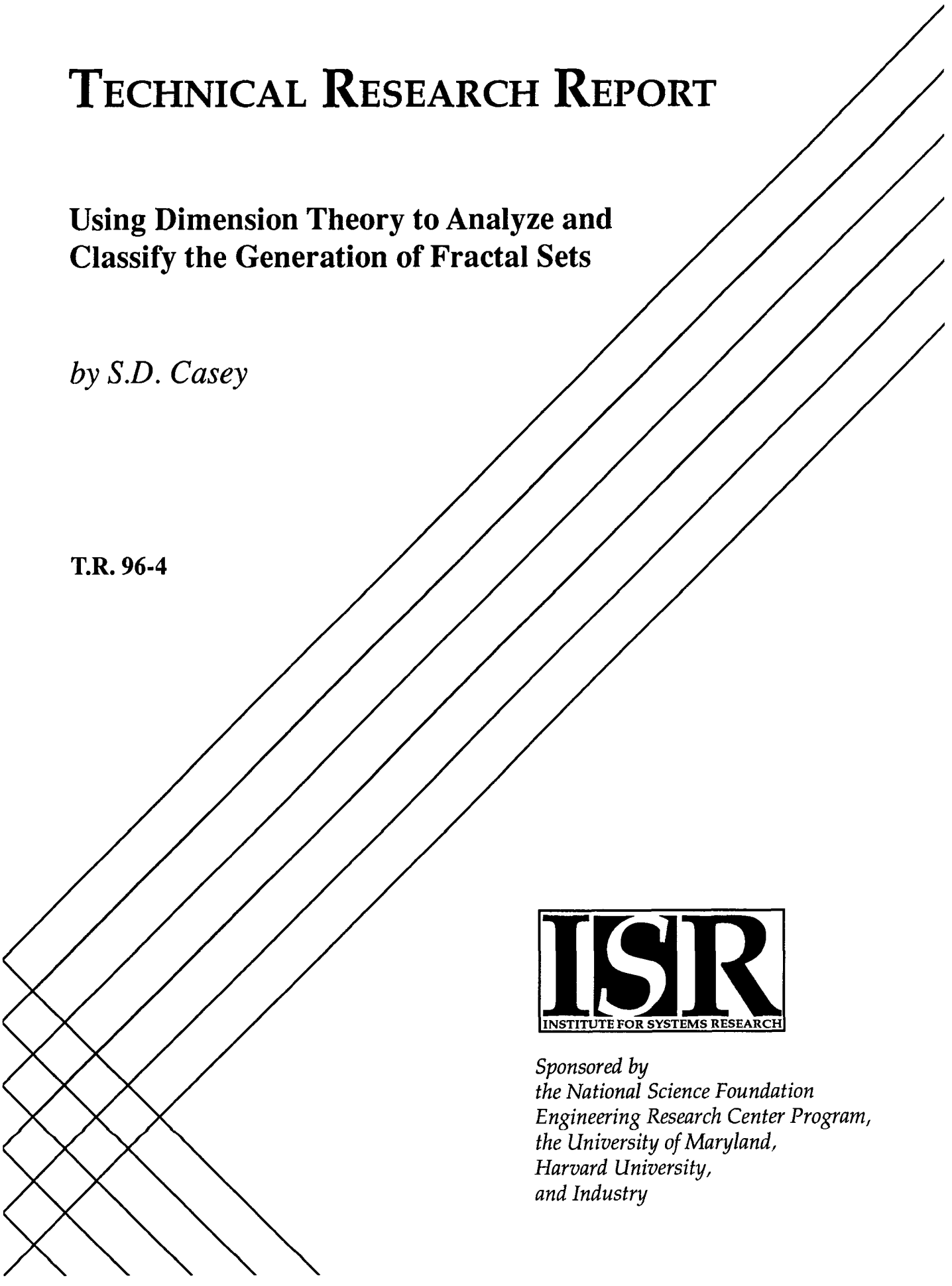
TECHNICAL RESEARCH REPORT

Using Dimension Theory to Analyze and Classify the Generation of Fractal Sets

by S.D. Casey

T.R. 96-4



Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

Using Dimension Theory to Analyze and Classify the Generation of Fractal Sets

Stephen D. Casey*†
Institute for Systems Research
University of Maryland
College Park, MD 20742–3311
e-mail: scasey@isr.umd.edu

January 21, 1996

Abstract

This article discusses the interplay in fractal geometry occurring between computer programs for developing (approximations of) fractal sets and the underlying dimension theory. The computer is ideally suited to implement the recursive algorithms needed to create these sets, thus giving us a laboratory for studying fractals and their corresponding dimensions. Moreover, this interaction between theory and procedure goes both ways. Dimension theory can be used to classify and understand fractal sets. This allows us, given a fixed generating pattern, to describe the resultant images produced by various programs. Thus, dimension theory can be used as a tool which enables us to predict and classify behavior of certain fractal generating algorithms.

We also tie these two perspectives in with some of the history of the subject. Four examples of fractal sets developed around the turn of the century are introduced and studied from both classical and modern viewpoints. Then, definitions and sample calculations of fractal and Hausdorff–Besicovitch dimension are given. We discuss various methods for extracting dimension from given fractal sets. Finally, dimension theory is used to classify images.

*On sabbatical leave from the Department of Mathematics and Statistics, The American University.
†Research partially supported by AFOSR Grant F49620–94–1–0196.

1 Introduction

Fractal geometry lies at an intersection of nature, art and many areas of science – mathematics, physics, computer science, electrical and mechanical engineering, chemistry, economics, etc. (see Pickover [15]). The theory has evolved into an extension (but by no means a replacement) of our current models of the world, which are based on Euclidean geometry and smooth generalizations of that geometry, e.g., differentiable manifolds. It includes highly detailed objects produced by repeated application of maps, while Euclidean geometry includes objects made by the conglomeration of Euclidean shapes – circles, lines, polygons, etc. – each of which can be generated by simple maps. Both visual inspection and statistical analysis[1] lead us to believe that much of the natural world (e.g., clouds, mountains, coastlines, trees (see Mandelbrot [12])) and many man–made objects (e.g., spatial distributions of communication networks, temporal flow of messages in these networks, economic distribution of resources (see Casey[3] and Mandelbrot [12])) can be modeled by this theory.

This article discusses a very interesting interplay in the theory of fractal geometry, occurring between computer programs for developing (approximations of) fractal sets and the underlying dimension theory. The computer is ideally suited to implement the recursive algorithms needed to create these sets, thus giving us a laboratory for studying fractals and their corresponding dimensions. In many articles, the flow of ideas is from procedure to theory. Images are generated, then analyzed. However, this interaction between theory and procedure goes both ways. In this paper we discuss how dimension theory can be used to classify and understand fractal sets, enabling us to understand the outcome of fractal generating procedures. We can, for example, start with a fixed generating pattern and predict and classify the resultant image, understanding its properties directly from its generator.

We will also tie these two perspectives in with some of the history of the subject. Four examples of fractal sets developed around the turn of the century are introduced and studied from

[1]Robust methods based on Pareto–Levy distributions can be used to measure data that exhibits a fractal structure (see [3]).

both classical and modern viewpoints. We include definitions and sample calculations of fractal and Hausdorff–Besicovitch dimension, and discuss various methods for extracting dimension from given fractal sets.

1.1 The *Fracmkr* Algorithm

Figures for the article were generated by a recursive algorithm (*Fracmkr*) developed by the author and Nicholas Reingold [4]. This is an efficient algorithm which produces (approximations of) self–similar fractal sets by a repeated scaling, translation, reflection, and/or rotation of a fixed pattern or set of patterns. The resultant sets strictly preserve fractal scaling. The algorithm complements Mandelbrot's *The Fractal Geometry of Nature*, in that it can be developed into a program which reproduces the self–similar fractals in Mandelbrot (approximately 45% of the figures in the book).

Fracmkr is a "pattern rewriting system" in which a given geometric pattern is drawn repeatedly after suitable scaling and placement. The system uses two patterns – an initial configuration, or base, and a generating pattern, or seed. Generating schemes are patterns with some built–in information on orientation and connection in later levels of iteration. This is exactly how one would begin to study self–similar patterns with a paper and pencil. In this system, if you can draw seed and base patterns on a piece of paper, identifying vertices and orientation schemes, then you can produce iterates of the seed on the base in the computer. The only overhead between levels of recursion is a single boolean variable. Consequently, the program used to generate the patterns is extremely portable. Figures for the article can be generated on any PC. See [4] for a more complete description of the procedure and a listing of the code.

1.2 Some Definitions

We now must establish some terminology. Let $\mathbf{R}^n$ denote n–dimensional Euclidean space.

We say a mapping $f : \mathbf{R}^n \to \mathbf{R}^m$ is a *contraction* if there exists an α, $0 \leq \alpha < 1$ such that

$$\|f(x) - f(y)\| \leq \alpha \|x - y\| \tag{1}$$

for all x, y in the domain of f. We say that a set S is *invariant* under a collection of contraction mappings, say $\{m_1, m_2, \ldots, m_n\}$, if

$$S = \bigcup_{i=1}^{n} m_i S \,. \tag{2}$$

Using these last two definitions, we define a *self–similar* set as a set which is invariant under a collection of contraction maps such that $m_i S \cap m_j S$ is at most a countable number of points when $i \neq j$ (see Hutchinson[2] [11]).

2 Fractal Geometry: Four Classical Examples

The theory of fractal geometry goes back at least to K. Weierstrass, who constructed a fractal set in his development of a continuous but nowhere differentiable curve[3] in 1872 (see Edgar [5]). The first fractal sets were developed as examples and counterexamples necessary in the natural development of mathematics, e.g., a totally disconnected perfect set, a curve of infinite length without a tangent contained in a finite region, or a space–filling curve. From our modern perspective, we see that they all exhibit both the scale invariance and intricate detail which group them together as fractals. It is important to note the years during which these mathematical constructs appeared. Although we may think of fractals as modern objects generated by computers, the mathematical foundations of fractal geometry are over one hundred years old (see Mandelbrot [12]). Moreover, our fascination with fractals began much earlier (see Pickover [15]).

The four examples below are important in the development of the theory in analysis, topology, and geometry (see Falconer [6, 7], Federer [8], Hocking and Young [9], Huriewicz and Wallman [10], and Munkres [13]). All are currently being used in the mathematical modeling of natural and man–made phenomena (see Casey [3], Mandelbrot [12], and West [17]). The sets are all self–similar. We include the collection of contraction mappings associated with each set in the description, and remark that the fractal sets may be constructed directly from the mappings by using M. Barnsley's iterated function systems (see Barnsley [1]).

[2]More sophisticated overlap conditions are possible (see [11]). This will be sufficient for our purposes.
[3]A curve with similar properties was developed by B. Bolzano in 1830. He did not publish his result.

2.1 The Cantor Set

G. Cantor (1884)

The ancient Greeks felt that in order to be indefinitely subdivisible a body had to be a continuum. In 1884, Cantor produced a set showing that this was not the case. This set, the Middle Thirds Set C, is constructed as follows. Given $[0,1]$, remove the open middle third segment $(\frac{1}{3}, \frac{2}{3})$. Then, remove the open middle thirds of the remaining two segments, i.e., remove $(\frac{1}{9}, \frac{2}{9})$ from $[0, \frac{1}{3}]$ and $(\frac{7}{9}, \frac{8}{9})$ from $[\frac{2}{3}, 1]$. Repeat. At the n^{th} stage of this process, there will be 2^n segments, each of length $(1/3)^n$. Continue ad infinitum.

The total length of the removed set is $(1/3)\sum_{i=0}^{\infty}(2/3)^n$, which equals one. However, the remaining set C is a perfect, and thus an uncountable set which is totally disconnected[4] (see [9]).

The Cantor set can be generated by the maps

$$m_1(x) = \frac{1}{3}x \ , \ m_2(x) = \frac{1}{3}x + \frac{2}{3} \ .$$

2.2 Space–Filling Curves

G. Peano (1890)

Peano developed a sequence of continuous curves, $\{P_n\}$, defined on $[0,1]$ which converge to a continuous curve P that hits every point in the square $[0,1] \times [0,1]$, i.e., P is space–filling.[5]

The curve P_1 is shown in Figure 1. To produce P_2, divide $[0,1] \times [0,1]$ into four equal subsquares, ordering them clockwise from the lower left hand corner. Then place four copies of the P_1 curve, scaled by $1/2$, and rotated by $-\pi/2$, 0, 0, and $\pi/2$, respectively, into the four subsquares. The curve P_{n+1} is produced similarly, using four scaled and rotated copies of P_n.

[4]Infinitely many other Cantor constructions are possible, e.g., removing the second and fourth sections of an interval divided into equal fifths.

[5]Technically, Peano curves are not fractal. However, they are constructed in the same fashion as self–similar fractals, and have the property that the limit curve P has a dimension which is an integer value greater than any curve P_n in the construction. (We can construct volume–filling curves, etc.) Mandelbrot [12] devotes chapter 7 of his book to a discussion of space filling curves.

The P_{n+1} curve is no further away than $(1/\sqrt{2})(1/2)^{n+1}$ from any point in $[0,1] \times [0,1]$ (also see [13], pp. 271–74). Various approximations of Peano curves can be produced using the algorithm.

Let $\vec{x} = [x\ y]^T$ be a vector in the Cartesian plane $\mathbf{R}^2$. Let $\theta \in [0, 2\pi)$ be an angle, and let

$$R_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

be the rotation matrix for angle θ. Then, the contraction mappings under which P is invariant are

$$m_1(\vec{x}) = \tfrac{1}{2}R_{(-\pi/2)}\vec{x} + [0\ \tfrac{1}{2}]^T\ , \quad m_2(\vec{x}) = \tfrac{1}{2}\vec{x} + [0\ \tfrac{1}{2}]^T\ ,$$
$$m_3(\vec{x}) = \tfrac{1}{2}\vec{x} + [\tfrac{1}{2}\ \tfrac{1}{2}]^T\ , \quad m_4(\vec{x}) = \tfrac{1}{2}R_{(\pi/2)}\vec{x} + [1\ 0]^T\ .$$

2.3 Snowflake Curves

H. von Koch (1904)

Koch's snowflake curve is an example of a Jordan curve (a closed, non–intersecting loop) of infinite length fitting inside a bounded region. The limit curve also has a tangent nowhere.

The snowflake curve[6] consists of three Koch arcs K joined at the vertices of an equilateral triangle. The Koch arc construction is to replace the segment $[0,1]$ with a seed pattern consisting of four segments of length $1/3$ with endpoints

$$(0,0),\ (\frac{1}{3},0),\ (\frac{1}{2},\frac{\sqrt{3}}{6}),\ (\frac{2}{3},0),\ (1,0).$$

Then, each of the four segments is replaced by the seed pattern, and so on. At the n^{th} stage the curve consists of 4^n segments of length $(1/3)^n$. As n goes to infinity, the total length also goes to infinity.

The contraction mappings under which K is invariant are

$$m_1(\vec{x}) = \tfrac{1}{3}\vec{x}\ , \quad m_2(\vec{x}) = \tfrac{1}{3}R_{(\pi/3)}\vec{x} + [0\ \tfrac{1}{3}]^T\ ,$$
$$m_3(\vec{x}) = \tfrac{1}{3}R_{(-\pi/3)}\vec{x} + [\tfrac{1}{2}\ \tfrac{\sqrt{3}}{6}]^T\ , \quad m_4(\vec{x}) = \tfrac{1}{3}\vec{x} + [\tfrac{2}{3}\ 0]^T\ .$$

[6]The snowflake curve is not self–similar. However, it consists of three self–similar pieces.

2.4 Sierpinski Gasket

W. Sierpinski (1915)

Starting with a filled–in equilateral triangle, perform a Cantor–like removal process by removing the interior of the middle equilateral triangle whose vertices are the midpoints of the three edges. Repeat this process on the three remaining filled–in triangles, and then the nine remaining ones, and so on. All of the points in the remaining set are vertices of some removed triangle, and thus are branch, or ramification, points. Figure 2 gives an approximation of the gasket by a non–intersecting arc.

This construction is due to Sierpinski. The set S_g is invariant under

$$m_1(\vec{x}) = \frac{1}{2}\vec{x}\ ,\ m_2(\vec{x}) = \frac{1}{2}\vec{x} + [\frac{1}{2}\ 0]^T\ ,\ m_3(\vec{x}) = \frac{1}{2}\vec{x} + [\frac{1}{4}\ \frac{\sqrt{3}}{4}]^T\ .$$

These four examples provide a good framework in which to introduce the definition of a fractal, given in the next section. In order to do this, we need to discuss dimension theory. Understanding even the definitions in this theory requires some technical background, which can be found in [9] and [13].

3 Fractals, Scaling, and Dimension

We can "define" a *fractal* as a set which :

(i.) *has detailed fine structure.* Fractals have the property that all magnifications of some or all of the set reveals intricate detail. This is unlike the differentiable curves and surfaces we encounter in calculus. Repeated enlargements of every differentiable curve or surface reveal a line or a plane, respectively.

(ii.) *has scale invariance.* All magnifications of the fractal reveal a set which is exactly the same (self–similarity) or statistically or asymptotically the same (statistical or quasi– self–similarity). This is related to the first property. However, there are self–similar sets with no detailed fine structure, e.g. a line.

(iii.) *is produced recursively.* We can generate these sets by repeated application of some collection of maps. These maps provide a simple encoding of these complicated sets.

(iv.) *can not be described easily in terms of Euclidean geometry.* Although the sets can be constructed from the objects studied under classical geometry, the structure of the limit set is an uncountable collection of points in a complicated arrangement.

(v.) *you know is a fractal when you see it.* This may be the only definition upon which everyone can agree.

The mathematical definition is more complicated and very technical.

Definition 3.1 *A* fractal *is a set whose Hausdorff–Besicovitch dimension strictly exceeds its topological dimension* (B. Mandelbrot (1975) [12], pg. 15).

3.1 Dimension Theory

In order to understand the definition of a fractal, we first need to calculate the *dimension* of a given set. Dimension is the primary tool used in classifying fractal sets. The topological and Hausdorff–Besicovitch dimensions (denoted D_T and D_{HB}, respectively) are very technical. Topological dimension can be thought of as an integer n which describes which Euclidean n–space $\mathbf{R}^n$ that a given set resembles locally. This dimension is invariant under homeomorphisms[7], and therefore is an accurate determination of topological information about the set. It does not, however, measure detailed geometry and scaling structure. For example, every Cantor set, any countable set, and a set consisting of a single point all have topological dimension 0. All non–intersecting curves have $D_T = 1$. Proof of these facts are difficult, and involve machinery from algebraic topology. Many of the fundamental results of topological dimension are developed in Huriewicz and Wallman's *Dimension Theory* [10]. Also see [4], [6], and [7].

Fortunately, there is a more accessible and simplified version of dimension — the *fractal* or *Kolmogorov–Mandelbrot dimension* D_{KM}. It was introduced by L. Pontrjagin (1932), and

[7] A homeomorphism is a one–to–one onto continuous mapping with a continuous inverse.

studied by A. Kolmogorov (1950's), L. Richardson (1960's), and Mandelbrot (1970's), among others (see [12] for more information). The fractal dimension of a set is a key ingredient in identifying a fractal set. Mandelbrot [12] assigns a D_{KM} value with most of the images in his book. For self–similar and nearly self–similar sets with an intricate geometry, fractal dimension can not only measure the structure of a set, but can also be determined experimentally (see [1], [3], [12], and section 3.2). Assuming this definition, we define a *fractal* as a set whose fractal dimension strictly exceeds its topological dimension. In most cases, the fractal dimension of a fractal is a non–integer number. We shall see that the definition of a fractal given above reflects properties (i.), (ii.), and (iv.) in the list above. A theorem of M. Barnsley and S. Demko gives that all fractals can be approximated arbitrarily closely by repeated applications of a set of maps (see [1], chapter 3).

Eight examples will be discussed. The first three are a point, the unit interval $I = [0, 1]$, and the unit square $I^2 = [0, 1] \times [0, 1]$. The next four were developed in the previous section – C, P, K, and S_g. The final example is the set $H = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\} \cup \{0\}$.

Consider any closed and bounded set in n–dimensional Euclidean space, $\mathbf{R}^n$. Let r be any positive number, and let $\mathcal{N}(r)$ be the *minimal* number of closed line segments, balls, or spheres of radius r needed to cover the set. Then we want to calculate a number D such that as $r \to 0$,

$$\mathcal{N}(r) \cdot r^D \sim 1 \quad \text{(Scaling Relationship)}. \tag{3}$$

Consider first a single point. For any $r > 0$, we need only one line segment to cover. Thus, $\mathcal{N}(r) \cdot r^D = 1 \cdot r^D = 1$ for $D_{KM} = 0$.

Next consider $[0, 1]$. Let $\lceil x \rceil$ be the smallest integer greater than or equal to any real number x. Given any $r > 0$, we need $\lceil 1/r \rceil$ line segments of length r to cover $[0, 1]$. Thus $\mathcal{N}(r) \cdot r^D = \lceil 1/r \rceil \cdot r^D \sim 1$, for $D_{KM} = 1$.

The set $[0, 1] \times [0, 1]$ can be covered by $\lceil 1/r \rceil^2$ balls of radius r, so in this case D_{KM} is 2.

In these first three examples, D_{KM} agrees with the topological dimension. The next example breaks from this notion.

Consider the Cantor Middle Thirds set constructed in 2.1. At the n^{th} stage of its con-

struction, we need 2^n line segments of radius $(1/3)^n$ to cover. As this is true for r restricted to $\{(1/3)^n\}$ and for all n, $n = 1, 2, \ldots$,

$$\mathcal{N}(r) \cdot r^D = 2^n (1/3)^{nD}.$$

Therefore,

$$\begin{aligned} & \mathcal{N}(r) \cdot r^D = 1 \\ \Leftrightarrow \quad & \log(2^n (1/3)^{nD}) = \log 1 \\ \Leftrightarrow \quad & n(\log 2 - D \log 3) = 0. \end{aligned}$$

(where log denotes the natural logarithm). Thus, for r restricted to $\{(1/3)^n\}$, $n = 1, 2, \ldots$, $D_{KM}(C) = \log 2 / \log 3$.

These examples motivate the following:

Definition 3.2 The fractal, or Kolmogorov–Mandelbrot, dimension $D_{KM}(X)$ *of a set* X *is:*

$$D_{KM}(X) = \lim_{r \to 0} \frac{\log(\mathcal{N}(r))}{\log(1/r)}. \tag{4}$$

The number D_{KM} is the value necessary to preserve the scaling relationship $\mathcal{N}(r) \cdot r^D \sim 1$ as $r \to 0$. Calculation of D_{KM} is simplified by the fact that the continuous variable r may be replaced by the discrete variable $r_n = \rho^n$, $0 < \rho < 1$, $n = 1, 2, \ldots$. That is, if

$$\Delta = \lim_{n \to \infty} \frac{\log(\mathcal{N}(r_n))}{\log(1/r_n)}, \tag{5}$$

then $\Delta = D_{KM}$ (see [1], pg. 176). This allows us to derive the fractal dimension of a self–similar fractal set directly from its generating seed. Let's look at the dimension of some of the sets from the previous section.

First we see that for the Cantor Middle thirds set, $D_{KM}(C) = \log 2 / \log 3$. Next, consider the Koch arc K described in section 2.3. At the n^{th} stage it consists of 4^n segments of length $(1/3)^n$. Therefore, for $r = r_n = (1/3)^n$,

$$\begin{aligned} & \mathcal{N}(r) \cdot r^D = 1 \\ \Leftrightarrow \quad & \log(4^n (1/3)^{nD}) = \log 1 \\ \Leftrightarrow \quad & n(\log 4 - D \log 3) = 0. \end{aligned}$$

Thus,

$$D_{KM}(K) = \frac{\log 4}{\log 3}.$$

At the n^{th} stage of construction, the Sierpinski gasket (2.4) consists of 3^n equilateral triangles with sides of length $(1/2)^n$. Therefore, calculating as above, we see that

$$D_{KM}(S_g) = \frac{\log 3}{\log 2}.$$

Finally, consider the construction of the Peano curve given in section 2.2. At the n^{th} stage it consists of $2 \cdot 4^n$ line segments of length $(\sqrt{2}/2)(1/2)^n$. So, for $r_n = (\sqrt{2}/2)(1/2)^n$,

$$D_{KM}(P) = \lim_{n\to\infty} \frac{\log(2 \cdot 4^n)}{\log\left(\frac{1}{(\sqrt{2}/2)(1/2)^n}\right)} = \lim_{n\to\infty} \frac{\log(2 \cdot 4^n)}{(1/2)\log(2 \cdot 4^n)} = 2.$$

We also introduce the Hausdorff–Besicovitch dimension. This dimension is an essential tool of geometric measure theory. Since it requires sophisticated mathematical tools, we give only the definition and the dimensions of the examples, referring the interested reader to Falconer [6, 7], Federer [8], and Taylor [16] for details. These references also include detailed discussions of other related dimensions.

Definition 3.3 The Hausdorff–Besicovitch Dimension $D_{HB}(X)$ *of a set* X*: Let* X *be a subset of* $\mathbf{R}^n$*, and consider the set of all countable covers of* X *by sets of diameter less than* $r > 0$*. Let* r_n *be the diameter of the* n^{th} *set in any given cover. Then, for* $\beta > 0$*, the* β^{th} Hausdorff–Besicovitch outer measure of X *is*

$$\mu^*_\beta(X) = \lim_{r\to 0}\left[\inf_{\{\text{covers}\}}\left[\sum_{n=1}^{\infty} r_n^\beta\right]\right]. \tag{6}$$

The Hausdorff–Besicovitch dimension is then

$$D_{HB}(X) = \inf\left\{\beta : \mu^*_\beta(X) = 0\right\}. \tag{7}$$

(F. Hausdorff (1919), A. Besicovitch (1928) [6], pp. 7–10)

Equivalently, we could define

$$D_{HB}(X) = \sup\left\{\beta : \mu^*_\beta(X) = \infty\right\}. \tag{8}$$

Clearly, the difference between D_{HB} and D_{KM} is the choice of the covering sets. The flexibility of choosing arbitrary sets of diameter less than r, as opposed to line segments, disks, or balls of radius r, makes the HB dimension a much finer measurement, albeit more difficult to calculate. However, for self–similar sets, we have the following useful theorem.

Theorem 3.1 (Hutchinson [11]) *If X is self–similar, then*

$$D_{KM}(X) = D_{HB}(X)\,. \tag{9}$$

Therefore, we see that for the examples above, we have that $D_{HB}(I) = 1$, $D_{HB}(I^2) = 2$, $D_{HB}(C) = \frac{\log 2}{\log 3}$, $D_{HB}(K) = \frac{\log 4}{\log 3}$, $D_{HB}(S_g) = \frac{\log 3}{\log 2}$, and $D_{HB}(P) = 2$. Also, as self–similar sets strictly preserve the scaling relationship given in the previous section, D_{HB} will reflect the scale invariant structure for these sets. It is also sufficiently robust to show this structure for statistically and quasi– self–similar sets.

In general,

$$D_{HB}(X) \leq D_{KM}(X)\,.$$

The set H is an example of a set for which $D_{HB}(X) < D_{KM}(X)$. In fact, $D_{HB}(H) = 0$, while $D_{KM}(H) = 1/2$ (see [7], pg. 45)[8]. Note that in this case, as H has only one "scaling point" (namely the point $\{0\}$), H is not a set which would be considered as a fractal. Thus, we see that the finer measurements used in calculating D_{HB} are needed to define a fractal.

3.2 Various Methods for Determining the Fractal Dimension

Sets that exhibit a scaling relationship can have this structure measured in various ways. The first and most obvious method is to extract this measure from a graphical representation of the set. This is most easily demonstrated by considering subsets of Euclidean n–space $\mathbf{R}^n$.

Let I be as in 3.1. Let X be the set we wish to measure. If has topological dimension 0 or 1, we can get the "fractal length" of X by measuring X with rulers of length $r \cdot I$. Let $\mathcal{L}(r)$ be

[8]For another example, let $T = \{1, \frac{1}{3}, \frac{1}{5}, \ldots\} \cup \{0\}$. Then, $D_{HB}(T) = 0$, while $D_{KM}(T) = 1/2$.

the minimal number of units $r \cdot I$ placed end to end needed to measure X. Then, X has "length" $\mathcal{L}(r) \cdot r$, which asymptotically gives

$$\mathcal{L} \sim \frac{1}{r^{D_{KM}-1}}, \text{ or } D_{KM} \sim \frac{\log \mathcal{L}(r)}{\log(1/r)} + 1. \tag{10}$$

Alternatively, we could calculate a "fractal box dimension." A closed circle of radius ρ can be contained in a closed square of side length 2ρ, which in turn can be contained in a closed circle of radius $\sqrt{2}\rho$. Therefore, if we let $\mathcal{N}(r)$ denote the minimal number of closed squares of side length ρ needed to cover the set, we get (by the "squeezing theorem" from calculus)

$$D_{KM} \sim \frac{\log \mathcal{N}(r)}{\log(1/r)}. \tag{11}$$

Finally, we can get a "fractal grid dimension" by using a grid of mesh size $\frac{1}{2^n}$. Let $\mathcal{G}(n)$ denote the number of elements of the grid (including boundary) that contain an element of X. Then, for $\mathcal{N}(r)$ as above, we have in $\mathbf{R}^n$, $m = 1, 2, \ldots$,

$$2^{-m}\mathcal{G}(n-1) \leq \mathcal{N}(1/2^n) \leq \mathcal{G}(k(n))$$

where $k(n)$ is the smallest integer satisfying $k \geq (n - 1 + \log_2(m))$. This inequality then in turn gives

$$D_{KM} \sim \frac{\log \mathcal{G}(n)}{\log(1/2^n)} \tag{12}$$

(see [1]).

A more general and more robust technique is to use statistical measurement. For the more complicated Hausdorff–Besicovitch measure, a related probability distribution function is the Pareto–Levy distribution, introduced by P. Levy and A. Khinchine (1936) (see West [17]). These distributions are stable, and most can be written down only in terms of their Fourier transforms. The Pareto–Levy is linked to the Hausdorff–Besicovitch dimension in that the distribution scales in the same manner as the outer measure, and, given a controlled random walk or more general process producing a family of curves or surfaces, those surfaces will have HB dimension which is a function of a parameter of the Pareto–Levy distribution.

Just as the HB dimension has its simpler counterpart in the KM dimension, so too does the Pareto–Levy, in that it reduces to the simpler Pareto distribution. This distribution was first studied by V. Pareto (1897). It has probability distribution function

$$f(x) = \begin{cases} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & x \geq x_0 \\ 0 & x < x_0 \end{cases} . \tag{13}$$

Thus, the cumulative distribution function is

$$F(x) = 1 - (\frac{x}{x_0})^{-\alpha} = 1 - \exp(-\alpha(\log(x) - \log(x_0))) . \tag{14}$$

Therefore, by mapping the domain onto log((domain)), we transform the Pareto into the exponential. Now, the exponential is "memoryless," i.e.,

$$P_e(x > a + b \,|\, x > b) = P_e(x > a) .$$

Thus, for the Pareto,

$$P_p(x > \tilde{a} \cdot \tilde{b} \,|\, x > \tilde{b}) = P_p(x > \tilde{a}) ,$$

where $\tilde{a}$, $\tilde{b}$ are of the form e^a, e^b and $\tilde{a}, \tilde{b} \geq x_0$, i.e., the Pareto is "scale invariant." Also note, the r^{th} moment of the Pareto about the origin is

$$\langle |x|^r \rangle = \frac{\alpha}{\alpha - r} ,$$

which converges for $r < \alpha$, and diverges for $\alpha \geq r$. Finally note that $f(x)$ is linear on a log–log scale. But

$$D = \frac{\log \mathcal{N}(x)}{\log(1/x)},$$

$$\log \mathcal{N}(x) = -D \log(x), \text{ and}$$

$$\mathcal{N}(x) = (x)^{-D} .$$

Therefore we have the relationship

$$\textbf{Pareto}(\alpha + 1) \longleftrightarrow \textbf{KM Dimension}(D) . \tag{15}$$

Both have a linear relation in log–log, namely

$$\begin{aligned} \log \mathcal{N}(x) &= -D \log(x) \\ \log f(x) &= -(\alpha + 1) \log(x) . \end{aligned} \tag{16}$$

The connection between the Pareto distribution and the fractal dimension has proven to be the most effective method of extracting fractal dimension (see [3]). Numerous data sets were examined. First, the log of the data was taken, and then this data was fit to an exponential by maximum likelihood estimation. The quality of this fit was then tested by χ^2 and Kolmogorov–Smirnov tests. A scaling structure was evident in many different sets, including errors in communication lines, delays in the delivery of messages, the structure of communication networks, and even European city sizes. (This last example gives more evidence to a theorem in human geography introduced by G. Zipf.) Work of Mandelbrot, Voss, and others (see [12], [14]) tells us that these methods also will extract useful measures of many varieties of noise.

We close this subsection with the following result, given in [3]. Assume that the data we have gathered is a subset of a larger data set which has a scaling structure, and that data is truncated, grouped into bins and/or consists of separate data points. Let N_B be the number of binned elements, and let B be the indexing set for these elements, except the elements in the last bin. Let N_D, D be the number of and the indexing set (again, except for the last bin) for the separate data points. Let n_j be the number if elements in each bin. Let x_j be the center of the j^{th} bin for $1 \leq j < K$. If the K^{th} bin is unbounded, let x_K be its lower endpoint. Else, let x_K be as above. We have the following.

Theorem 3.2 *Let X be a truncated, binned and/or isolated subset of a set with a scaling structure of parameter D. Then the maximum likelihood estimator for D is $\widehat{D} = \widehat{\alpha} + 1$, where*

$$\widehat{\alpha} = \frac{N_B + N_D - n_K}{\sum_{j \in B,D} n_j x_j + n_K x_K}. \tag{17}$$

The proof of this theorem, given in [3], is constructive, and provides a procedure for extracting the fractal dimension from the truncated and binned data.

3.3 Classifying General Fractal Sets Using Dimension Theory

We commented earlier that dimension is the primary tool for identifying fractals. In this subsection, we mention some of the theorems in which dimension is used to classify fractals. These

theorems are for a very general class of "reasonable" subsets of Euclidean space, namely Borel sets (see [7]). Federer's encyclopedic *Geometric Measure Theory* [8] and Falconer's *The Geometry of Fractal Sets* and *Fractal Geometry: Mathematical Foundations and Applications* [6, 7] contain many of the fundamental results in geometric measure theory, and applications of the theory to the analysis of fractals.

We begin with a well known result.

Proposition 3.1 *A set F with a countable number of elements has HB dimension 0.*

If F is uncountable and has HB dimension less than one, it has qualities in common with the Cantor set.

Theorem 3.3 ([7], pg. 30) *A set $F \subset \mathbf{R}^n$ with $D_{HB}(F) < 1$ is totally disconnected.*

Jordan curves with dimension between 1 and 2 all share a common property with the Koch curve.

Theorem 3.4 ([7], pg. 80) *If $\mathcal{C}$ is a Jordan curve in $\mathbf{R}^n$ with $1 < D_{HB}(\mathcal{C}) < 2$, then $\mathcal{C}$ contains a subcurve $\mathcal{C}_s$ such that at almost all points of $\mathcal{C}_s$, no tangent exists.*

Theorem 3.4 generalizes to higher dimensions. If we have a surface of topological dimension n, we say that it has a tangent at a point x if there exists an n–dimensional hyperplane approximating the surface at that point.

Theorem 3.5 ([8], section 3.3) *If $\mathcal{S}$ is an n–dimensional non–intersecting surface in $\mathbf{R}^m$, $n < m$, with $n < D_{HB}(\mathcal{S}) < n+1$, then $\mathcal{S}$ contains a subsurface $\mathcal{S}_s$ such that at almost all points of $\mathcal{S}_s$, no tangent exists.*

We will say that a curve is space–filling if it fills in any planar area. Combining the following with our computation above proves that P is space–filling.

Theorem 3.6 ([12], II, 7) *If $\mathcal{C}$ is a curve in $\mathbf{R}^n$ such that $D_{HB}(\mathcal{C}) = 2$, then $\mathcal{C}$ is space–filling.*

Theorem 3.6 also generalizes to higher dimensions. If have a curve $\mathcal{C}$ in $\mathbf{R}^n$ such that $D_{HB}(\mathcal{C}) = k$, where k is an integer such that $3 \leq k \leq n$, then we say that C is volume–filling.

We can create fractal sets as Cartesian products or as intersections of other fractals. Computing dimensions in these cases may be tricky. In general, we have the following.

Theorem 3.7 ([7], pg. 94) *If F and G are subsets of Euclidean space, then*

$$D_{HB}(E \times F) \geq D_{HB}(E) + D_{HB}(F)\,. \tag{18}$$

Next, assume $E, F \subset \mathbf{R}^n$, and let $F + x$ be the translation of F by the vector x.

Theorem 3.8 ([7], pg. 102)

$$D_{HB}(E \cap (F + x)) \leq \max\{0, D_{HB}(E \times F) - n\} \text{ for almost all } x \in \mathbf{R}^n\,. \tag{19}$$

It becomes increasingly difficult to classify fractals as D_{HB} increases. In fact, the subtlety of these types of results increases exponentially with the increase in dimension. General fractal sets can have small subsets exhibiting one type of behavior, other subsets exhibiting another, and so on. The dimension of the set is given by the component with the highest dimension. Thus, a curve could have $D_{HB} > 1$, and yet be differentiable along most of its length. Self–similar fractals, however, are homogeneous. Thus, the dimension of a subset equals the dimension of the set.

The theorems in this section are just a few of the results in the area of geometric measure theory. Classification of fractals using geometric measure theory is by no means complete, nor is the study of fractal sets the only topic of research in this area.

We will be able to use the results quoted in this subsection to describe self–similar sets directly from their generators.

4 Dimension and Classification of Self–Similar Fractals

The rigorous study of fractal sets involves sophisticated mathematical tools, some of which where introduced in the previous section. It was also noted that fractals can be created via recursion,

and therefore computers are an ideal tool for studying these sets. The computers can be used as tools in proving theorems about these sets, generating examples and checking calculations, but cannot prove theorems. However, the theory can be used to predict the output of a computer program, as we will now demonstrate. Our discussion in this section is limited to self–similar fractals. Recall the definition of a self–similar set given in the section 1.2. The first theorem we mention is a theorem of J. Hutchinson (Theorem 3.1). It tells us that for self–similar sets X,

$$D_{HB}(X) = D_{KM}(X)\,.$$

Thus for self–similar sets, the "proper" form of dimension can be calculated by the methods used to calculate D_{KM}. Furthermore, this number can be calculated directly from the maps used to define the set.

Theorem 4.1 (Hutchinson [11]) *Let X be a self–similar set defined by contraction mappings $\{m_i\}_{i=1}^n$. Let $\{\alpha_i\}_{i=1}^n$ be the set of contractions associated with these mappings. Then $D_{HB}(X) = D$, where D is the solution to the equation*

$$\sum_{i=1}^{n} \alpha_i^D = 1\,. \tag{20}$$

Using Hutchinson's theorems, we can calculate the HB dimension of a self–similar set in several ways. For example, given the associated contraction mappings, we can use equation (20) to compute the KM dimension. Hutchinson's first theorem then gives us that this is in fact equal to the HB dimension. The theory in section 3.3 then comes into play, classifying the set.

A variation on this also works for the resultant self–similar fractal sets. The trick is as follows. We look for pieces of the image which are scaled, rotated, and/or translated miniatures of the whole, and get the affine contraction mapping that maps the whole image onto this piece. We do this so that the entire image is made up of a collage of essentially non–overlapping miniatures of itself. This is the basic idea of Barnsley's Collage Theorem (see [1], pp. 93–110). Once we have the mappings, we again use (20).

We can also compute the dimension of a self–similar set X approximated by the pattern rewriting system (see section 1.1) directly from the generating seed. The set X we refer to is

the theoretical limit set which is produced by iterating the pattern rewriting system infinitely often. Each seed needs to be normalized by drawing it so that it begins at $(0,0)$ and ends at $(1,0)$. We also have to guarantee that the set we produce is self–similar. Our seed cannot have line segments in it of length greater than or equal to one. (If you gave the pattern rewriting system such a seed, it would produce a pattern – it just would not converge to a new self–similar pattern.[9]) We also cannot have the line segments overlap each other except at the endpoints. This can be determined by examining two generations of the pattern.

Theorem 4.2 *Let $\mathcal{S}$ be a normalized seed pattern consisting of n line segments which generates a self–similar set X. Let ℓ_i denote the length of the i^{th} line segment in $\mathcal{S}$, $i = 1, \ldots, n$. Let D be the solution to the equation*

$$\sum_{i=1}^{n} \ell_i^D = 1\,. \tag{21}$$

Then $D_{HB}(X) = D$.

The theorem follows directly from Hutchinson's second theorem (Theorem 4.1). Given a normalized seed, we can construct the contraction mappings associated with the pattern that this seed generates by mapping the unit interval $[0,1]$ onto each line segment in the seed. Each map is the composition of a contraction, a rotation, and a translation. The contraction factor α_i for the i^{th} map equals the length ℓ_i of the i^{th} line segment.

If we have a normalized seed and we have that $\sum_{i=1}^{n} \ell_i^D = 1$ has solution $D = 2$, then X will be be space–filling. Moreover, if we relax the overlap condition and have $D > 2$, iterates of the seed will overwrite each other. In fact, we can think of the number $\frac{D}{2}$ as the number of times that the iterates will overwrite each other. We discuss an example of this below. We note that the need for an overlap condition is clear from the theorem. If we allowed for overlapping segments in our defining seed, then we would be counting segments more than once in our formula. We also remark that in general we cannot solve the equation (21) in closed form unless all of the

[9]There is, of course, the case where we have $[0,1]$ as a seed. Repeated iteration of the pattern produces nothing new.

lengths of the line segments are equal. However, a numerical solution can be easily calculated, using, say, the Newton–Raphson algorithm.

Again, the theory in section 3.3 now comes into play. The dimension predicts and classifies the resultant fractal set X, telling us what sort of pattern will emerge after repeated iterations of a given seed. If $D_{HB}(X) = 0$, we will see only isolated dots. For $0 < D_{HB}(X) < 1$, X will be is totally disconnected. When $D_{HB}(X) = 1$, the classification begins to get complicated. If X is a Jordan curve or arc, then X is a rectifiable curve or arc i.e., having finite length and not oscillating too much, which has a tangent at almost all of its points. However, X can also be a product of two (or more) Cantor sets. Complete characterizations are possible if we split the set into curve and curve–free components (see [7]). For $1 < D_{HB} < 2$, several things can happen. If X is a Jordan curve or arc, then we will approximate a curve or a set of curves with the property that the curves have tangents almost nowhere, similar to the Koch arc. If X is intersecting, intersection points are ramification points, similar to the Sierpinski gasket. The set X could be a product of Cantor sets, or Cantor sets with lines or circles, etc. If $D_{HB} = 2$, iterates of the seed will, in the limit, be space–filling. Finally, if $D > 2$, iterates of the seed will quickly fill up regions, "overlapping" many points $\frac{D}{2}$ times.

We give the following examples. The first two come from Mandelbrot [12]. The "monkey's tree," plates 31 and 146 from [12], is made up of six segments of length $\frac{1}{3}$ and five segments of length $\frac{\sqrt{3}}{9}$, with various orientations. Let M_1 be the set which results from iteration of the seed (see Figure 3). Then, by equation (21), $D_{HB}(M_1) = D$, where D is the solution of

$$6(\frac{1}{3})^D + 5(\frac{\sqrt{3}}{9})^D = 1\,.$$

By noting that $(\frac{1}{\sqrt{3}})^2 = \frac{1}{3}$ and $(\frac{1}{\sqrt{3}})^3 = \frac{\sqrt{3}}{9}$, we see that $(\frac{1}{3})^{D/2} = x_0$, where x_0 is the positive solution of $5x^3 + 6x^2 - 1 = 0$. Solving, we get that

$$D = \frac{2\log(\frac{10}{\sqrt{21}-1})}{\log(3)} = 1.868726764\ldots\,.$$

Thus, M_1 will be a curve which has a tangent nowhere, but is not space filling (in spite of the fact that iterates of the seed will "fill up" patches of the computer screen). The "snowflake sweep,"

plate 68 from [12], is made up of six segments of length $\frac{1}{3}$ and one segment of length $\frac{\sqrt{3}}{3}$, with various orientations. Let M_2 be the set which results from iteration of the seed (see Figure 4). Then, $D_{HB}(M_2) = 2$, because

$$6(\frac{1}{3})^2 + (\frac{\sqrt{3}}{3})^2 = 1\,.$$

The curve M_2 fills the interior of the Koch snowflake.

We also can produce curves with variable dimension using a "fat" Cantor set, a "sliding" Koch generator, and two sides of a isosceles triangle.

A "fat" Cantor set can be produced by a seed consisting of the line segments

$$[0, \frac{3-\alpha}{6}]\,,\ [\frac{3+\alpha}{6}, 1]\,,\ 0 < \alpha < 1\,.$$

Here, we are removing middle interval of length $\frac{\alpha}{3}$. If C_α is the set that is produced by iteration of the seed, equation (21) gives that

$$D_{HB}(C_\alpha) = \frac{\log(2)}{\log(\frac{6}{3-\alpha})}\,.$$

Since $0 < D_{HB}(C_\alpha) < 1$ for all α, $0 < \alpha < 1$, C_α is totally disconnected. If the base is $[0,1]$, the total mass of C_α, as measured by Lebesgue measure,[10] is $1 - \alpha$. Note that we can make α arbitrarily close to zero. Interesting variations on this set can be created by taking its Cartesian product with line segments or circles (see [7, pp. 95–96]).

We can make a "sliding" Koch generator as follows. Let $\alpha \in (1/4, 1/2)$, and let the generator be given by the four line segments with endpoints

$$(0,0),\ (\alpha,0),\ (\frac{1}{2}, \sqrt{\alpha - \frac{1}{4}}),\ (1-\alpha,0),\ (1,0).$$

Proceed as in the generation of the Koch arc, i.e., each of the four segments is replaced by the seed pattern, and so on.

Let K_α be the set that is produced by iteration of the seed. Using (21), we compute that

$$D_{HB}(K_\alpha) = \frac{\log(4)}{\log(\frac{1}{\alpha})}\,.$$

[10] See [8] for reference.

As $\alpha \longrightarrow 1/4$, K_α approaches the line segment $[0,1]$. As $\alpha \longrightarrow 1/2$, K_α becomes space-filling. This is consistent with the theory. For $\alpha = 1/4$, $D_{HB}(K_\alpha) = 1$, while if $\alpha = 1/2$, $D_{HB}(K_\alpha) = 2$. For $\alpha \in (1/4, 1/2)$, $1 < D_{HB}(K_\alpha) < 2$, and the resultant non–intersecting arc is of infinite length and has a tangent nowhere.

Our final example gives a simple generator which produces a wide range of behavior. We will use two sides of a isosceles triangle. We let these be the line segments with endpoints $(0,0)$, $(\frac{1}{2}, \alpha)$, $(1,0)$, $(\frac{1}{2}, \alpha)$, respectively. With the orientation produced by connecting the line segments in the order they were written, and $\alpha = \frac{1}{2}$, iterates of this produce Mandelbrot's "dragon sweep" (see plate 66 of [12]), which is a space–filling curve. For variable α, $0 \leq \alpha \leq \frac{1}{2}$, the general seed pattern will produce self-similar curves. With α in this range, we have that

$$D_{HB}(X) = \frac{\log(2)}{\log(\frac{1}{\sqrt{\alpha^2 + \frac{1}{4}}})}.$$

If $\frac{1}{2} < \alpha < \frac{\sqrt{3}}{2}$, then iterates of this pattern will produce an image by overwriting it. In fact, if ℓ is the length of each side of the triangle, and if D is the solution of $2\ell^D = 1$, then we can produce a set with $\frac{D}{2}$ overlap by setting

$$\alpha = \sqrt{\frac{1}{2^{\frac{2}{D}}} - \frac{1}{4}}$$

for $2 < D < \infty$, which gives a range of $\frac{1}{2} < \alpha < \frac{\sqrt{3}}{2}$. Note that we can make D arbitrarily large. For example, we can let $\alpha = \sqrt{\frac{1}{\sqrt[50]{2}} - \frac{1}{4}}$, which then makes $\ell = \frac{1}{\sqrt[100]{2}}$ and $D = 100$. Iterates of the seed overwrite each other 50 times.

Acknowledgements

The author would like to thank Lawrence Crone and Richard Holzsager of American University and Nicholas Reingold of AT&T Bell Labs for mathematical and programming advice.

References

[1] M. F. Barnsley, *Fractals Everywhere*, Academic Press, Boston (1988).

[2] S. D. Casey, "Formulating fractals," *Computer Language* **4**, 4, 28–40, cover (1987).

[3] S. D. Casey, "Analysis of fractal and Pareto–Levy sets: theory and application," *Proceedings of EFTF*, 205–211 (1990).

[4] S. D. Casey and N. F. Reingold, "Self–similar fractal sets: theory and procedure," *IEEE Computer Graphics and Applications* **14**, 3, 73–82 (1994).

[5] G. A. Edgar, ed., *Classics on Fractals*, Addison–Wesley, Reading, MA (1993).

[6] K. Falconer, *The Geometry of Fractal Sets*, Cambridge University Press, Cambridge (1985).

[7] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley and Sons, Chichester (1990).

[8] H. Federer, *Geometric Measure Theory*, Springer–Verlag, New York (1969).

[9] J. Hocking and G. Young, *Topology*, Dover Publications, New York (1988).

[10] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton University Press, Princeton, N.J. (1941).

[11] J. E. Hutchinson, "Fractals and self–similarity," *Indiana Univ. Math. J.* **30**, 713–747 (1981).

[12] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York (1983).

[13] J. R. Munkres, *Topology: A First Course*, Prentice–Hall, Englewood Cliffs, N.J. (1975).

[14] H.–O. Peitgen and D. Saupe, *The Science of Fractal Images*, Springer–Verlag, New York (1988).

[15] C. A. Pickover, ed., *The Pattern Book: Fractals, Art, and Nature*, World Scientific, Singapore (1995).

[16] S. J. Taylor, "The measure theory of random fractals," *Math. Proc. Camb. Phil. Soc.* **100**, 383–406 (1986).

[17] B. West, *An Essay on the Importance of Being Non–Linear*, Springer–Verlag, Berlin (1985).

Figure 1 – A Peano construction P.

$$D_{HB}(P) = \lim_{n\to\infty} \frac{\log(2\cdot 4^n)}{\log \frac{1}{(\sqrt{2}/2)(1/2)^n}} = \lim_{n\to\infty} \frac{\log(2\cdot 4^n)}{(1/2)\log(2\cdot 4^n)} = 2.$$

Figure 2 – Arc A_{S_g} approximating the Sierpinski Gasket.

$$D_{HB}(A_{S_g}) = \frac{\log(3)}{\log(2)} = 1.584962501\ldots .$$

Figure 3 – Mandelbrot's "Monkey's Tree" M_1.

$$D_{HB}(M_1) = \frac{2\log(\frac{10}{\sqrt{21}-1})}{\log(3)} = 1.868726764\ldots .$$

Figure 4 – Mandelbrot's "Snowflake Sweep" M_2.

$$D_{HB}(M_2) = 2, \text{ because } 6(\frac{1}{3})^2 + (\frac{\sqrt{3}}{3})^2 = 1\,.$$

Figure 5 – Peano–Gosper Curve G.

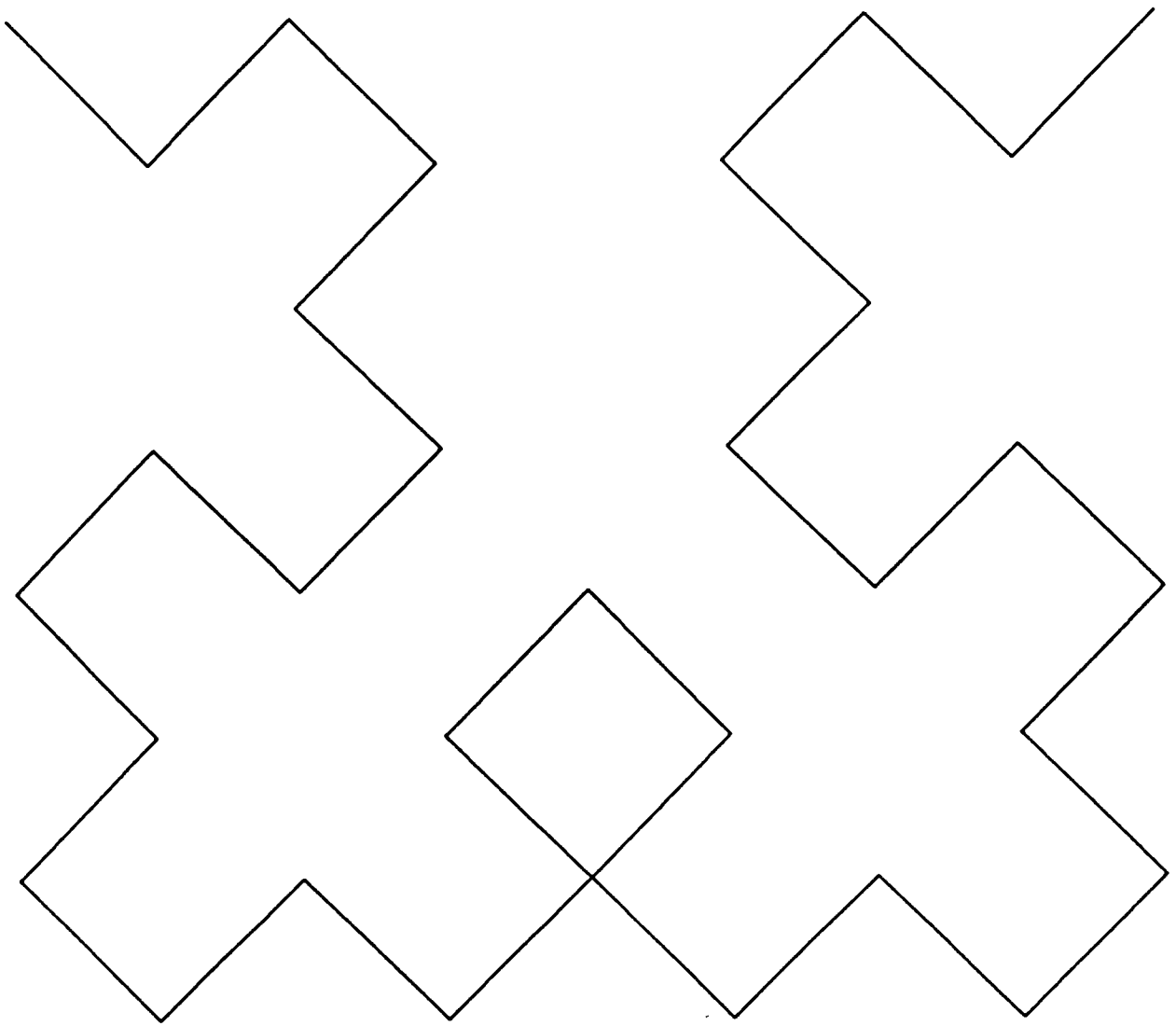
$$D_{HB}(G) = \frac{\log(7)}{\log(\sqrt{7})} = 2\,.$$

Figure 6 – Antenna A.

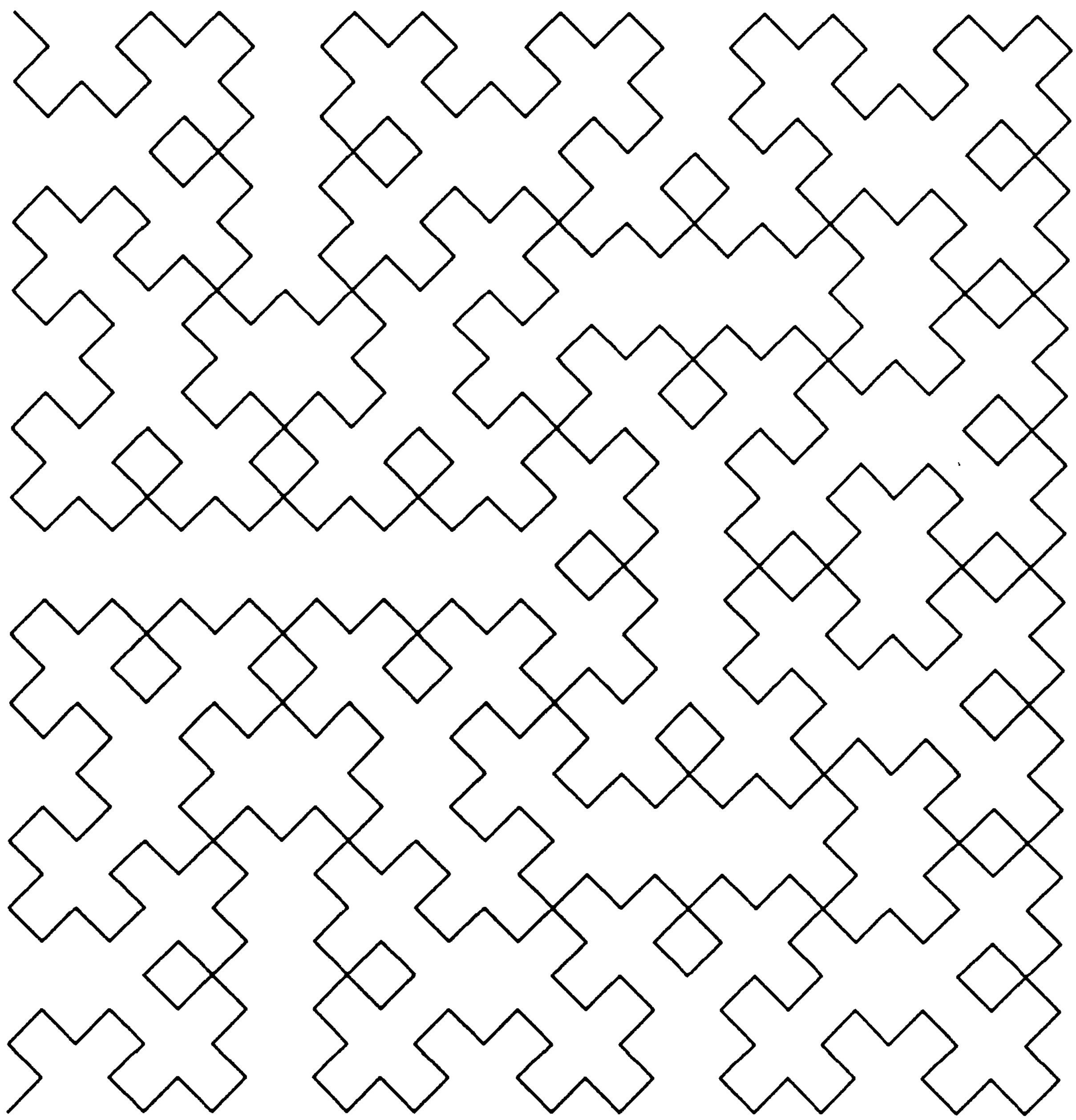
$$D_{HB}(A) = D = 1.746141227\ldots, \text{ because } 2\left(\frac{1}{2}\right)^D + 2\left(\frac{2}{5}\right)^D = 1\,.$$

Figure 7 – Asymmetric Sierpinski Carpet S_c (Also see [15]).

$$D_{HB}(S_c) = \frac{\log(6)}{\log(3)} = 1.630929754\ldots .$$



1.c)



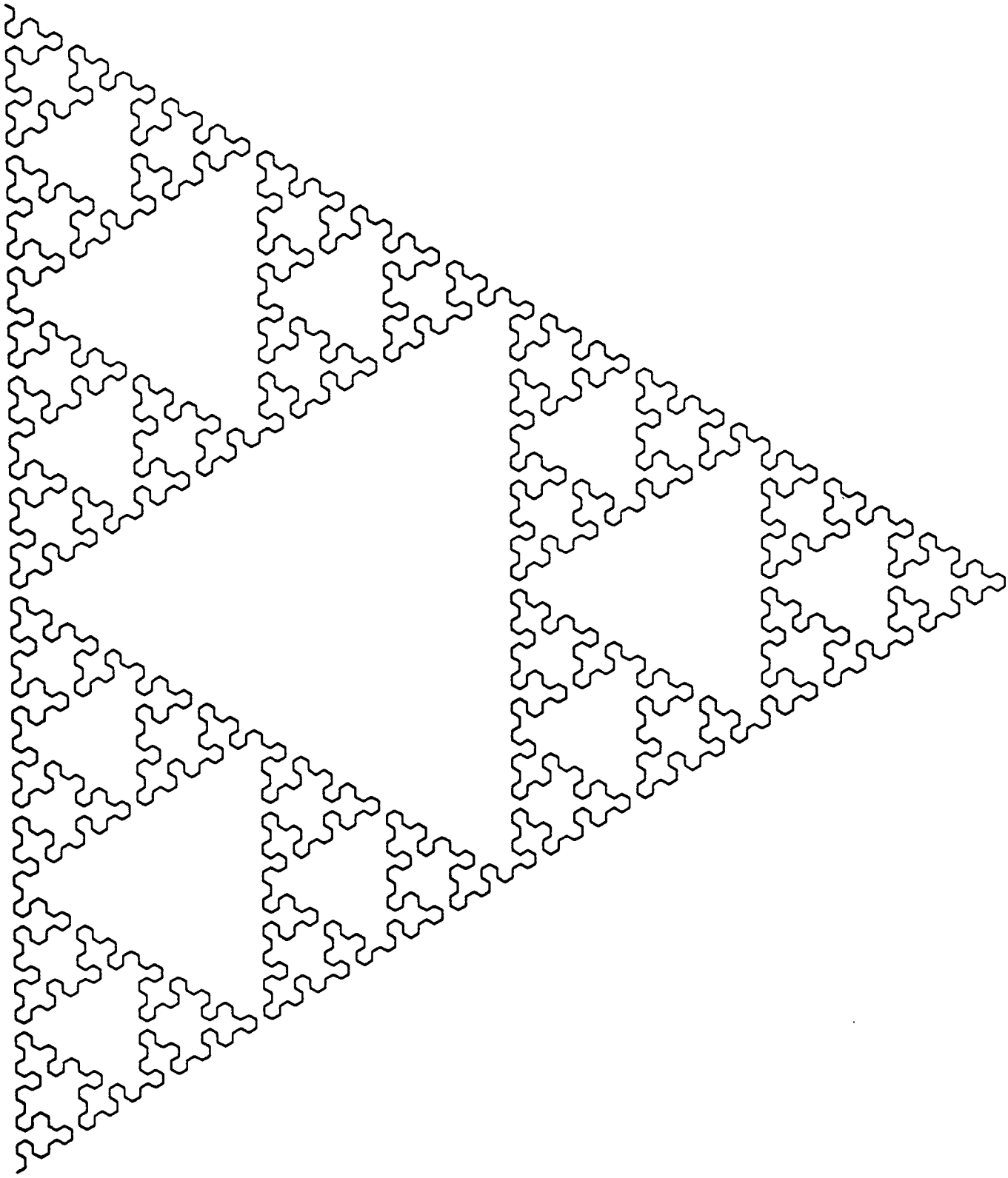
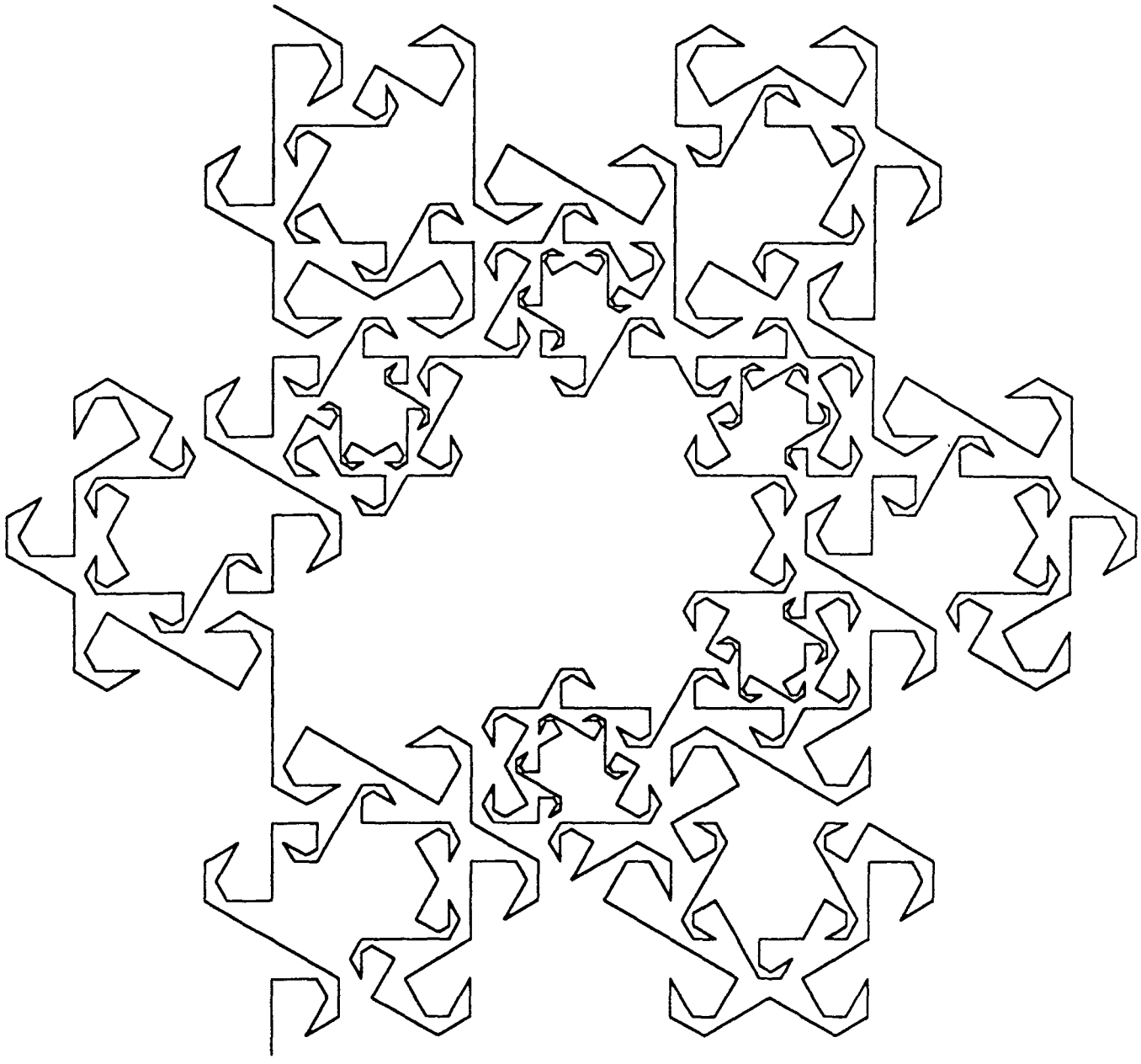
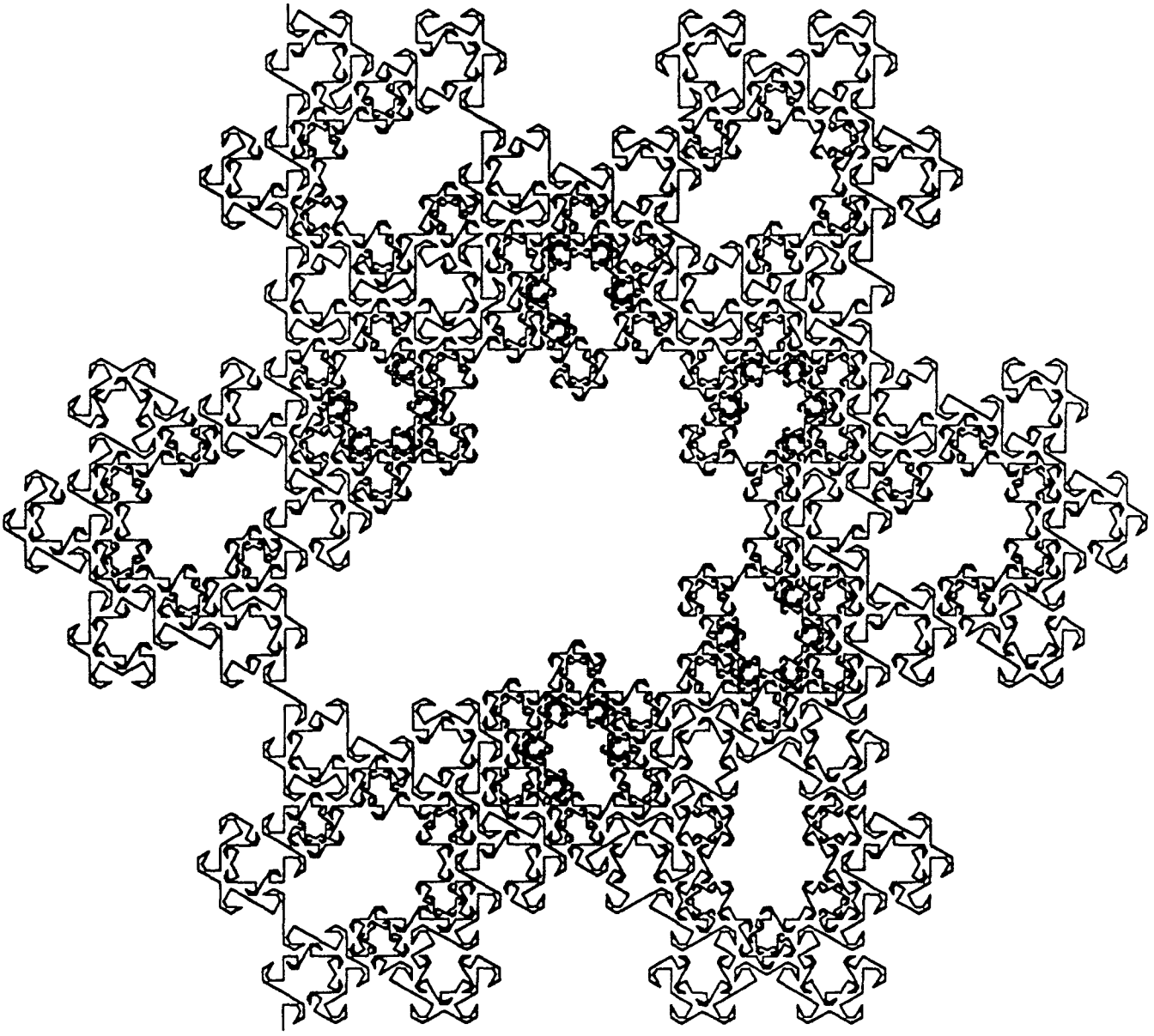
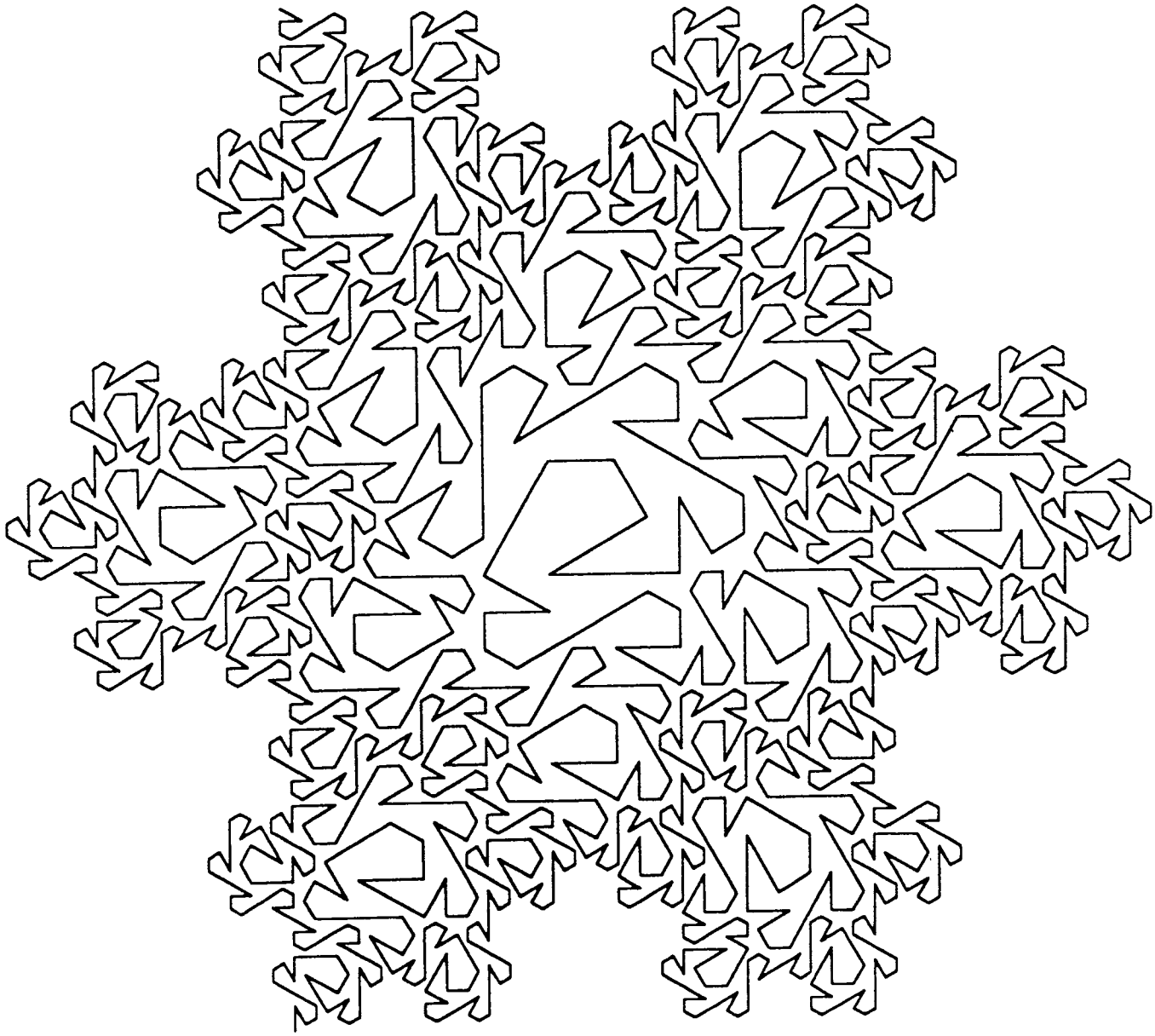
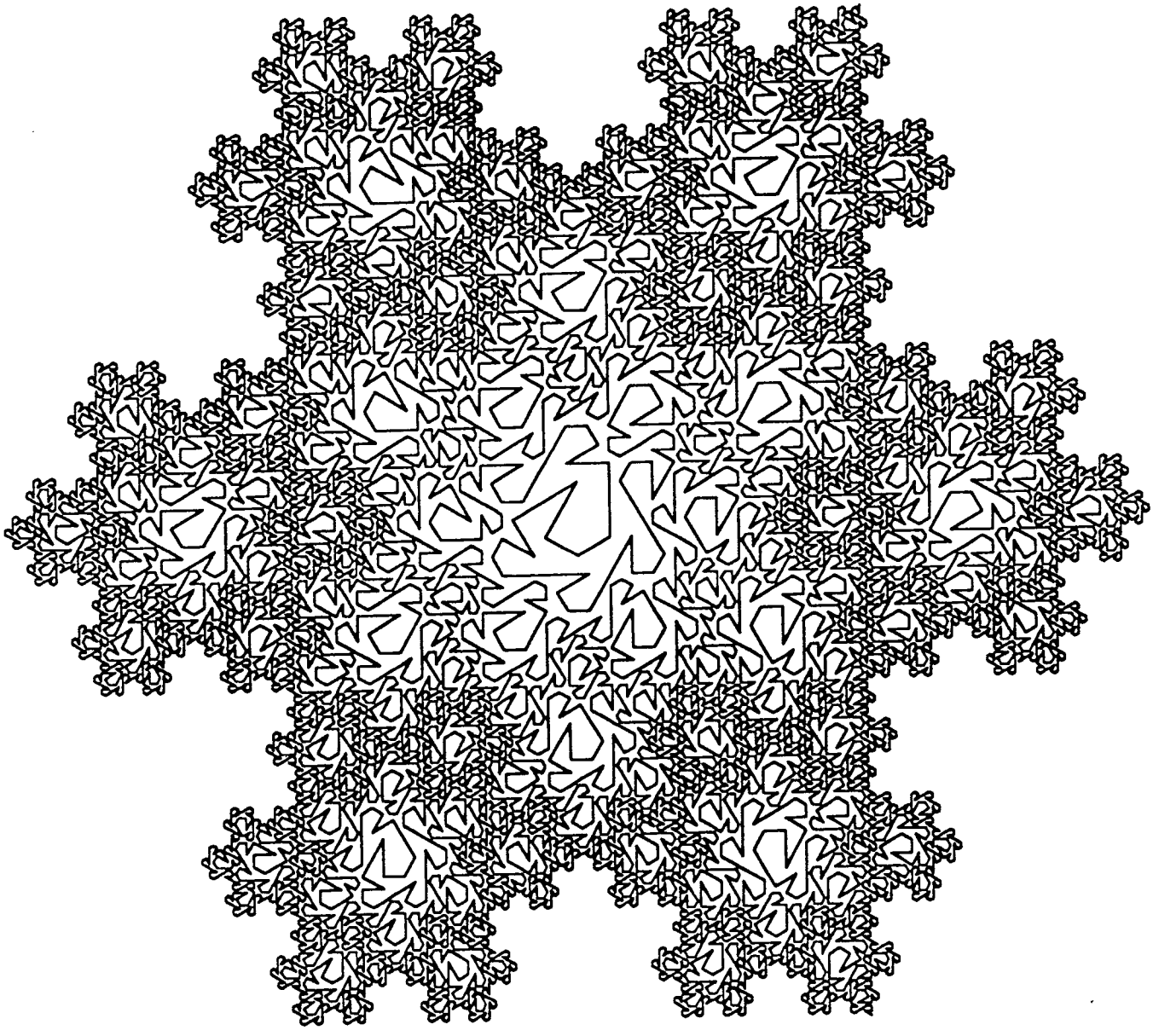


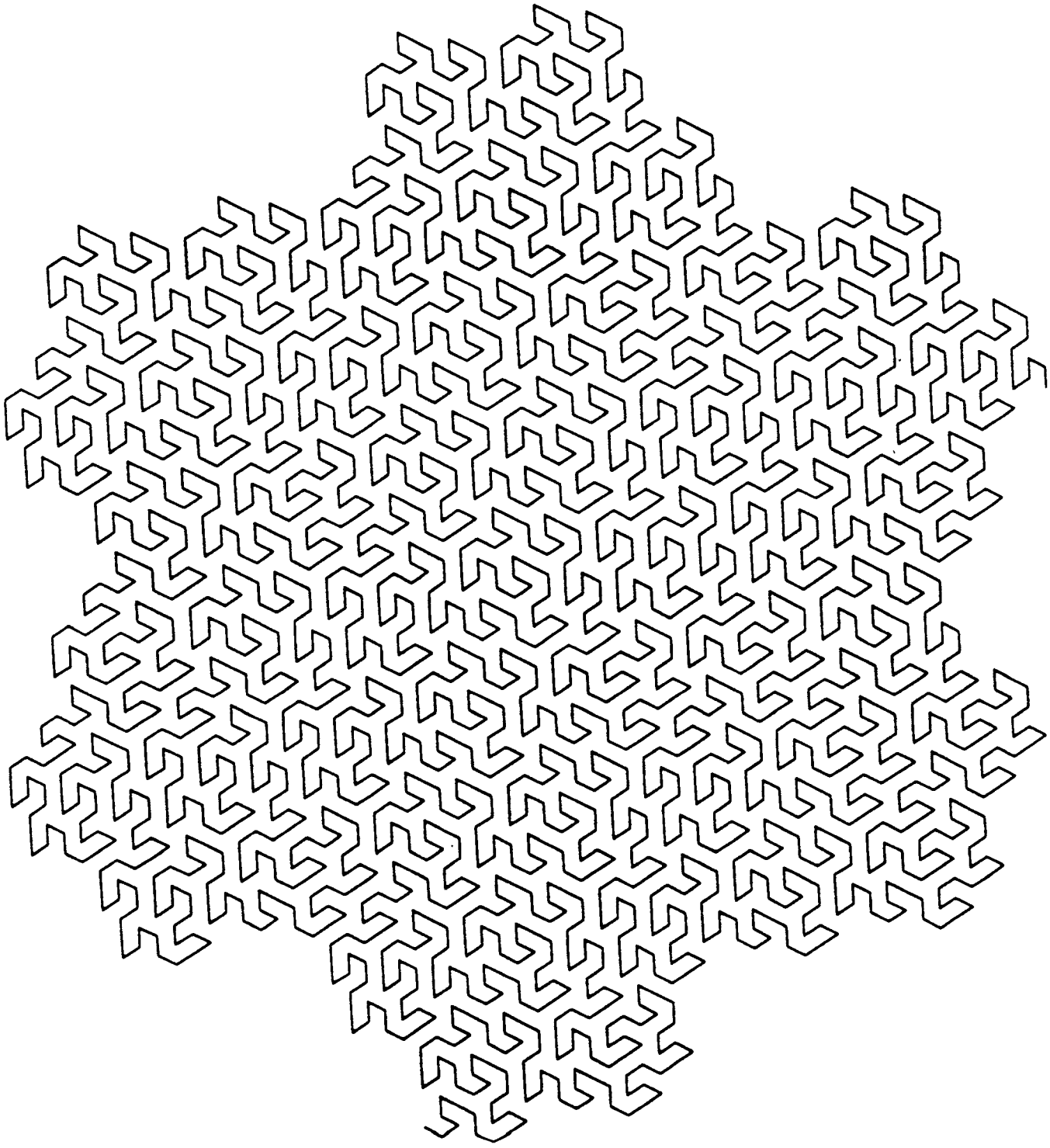
Fig. 1



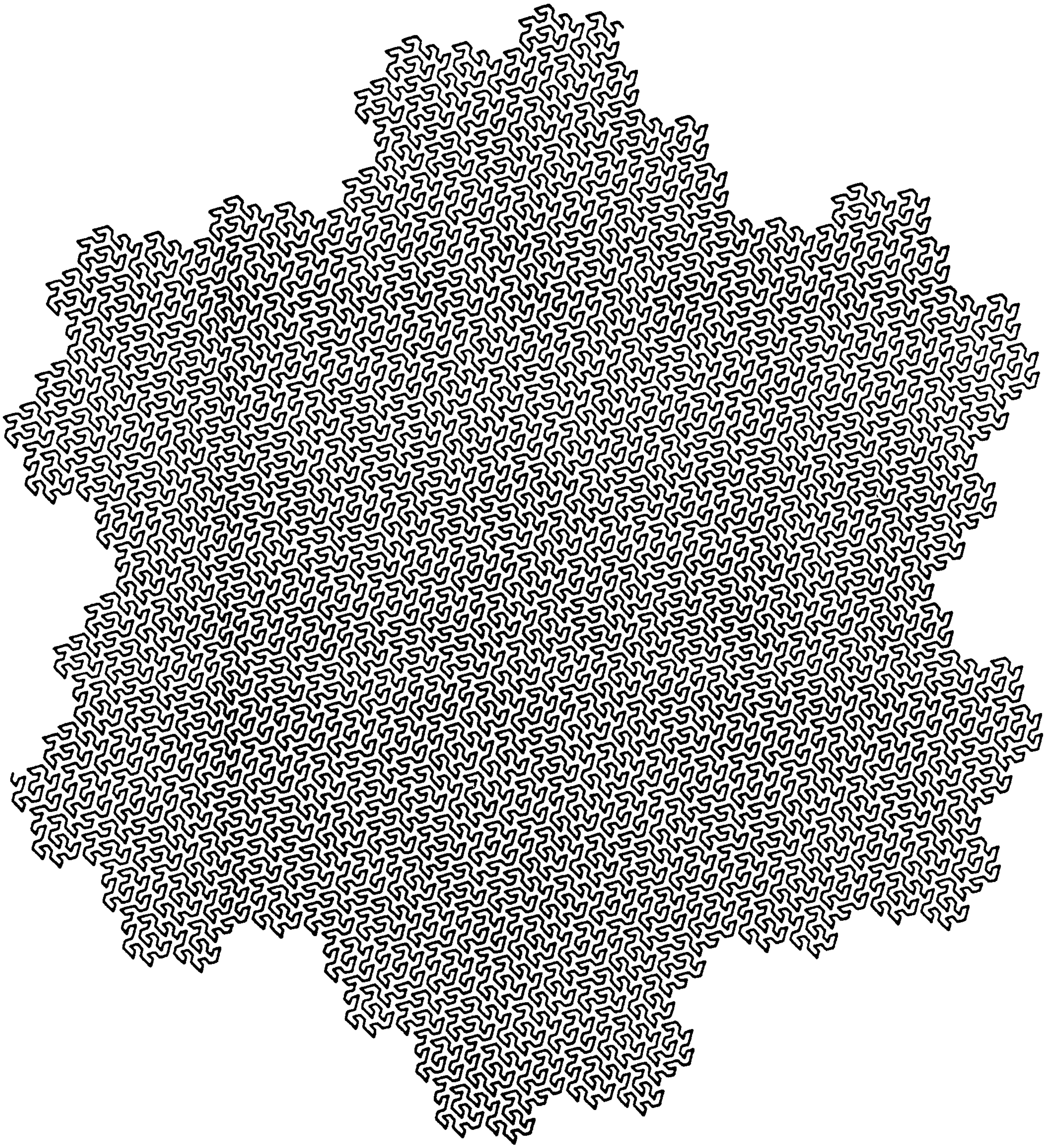


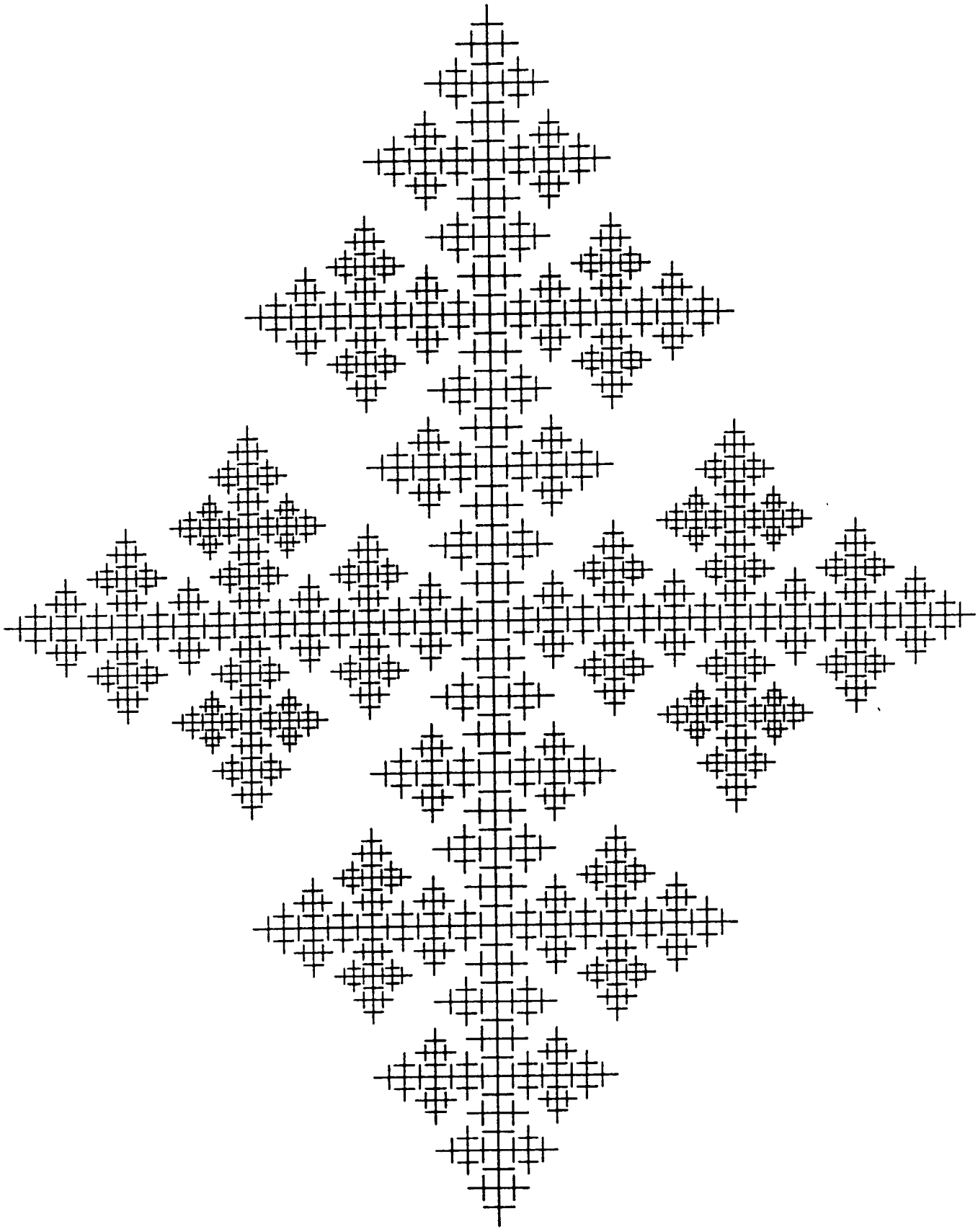






5a.)





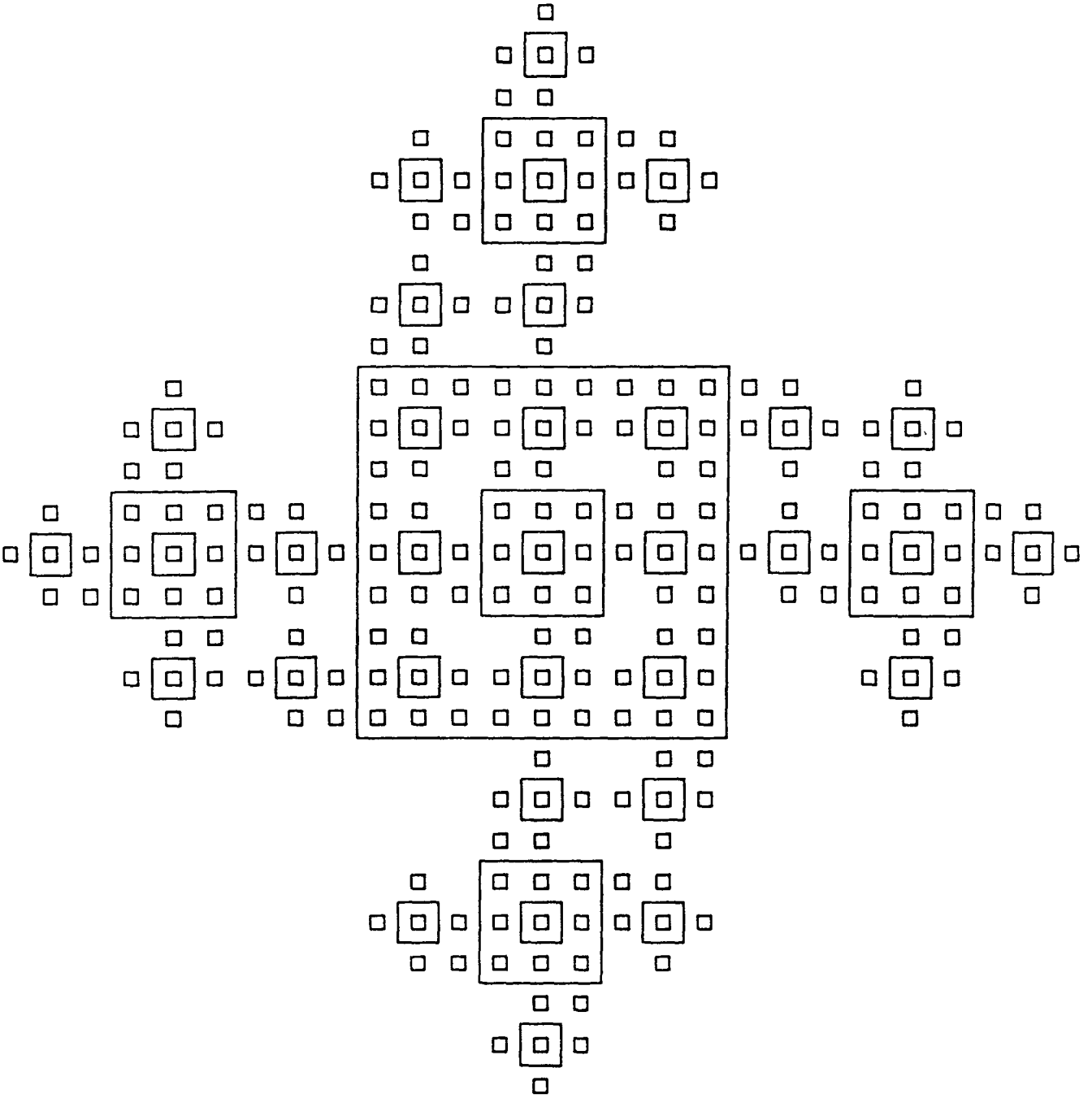


Fig.