

# Fractals and the Large-Scale Structure in the Universe

## 1. Introduction and Basic Concepts

*A K Mittal and T R Seshadri*



A K Mittal is currently a visiting scientist at Harish-Chandra Research Institute, Allahabad, on leave from the University of Allahabad. His research interests include chaos, fractals and their applications, particularly in cosmology and in weather prediction.



T R Seshadri is at the Department of Physics and Astrophysics, University of Delhi. He is presently on lien from Harish-Chandra Research Institute, Allahabad. Areas of his interest include cosmic microwave background radiation, large scale structures in the Universe and application of fractals in these.

During the last decade it has been argued by some investigators that the distribution of galaxies in the Universe is a fractal. This is contrary to the standard view that the Universe is homogeneous and isotropic on large scales. In this article, the basic concepts of fractals and their characterization are introduced with the help of simple examples. The applications of these concepts in the context of cosmology will be discussed in the second part.

### Introduction

One of the most interesting aspects of our universe is that there are structures of different sizes like planets, stars, star clusters, galaxies, clusters of galaxies, etc. Of these structures, the ones smaller than galaxies are believed to have been formed due to a number of physical processes in addition to gravity. Structures of galactic size and larger are supposed to have been formed predominantly due to gravitational condensation of matter in the universe. The latter structures are broadly referred to as large-scale structures and understanding the physics operating at these scales is an important challenge to present day cosmology.

There have been various possible descriptions of the matter distribution of the Universe. The simplest and the most common assumption has been that the universe is homogeneous over very large scales and within this framework of the homogeneous Universe, we have



structures of different sizes. Another view is that the Universe has a fractal distribution of matter.

With the aim of understanding the origin of this controversy, in this part we will introduce some basic notions of fractals.

### Measuring the Length of a Curve

Consider the problem of measuring the length of a curve. The standard procedure to measure the length of a curve would be the following: We use a straight ruler of length  $l$ . Starting from one end of the curve, we walk the ruler along the curve for its entire length. In doing so let us denote by  $N$  the total number of steps we had to take to go from one end of the curve to the other. [The meaning of ‘walking the ruler’ along the curve is given in *Box 1*.] Measured like this the length of the curve turns out to be approximately  $Nl$ . Needless to say that this measure is approximate. The inaccuracy arises from the large value of  $l$ . So we now repeat this exercise using a scale of smaller length. Naive logic would suggest that as we go to the limit of  $l$  tending to 0, the product  $Nl$  will gradually become more and more independent of  $l$  and this product is what we term the length of the curve. In other words,  $N$  is proportional to  $1/l$  as  $l \rightarrow 0$ . An example of this method for measuring the circumference of a circle is given in *Box 2*. Hidden in the above method of measuring length is an important assumption; that of the smoothness of the curve (*Box 3*). When we use

#### Box 1. Walking the Ruler along the Curve

We explain here, what we mean by walking the ruler along the curve. If we have a closed curve, then we choose a point on it (say O) and keep an end of the ruler at this point. With O as the pivot, we swing the ruler so that the other end of the ruler cuts the curve at a point (say A). With A as the pivot, we again swing the ruler as to intersect the curve at B. Once again we use B as the pivot and proceed in this manner till we reach the point where we started (or the nearest to it without overshooting).

Keywords.

Fractals, large-scale structures in the Universe, cosmological principle, conditional cosmological principle.

#### Box 2. Measuring Circumference of a Circle

As an example of measuring the length of a smooth curve, we consider the measurement of the circumference of a circle. Consider a circle of radius  $R$ . Inscribe in it a regular polygon of  $n$  sides. Let the length of each side of the polygon be  $l$ . So the perimeter of the polygon is  $nl$ . If the chord formed by the side of the polygon subtends an angle  $\theta$ , the number of sides,  $n = 2\pi/\theta$ . Further, the side  $l = 2R \sin(\theta/2)$ . So the perimeter comes out to be,  $nl = 2\pi R \sin(\theta/2)/(\theta/2)$ . In the limit of  $l$  tending to zero,  $\theta$  also tends to zero and  $\sin(\theta/2) \rightarrow \theta/2$ .



smaller and smaller rulers, basically we are probing the curve on smaller and smaller scales. As we have argued in *Box 3*, if the curve is smooth, on sufficiently small scales, the curve can be approximated to a straight line. So the deviation of the curve from the ruler (at every step of walking it along the curve) keeps reducing as the size of the ruler is reduced. In other words, the chord formed by the ruler is a good approximation to the arc of the curve over that small scale.

One can, however, envisage a situation, where such a nice limit does not exist. This can happen when the curve is such that it remains wiggly even on smaller scales. A practical situation when such a thing can happen is while measuring the length of the border between two countries. The lengths of the common frontiers between Spain and Portugal, or between Belgium and Netherlands, as reported in these neighbours' encyclopedias differ by 20 percent. This difference can arise if

### Box 3. Smoothness of a Curve

We have used the phrase 'a smooth curve'. In this box we explain what is meant by this in the present context. Consider a curve described by the equation

$$y = f(x). \quad (1)$$

We further assume that the function  $f(x)$  is differentiable everywhere. We know from calculus that  $y$  at a point  $x$  in the neighborhood of  $x_0$  can be expressed as a Taylor series in terms of  $y$  at  $x_0$ .

$$y(x) = y(x_0) + (x - x_0)[df/dx]_{x_0} + \text{higher degree terms}. \quad (2)$$

For sufficiently small interval  $(x - x_0)$  we can approximate this as

$$y(x) \sim y_0 + [df/dx]_{x_0}(x - x_0). \quad (3)$$

As the reader would immediately recognize, this is the equation for a straight line. In other words, the curve can be approximated as a straight line at sufficiently small interval in  $x$ . This is the reason, why we are able to measure their lengths by using straight rulers if the rulers are sufficiently small. Clearly this is possible only if the function is Taylor expandable about any point. And this in turn is possible only if the function  $f$  is differentiable everywhere. Such curves we call smooth.



the length of the ruler differs just by a factor 2. Mathematically, we can have a situation where this happens at all scales, however small. In these cases,  $L(\epsilon) \rightarrow \infty$  as  $\epsilon \rightarrow 0$ , where,  $L(\epsilon)$  is the length as measured by a ruler of size  $\epsilon$ . But it is not of much use to say that the length of these curves is infinite. It is more appropriate to say that the length is not a useful ‘measure’ of these non-rectifiable curves. The curve does not become smooth at however small a scale we probe it. If we consider 2 points on such a curve, and we keep bringing the points closer and closer to each other, but always keeping them on the curve, we will never reach a situation when the portion between the two points tends to a straight line. Put mathematically, these curves are not differentiable anywhere. They cannot be Taylor expanded about any point. Such a curve is an example of a fractal.

To understand such structures, we need to take a more quantitative look at these. In the next section, we describe an algorithm to generate a particular type of fractal, the Koch curve. This example will help us appreciate the unfamiliar behaviour of these structures.

For a quantitative discussion, we need to discuss how to generate a fractal in a more systematic way and how to characterize a given fractal.

We first give some systematic algorithms to generate some fractals.

### Generating a Simple Fractal: The Koch Curve

The concept of fractals is most easily explained through a simple example. (For a good discussion on fractals see references [1, 2]). A typical example of a fractal curve is the Koch curve. This curve is constructed as follows: First consider a straight line of unit length as shown in *Figure 1*.

Now, with the middle third of this line as base, construct an equilateral triangle and remove the base to get stage





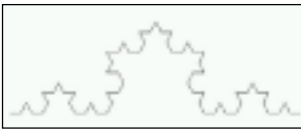
**Figure 1. The seed segment.**



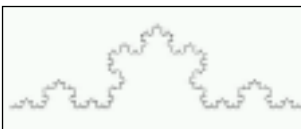
**Figure 2. Stage 1 of Koch curve.**



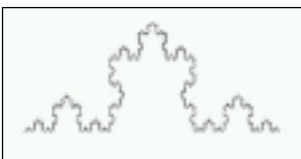
**Figure 3. Stage 2 of Koch curve.**



**Figure 4. Stage 3 of Koch curve.**



**Figure 5. Stage 4 of Koch curve.**



**Figure 6. Stage 5 of Koch curve.**

1 as shown in *Figure 2*.

This figure is made up of four segments of length  $1/3$  each, so that the total length is  $4/3$  units. Now, with the middle third of each segment as base, construct equilateral triangles and remove the bases to get stage 2 as shown in *Figure 3*. This figure is made up of  $4^2$  segments of length  $1/3^2$  each, so that the total length is  $(4/3)^2$  units. Continuing this process, we get stage 3 as shown in *Figure 4*. Continuing further, we get stage 4 as shown in *Figure 5* and stage 5 as shown in *Figure 6*. As this process is continued the length of the curve goes on increasing by a factor of  $4/3$  at each stage, although the curve is bounded, i.e. confined to a finite region of the plane. The curve becomes more and more winding; and approaches a limiting curve, called the Koch curve, which has no tangent anywhere. As the size of the ruler used to measure the length of the curve goes on decreasing, the computed length does not approach a finite limit. More specifically if a ruler of length  $3^{-k}$  is walked over the Koch curve, it will walk over the  $k$ th stage of the above construction as the structure introduced after the  $k$ th stage is smaller than the ruler resolution. Thus  $4^k$  steps will be required and the length will be computed as  $(4/3)^k$ . Clearly as  $k$  increases indefinitely the computed length does not approach a finite limit. Thus the usual notion of length fails to be useful for a curve of this type. The primary reason being that one cannot uniquely define this parameter for fractal curves. It is tied down to the size of the measuring scale. However, there is a related parameter which is termed 'dimension' and which is useful to describe these structures. Although the length of the curve depends on the size of the scale, we can ask how the measured length varies with the ruler size. The notion of dimension is related to the scaling relation of the curve 'length' with the ruler size.



## Fractal Characterization: Generalizing the Notion of Dimension

Let us consider a curve. We enclose it in a cube of size  $L$ . We next divide the cube into equal cubes each of size  $L/k$ . Clearly the number of cubes is  $k^3$ . However, the curve need not pass through each of these cubes. Let the number of cubes through which the curve passes be  $N_1$ . This is the first generation. We now get the second generation characterization by dividing each of these small cubes further to a size  $L/(2k)$ . The total number of cubes is now  $(2k)^3$ . Some of these cubes are empty. We can again compute the number of non-empty cubes. In this way we get larger and larger number of smaller and smaller cubes. In general we find that the total number of non-empty cubes denoted by  $N$  is proportional to  $l^D$ , where  $l$  is the linear size of the smallest elemental cube. So given a distribution of points (a curve being a special case) we evaluate  $D$  by computing,

$$D = \lim_{l \rightarrow 0} \frac{d \ln(N)}{d \ln(l)}. \quad (1)$$

The reader may rightly wonder why it is called a dimension. After all what we have computed is merely some parameter related to how the number of non-empty boxes scales with the size of the box. The reason why it is called a dimension is that if we consider a one-dimensional smooth curve and do this exercise, the exponent  $D$  will turn out to be 1. Similarly if we do this to a two dimensional smooth surface embedded in 3 dimension and perform the above procedure then the exponent will turn out to be 2. So we generalize the concept of dimension and state that this exponent computed in the above way will be called dimension of the object. The interesting thing is that in the case of fractals, this exponent need not be an integer. It can even be a fraction. This is the origin of the nomenclature of such objects. In *Box 4* we give an example of computing the dimen-



**Box 4. Box-counting Dimension**

If we think that the given points constitute a 3-dimensional object and not a curve then we will not try to determine the length, we will try to determine its volume instead. To do this we will take a grid of cubes of side  $\epsilon$ , find  $N(\epsilon)$ , the number of cubes that contain some part of the object, and estimate the volume by  $\epsilon^3 N(\epsilon)$ . We will expect this estimate to improve as the grid size  $\epsilon$  is made smaller and smaller. Suppose that the given set of points constituted a 1-dimensional object but we tried to find its volume by the above procedure. It is easy to see that in this case  $N(\epsilon)$  will be proportional to  $\epsilon^{-1}$  so that the limit  $\epsilon^3 N(\epsilon)$  as  $\epsilon \rightarrow 0$  will be 0 and the volume of a 1-dimensional object will be computed as 0. The interesting measure of a 1-dimensional object is its length and not the volume. However, we need not walk a ruler on this object to find its length. We can find the length by using the procedure for finding the volume. To do this we simply take the limit  $\epsilon N(\epsilon)$  as  $\epsilon \rightarrow 0$ .

To illustrate this method in detail, draw a circle of radius, say 8 cm. Let a grid of side  $\epsilon = 8$  cm be drawn. Then  $N(\epsilon) = 4$  squares will contain some part of the circle. Reduce the grid side to  $\epsilon = 4$  cm. Then  $N(\epsilon) = 12$  squares are found to contain some part of the circle. Reducing the grid side to  $\epsilon = 2$  cm,  $N(\epsilon) = 28$  squares are found to contain some part of the circle. Further reducing grid side to  $\epsilon = 1$  cm,  $N(\epsilon) = 56$  squares are found to contain some part of the circle. In this way one can see that  $N(\epsilon)$  is proportional to  $\epsilon^{-1}$ .

sion of a circle in this way and arriving at the intuitive result of 1. In *Box 5* we do this exercise for a Koch curve and we show that this exponent turns out to be a fraction. The above definition of dimension is called box-counting dimension. This is just one of the many definitions. An obvious question arises at this point. If we use different ways to compute dimension, do we get a unique result? In other words, do all these results give the same value of dimension? The answer to this is that, this may or may not be so. And whether or not this result is unique defines the basis of dividing the fractals into two classes. Let us see what are the other definitions that generalize the standard notion of dimension.

**Self-Similarity Dimension**

Another useful definition of dimension is the self-similarity dimension. At each stage in the construction of the Koch curve, the previous figure is reduced by a factor of three, and four such reduced figures make the next



figure. The limiting figure that will be obtained if the process is continued through infinite stages will be such that if reduced by a factor of three, four such reduced figures will reproduce the original figure. Self-similarity dimension  $D_s$  depends on these two numbers. If the figure is reduced by a factor  $r$ , and  $n$  such reduced figures reproduce the original figure, we say that the figure is self-similar with self-similarity dimension

$$D_s = \log(n)/\log(r). \quad (2)$$

Thus the self-similarity dimension for the Koch curve is

$$D_s = \log(4)/\log(3). \quad (3)$$

Similarly for the case of a Cantor set (*Box 6*), the similarity dimension turns out to be  $\ln(2)/\ln(3)$ . We can see this by noting that if we begin with a set  $C_0 = [0, 1]$ , then after  $n$  steps,  $C_n$  will consist of  $2^n$  intervals each of length  $3^{-n}$ . Moreover  $C_n$  can be made from  $2^k$  copies of  $C_{n-k}$  scaled by a factor of  $3^{-k}$ . Another example of a fractal is the Sierpinski gasket (*Box 7*).

### Correlation Dimension

The box counting dimension counts every box that has a point of the object inside the box; it does not care how many points of the object there are in the box. For a mathematical fractal, quite obviously, if a box contains one point, it contains infinite number of points, because the scaling behavior continues indefinitely. In physical applications, however, there is usually a lower and upper cutoff beyond which the fractal-scaling behavior does not hold. Moreover, instead of all the points of the object, one has access to only a relatively small sample. The concept of correlation dimension is useful for handling such situations. *The correlation dimension describes the scaling of  $N(r)$  (the fraction of point pairs of the object that have a distance less than  $r$  from each other) with  $r$ .* If  $N(r)$  is proportional to  $r^D$ ,  $D$  is called the correlation dimension.

#### Box 5. Box-counting Dimension for the Koch Curve

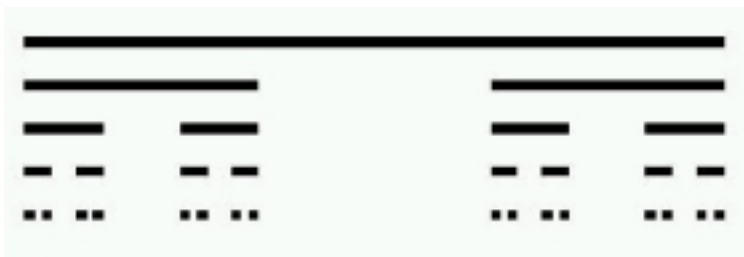
Let us enclose the segment of unit length given in *Figure 1* inside a square such that it forms the diagonal. Once we generate the first step of the Koch curve on this segment, we divide the square into nine equal parts by dividing each side into 3 parts. So the value of  $k$ , i.e. the factor by which we reduce each side is 3. We clearly see that the number of occupied squares is now four, i.e., the number of occupied squares increases by a factor of 4. There is a slight ambiguity here because some of the portions of the curve are along the borders of the elemental squares. We can lift this ambiguity by prescribing the rule that in such cases identify the portion with the square above. Box-counting dimension of the Koch curve is  $\ln(4)/\ln(3) \sim 1.2618$ .



**Box 6. The Cantor Set**

Another example of a self-similar object is that of the Cantor set. First begin with the set  $C_0 = [0, 1]$ , *i.e.*, a line of unit length. From this remove the middle one third to get  $C_1 = [0, 1/3]$  and  $[2/3, 1]$ . The set  $C_1$  consists of two lines of length  $3^{-1}$ . We see that  $C_1$  is made from two copies of  $C_0$  scaled by a factor  $3^{-1}$ . From each of the two intervals of  $C_1$  again remove the middle one third to get the set  $C_2 = [0, 1/9], [2/9, 1/3], [2/3, 7/9], [8/9, 1]$ . Again we see that the set  $C_2$  is made from two copies of  $C_1$  scaled by a factor of  $3^{-1}$ . We also see that the set  $C_2$  is made of  $2^2$  copies of  $C_0$  reduced by a factor of  $3^{-2}$ . Continuing the process we notice that  $C_n$  will consist of  $2^n$  intervals each of length  $3^{-n}$ . Moreover  $C_n$  can be made from  $2^k$  copies of  $C_{n-k}$  scaled by a factor of  $3^{-k}$ . These facts can be easily verified from *Figure A*.

The limiting set obtained in the limit  $n \rightarrow \infty$  is called the Cantor set  $C$ . It follows from the arguments given above that the limiting Cantor set is self-similar and  $2^k$  pieces of this set scaled by a factor of  $3^{-k}$  are required to obtain the original Cantor set. The self-similarity dimension is  $D_s = \ln(2)/\ln(3)$ .



*Figure A. Cantor set.*

**Some Common Misconceptions**

*Self-Similar Structures must always be Fractals – False*

One might think that a self-similar structure must necessarily be a fractal. But this is not true. For example, line segments, or squares, or cubes are all self-similar but not fractal. In these cases the reduction factor can be arbitrary. But for fractals the reduction factors are characteristic. For example, the Koch curve can be reduced only by factors of  $1/3^n$  ( $n$  integral), to obtain self-similarity.

However, for all self-similar structures, whether fractal or not, there is always a relation between the scaling factor and the number of scaled down pieces into which the structure is divided. Consider a line segment of length,

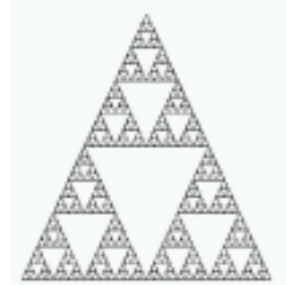


**Box 7. The Sierpinski Gasket**

Sierpinski gasket is another example of a self-similar object. Begin with an equilateral triangle of unit side. Join the middle points of the sides so that the triangle gets divided into four parts. Remove the central triangle (*i.e.*, the triangle made by joining the middle points). The figure now consists of three equilateral triangles of side  $2^{-1}$ . At the next stage, in each of these triangles, the central triangle, obtained by joining the middle points of that triangle, is removed. At this stage the figure consists of  $3^2$  equilateral triangles of side  $2^{-2}$ . As this process is continued at the  $n$ th stage we have a figure consisting of  $3^n$  equilateral triangles of side  $2^{-n}$ . The limiting figure of this process is called the Sierpinski gasket. It is easy to see that if one scales the Sierpinski gasket by a factor of  $2^{-k}$  then  $3^k$  such scaled figures will be required to reproduce the original Sierpinski Gasket. Hence by definition, the self-similarity dimension of the Sierpinski gasket is

$$\ln(3^k)/\ln(2^k) = \ln 3/\ln 2. \quad (1)$$

The  $n = 5$  approximation of the Sierpinski gasket is shown in *Figure B*.



**Figure B. Sierpinski gasket.**

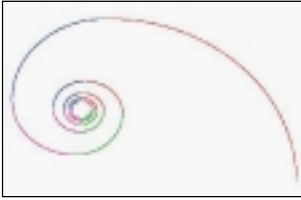
say 8 cm. If it is scaled down by a factor of  $1/2$  we will get a line segment of length 4 cm. Two such scaled down line segments put together produce the original line segment of length 8 cm. More generally, if a line is scaled down by a factor  $1/n$  then  $n$  such scaled down pieces are required to obtain the original line. The self-similarity dimension will be

$$D_s = \frac{\ln(n)}{\ln(\frac{1}{1/n})} = 1. \quad (4)$$

Consider a square of side, say 8 cm. If it is scaled down by a factor of  $1/2$ , we will get a square of side 4 cm. Four such scaled down squares put together produce the original square of side 8 cm. If a square is scaled down by a factor  $1/n$  then  $n^2$  such scaled down squares will be required to obtain the original square. The self-similarity dimension will be

$$D_s = \frac{\ln(n^2)}{\ln(\frac{1}{1/n})} = 2. \quad (5)$$





**Figure 7.** This curve is of infinite length and is in a bounded region but is not a fractal.

Consider a cube of side, say 8 cm. If it is scaled down by a factor of 1/2, we will get a cube of side 4 cm. Eight such scaled down cubes put together produce the original cube of side 8 cm. More generally, if a cube is scaled down by a factor 1/n, then  $n^3$  such scaled down cubes will be required to obtain the original cube. The self-similarity dimension will be

$$D_s = \frac{\ln(n^3)}{\ln(\frac{1}{1/n})} = 3. \tag{6}$$

If a Koch curve is scaled down by a factor  $3^{-n}$ , then  $4^n$  such scaled down Koch curves will be required to obtain the original Koch curve. The self-similarity dimension will be

$$D_s = \frac{\ln(4^n)}{\ln(\frac{1}{3^{-n}})} = \frac{\ln(4)}{\ln(3)}. \tag{7}$$

The above examples indicate that the self-similarity dimension is a generalization of our ordinary understanding of dimension in as much as ‘nice’ objects such as line, square and cube have self-similarity dimension equal to their ‘ordinary’ dimension.

*A Bounded Curve having Infinite Length must be a Fractal – False*

Consider a spiral made up of circle segments of radii  $a_k$  (see Figure 7) so that the arc length  $s_k$  of the  $k$ th circle segment is  $(\pi/2)a_k$ . Then the length of the spiral is given by

$$l = \sum_{k=1}^{\infty} s_k = (\pi/2) \sum_{k=1}^{\infty} a_k, \tag{8}$$

which diverges if  $a_k = 1/k$ .

*Fractal Dimension must be Fractional – False*

It is a common mistake to think of fractals as objects having non-integer dimensions. *Devil’s staircase and Peano curve are examples of fractal curves that have*

**Suggested Reading**

- [1] J Feder, *Fractals*, Plenum Press, New York, 1988.
- [2] H O Peitgen, H Jurgens and D Saupe, *Chaos and Fractals: New Frontiers in Science*, Springer Verlag, New York, 1992.



*integer dimensions.* The Peano curve is constructed as follows: At the zeroth stage we begin with a line of unit length. At stage 1, we replace this line by the curve obtained by joining the points  $[0, 0]$ ,  $[1/3^-, 0]$ ,  $[1/3^-, 1/3^-]$ ,  $(2/3^-, 1/3^-)$ ,  $(2/3^-, 0)$ ,  $(1/3^+, 0)$ ,  $(1/3^+, -1/3^+)$ ,  $[2/3^+, -1/3^+]$ ,  $[2/3^+, 0]$ ,  $[1, 0]$  successively as shown in *Figure 8*.

The generator of this curve consists of 9 line segments of length  $3^{-1}$ . At the next stage we replace each of these line segments by the generator scaled by a factor of  $3^{-1}$  to get a curve consisting of  $9^2$  line segments of length  $3^{-2}$ . This process is continued indefinitely and the limiting curve is called the Peano curve. It is clear that if the Peano curve is scaled by a factor of  $3^{-1}$  then 9 copies of the scaled curve will be required to reproduce the Peano curve. So the self-similarity dimension of the Peano curve is

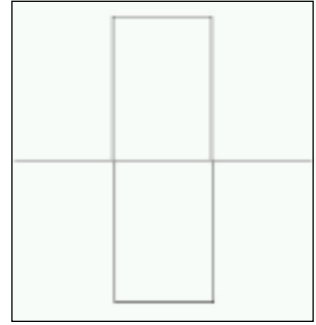
$$D_s = \ln(9)/\ln(3) = 2. \quad (9)$$

The box counting dimension is a widely used dimension in the empirical studies of fractal properties. The reason for this is that it is readily applicable and easily computable.

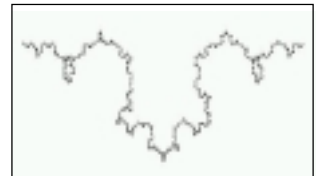
### Random Fractals

So far we have considered what are called deterministic fractals. Such fractals do not occur in Nature. Nature presents what are called statistical fractals. An example of a statistical fractal is obtained if we modify the construction of the Koch curve. When we replace a line segment by the generator shown in *Figure 2* we let this generator lie on either side of the line with equal probability. A stage 5 construction of this curve is shown in *Figure 9*.

This curve looks more like a naturally occurring object, e.g. a coastline, than the corresponding deterministic curve shown in *Figure 6*.



**Figure 8. Generator of a fractal of integer dimension: the Peano curve.**



**Figure 9. A random fractal.**

*Address for Correspondence*  
A K Mittal and T R Seshadri  
Harish-Chandra Research  
Institute  
Chhatnag Road, Jhusi  
Allahabad 211 019, India.

