Dynamical systems, information and time series

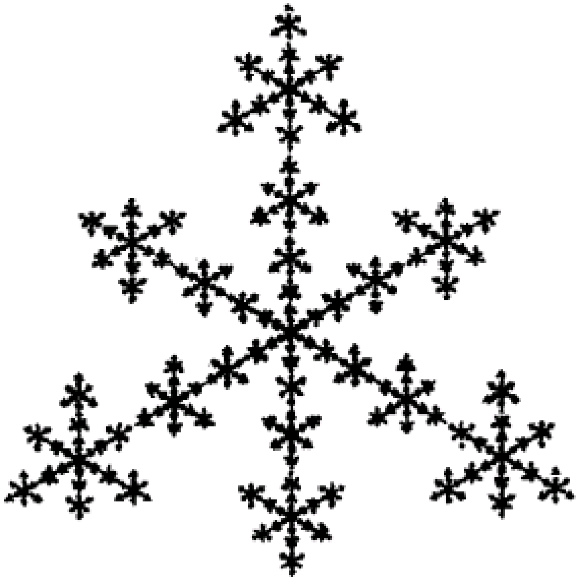
Stefano Marmi
Scuola Normale Superiore
http://homepage.sns.it/marmi/
Lecture 5- European University Institute
November 6, 2009

- Lecture 1: An introduction to dynamical systems and to time series. Periodic and quasiperiodic motions. (Sept 18)
- Lecture 2: A priori probability vs. statistics: ergodicity, uniform distribution of orbits. The analysis of return times. Kac inequality. Mixing (Sep 25)
- Lecture 3: Shannon and Kolmogorov-Sinai entropy. Randomness and deterministic chaos. Relative entropy and Kelly's betting. (Oct 9)
- Lecture 4: Time series analysis and embedology: can we distinguish deterministic chaos in a noisy environment? (Tuesday, Oct 27, 11am-1pm)
- Lecture 5: Fractals and multifractals. (Nov 6, 3pm-5pm)

Self-similarity and fractals

A subset **A** of Euclidean space will be considered a “fractal” when it has most of the following features:

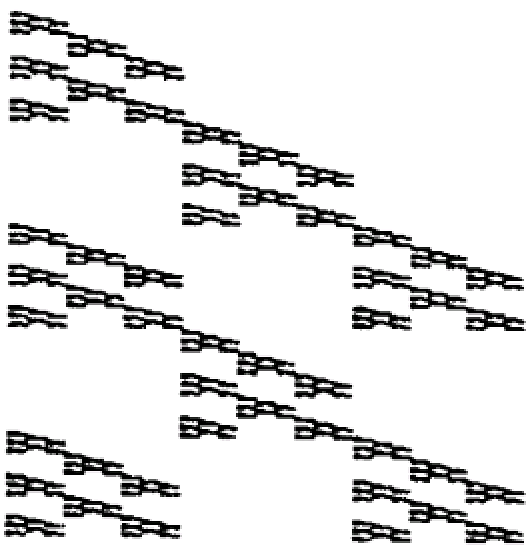
- **A** has fine structure (wiggly detail at arbitrarily small scales)
- **A** is too irregular to be described by calculus (e.g. no tangent space)
- **A** is self-similar or self-affine (maybe approximately or statistically)
- the fractal dimension of **A** is non-integer
- **A** may have a simple (recursive) definition
- **A** has a “natural” appearance: “Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in a straight line . . .” (B. Mandelbrot)



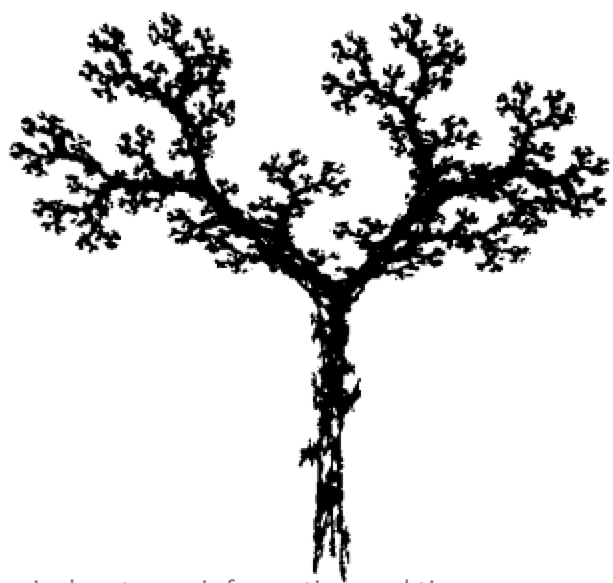
(a)

(b)

self-similar fractals



(c)



(d)

self-affine fractals

From: K. Falconer, Techniques in Fractal Geometry, Wiley 1997

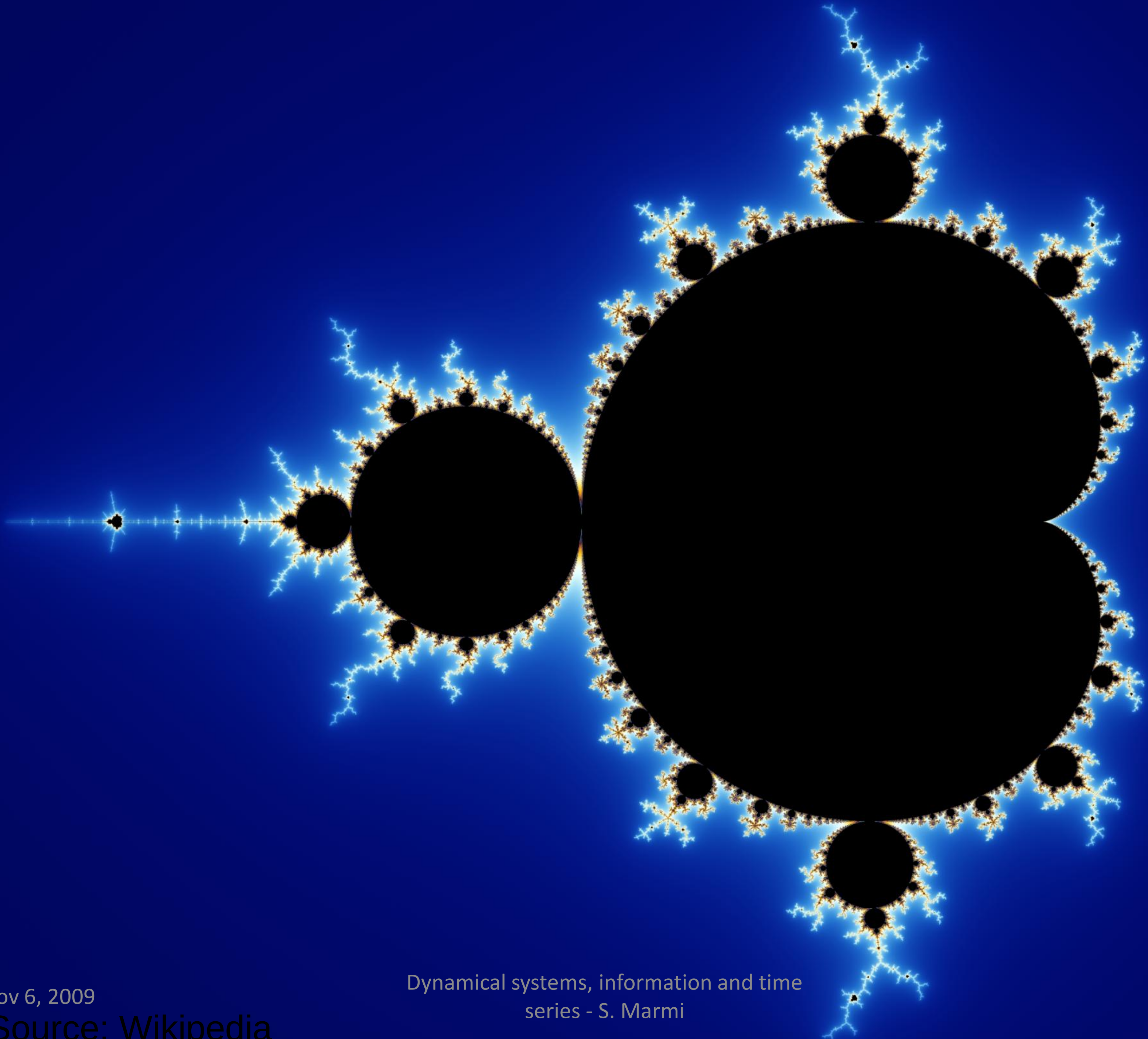
(e)

(f)

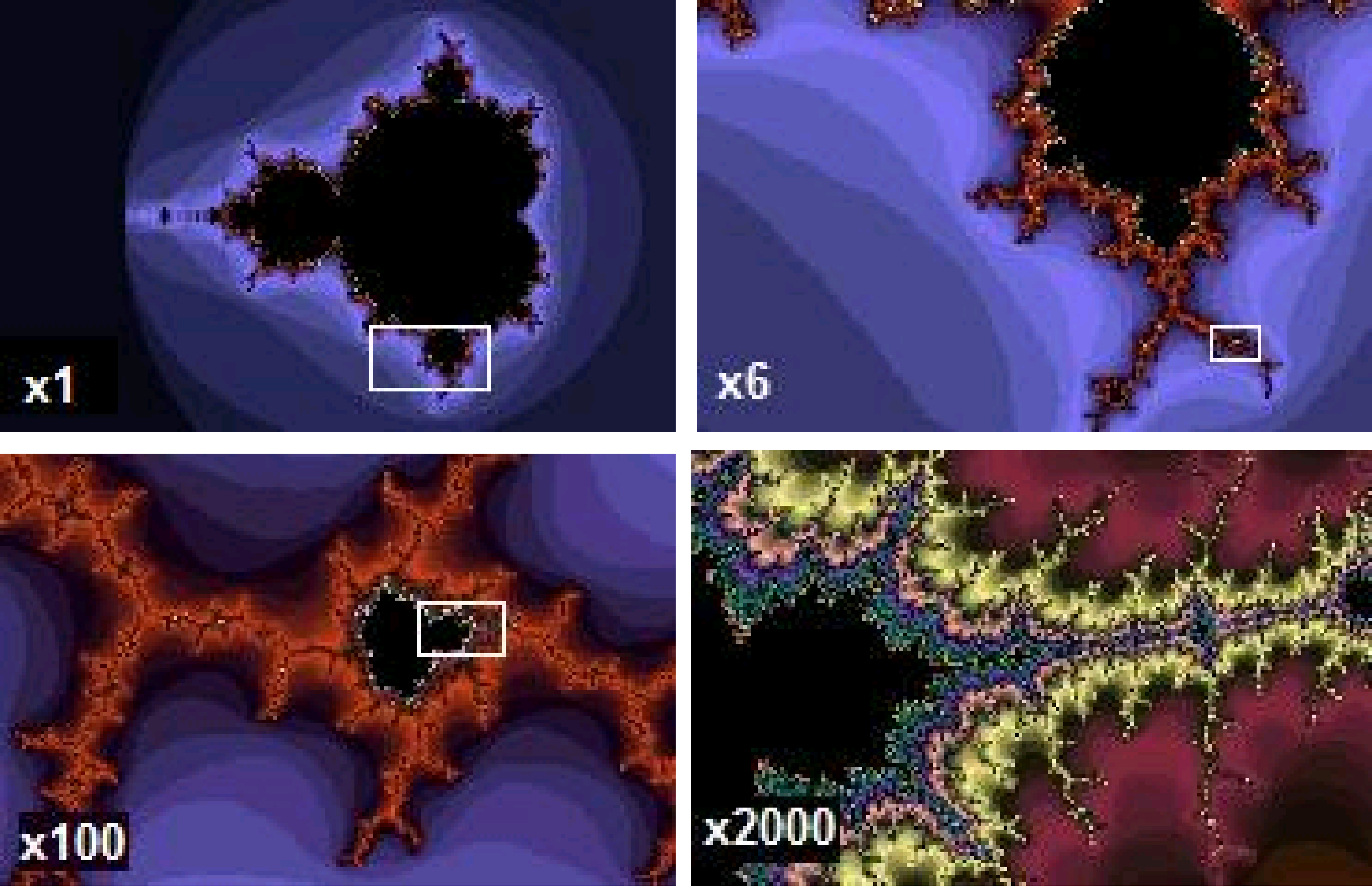
self-conformal
fractals

Statistically
self-similar
fractals

From: K. Falconer, Techniques in Fractal Geometry, Wiley 1997

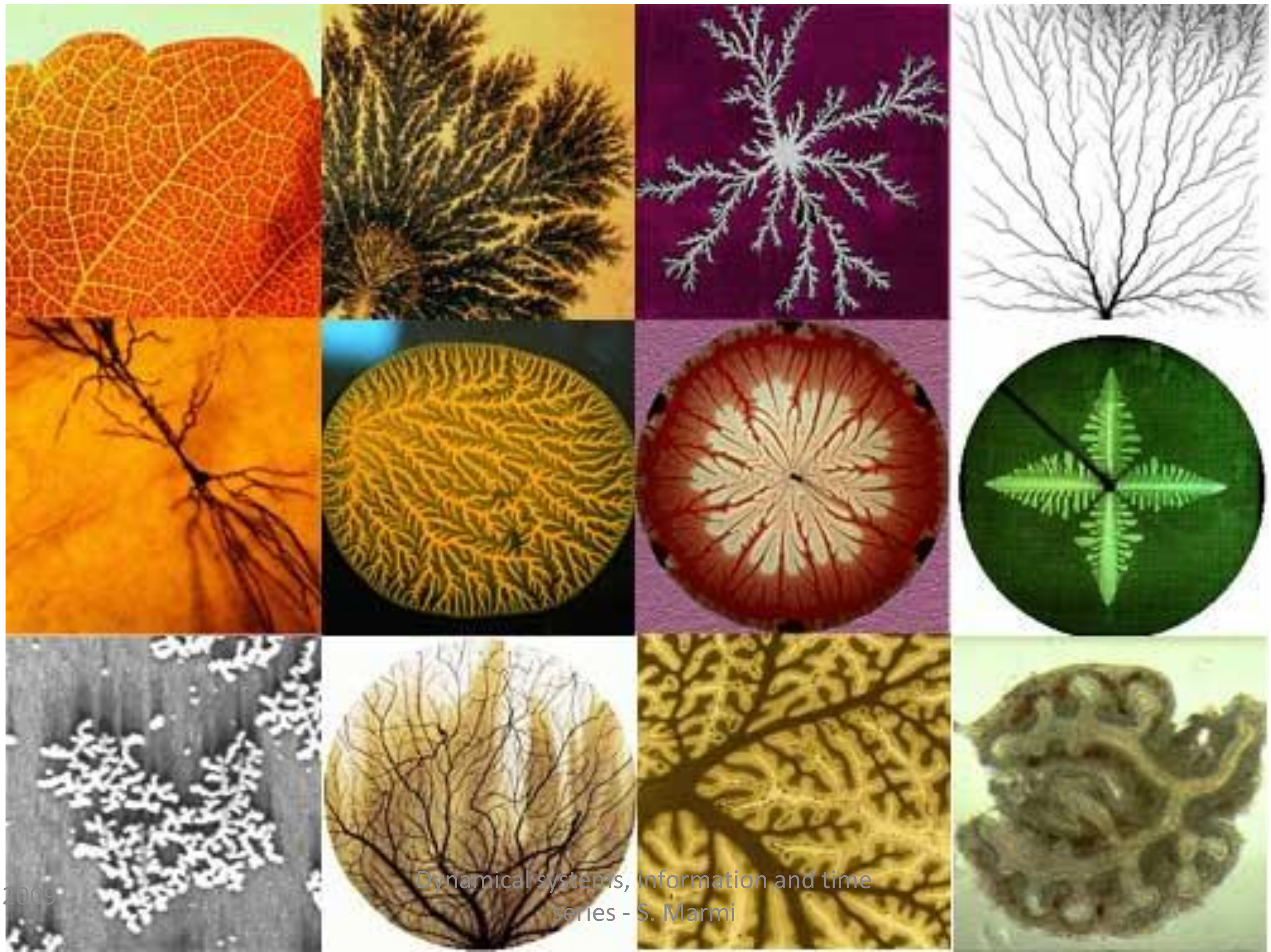


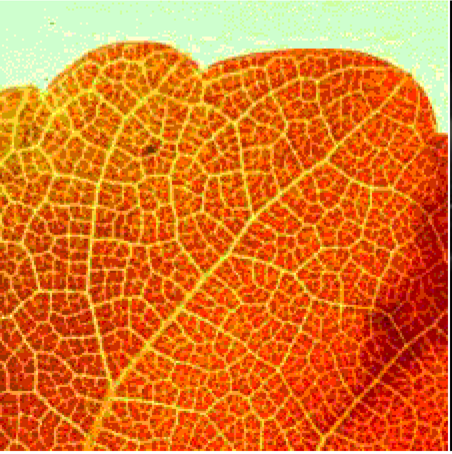
Source: Wikipedia



Source: Wikipedia

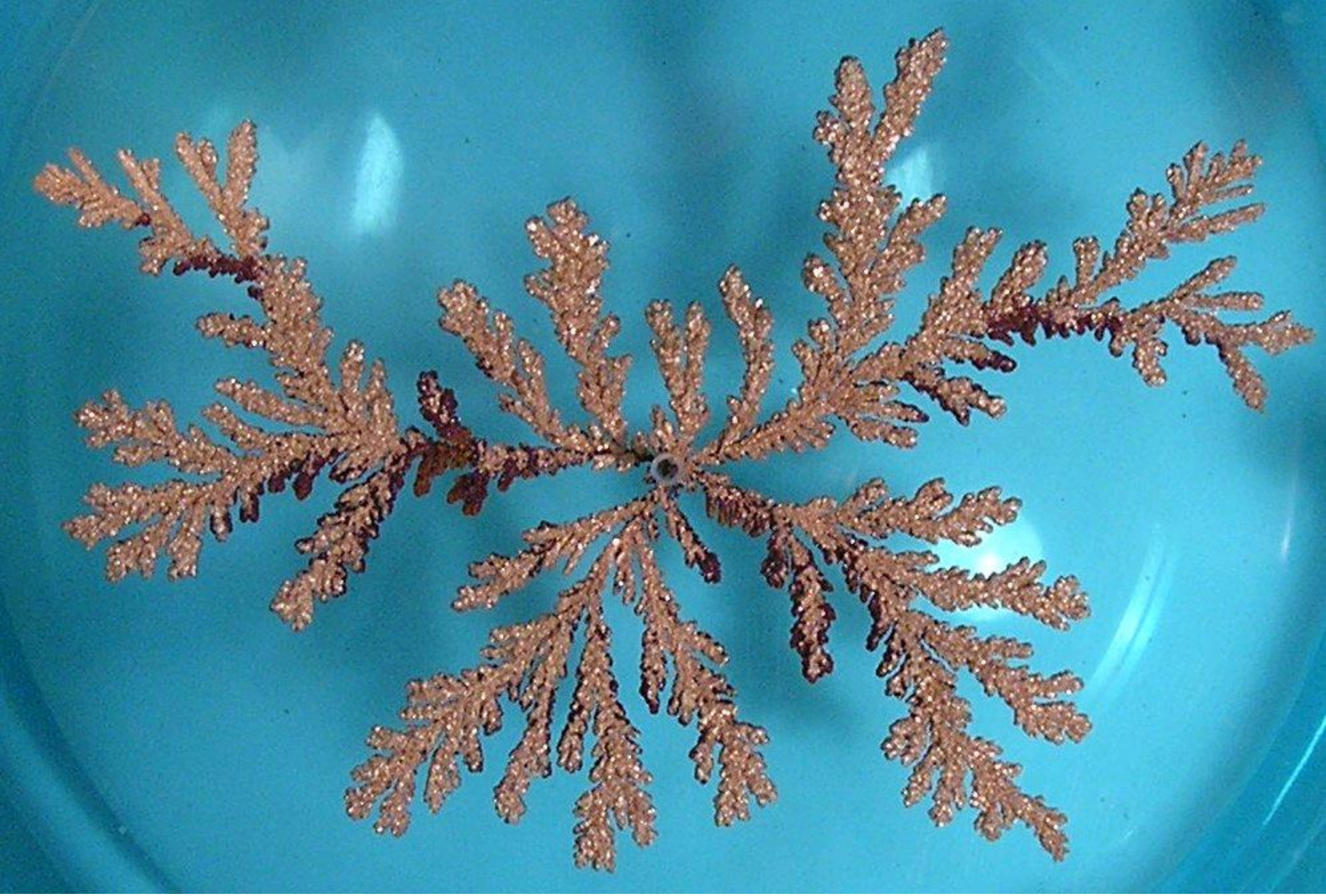
Mathematics, shapes and nature





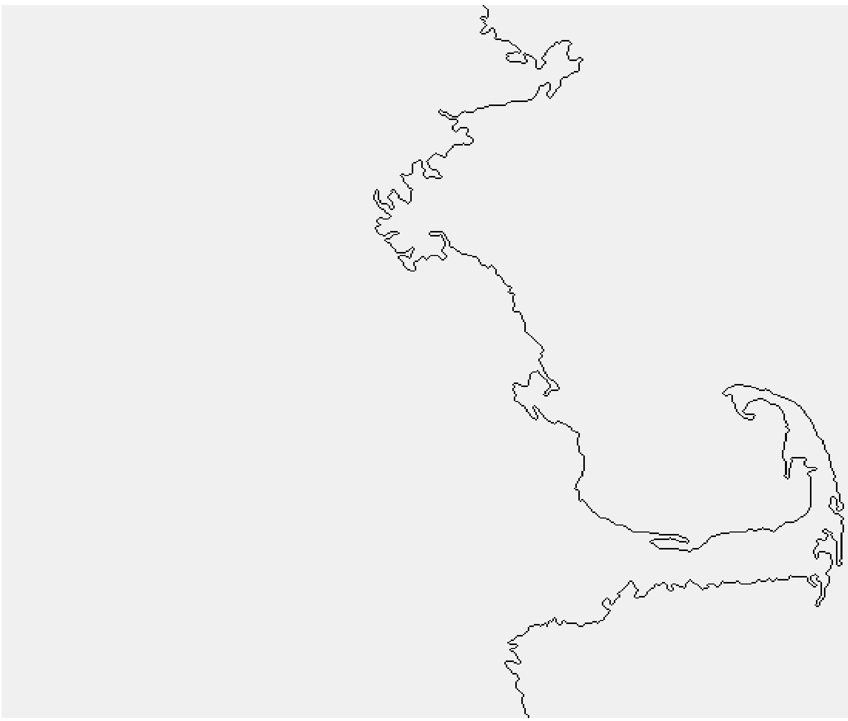
http://classes.yale.edu/fractals/Panorama/Nature/NatFracGallery/Gallery/Gallery.html

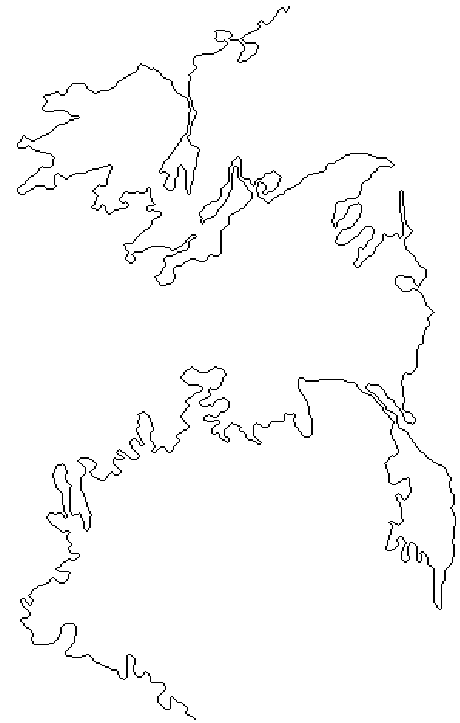
From http://en.wikipedia.org/wiki/Image:Square1.jpg
Lichtenberg Figure
High voltage dielectric breakdown within a block of plexiglas creates a beautiful fractal pattern called a Lichtenberg_figure. The branching discharges ultimately become hairlike, but are thought to extend down to the molecular level.
Bert Hickman, http://www.teslamania.com



A diffusion-limited aggregation (DLA) cluster. Copper aggregate formed from a copper sulfate solution in an electrode position cell. Kevin R. Johnson, Wikipedia

Coastlines





Massachusetts

D=1.15

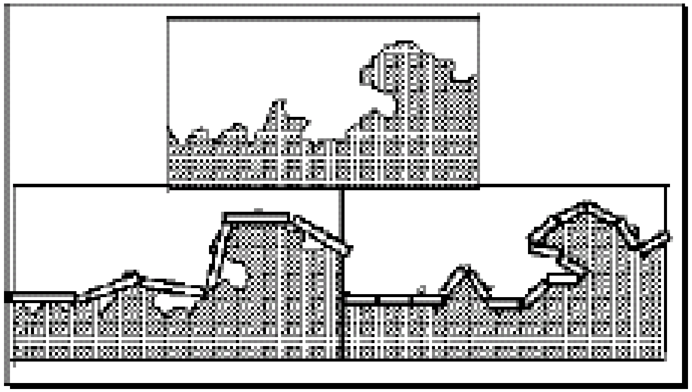
Greece

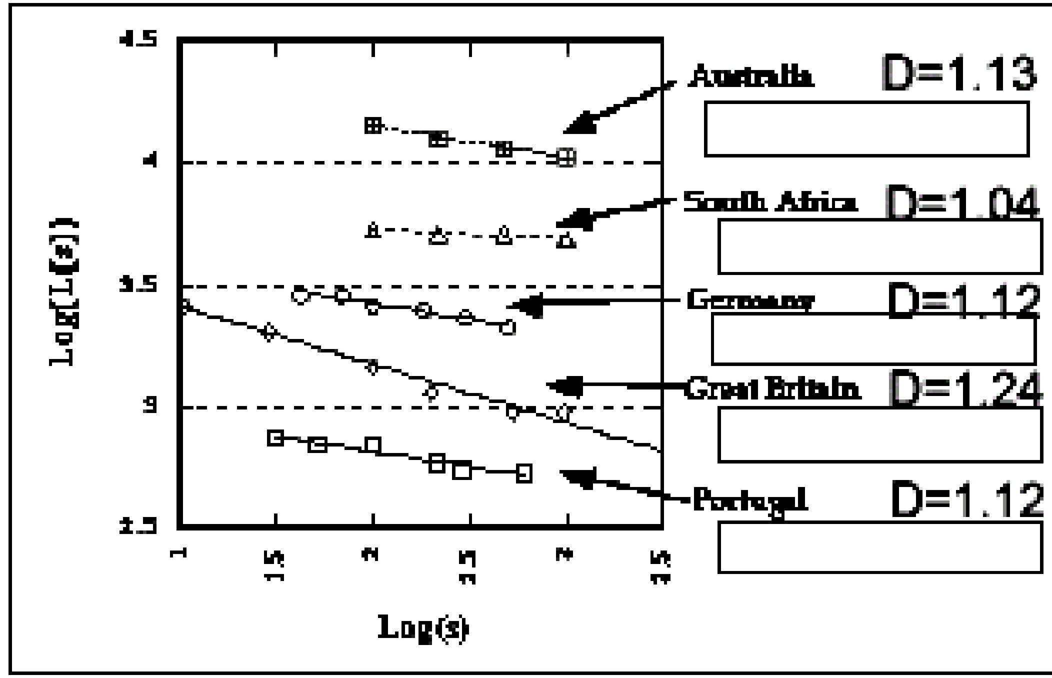
D=1.20

200 km 100 km 50 km

http://upload.wikimedia.org/wikipedia/commons/2/20/Britain-fractal-coastline-combined.jpg

How long is a coastline?





The answer depends on the scale at which the measurement is made: if s is the reference length the coastline length L(s) will be

Log L(s) = (1-D) log s + cost

(Richardson 1961, Mandelbrot Science 1967)

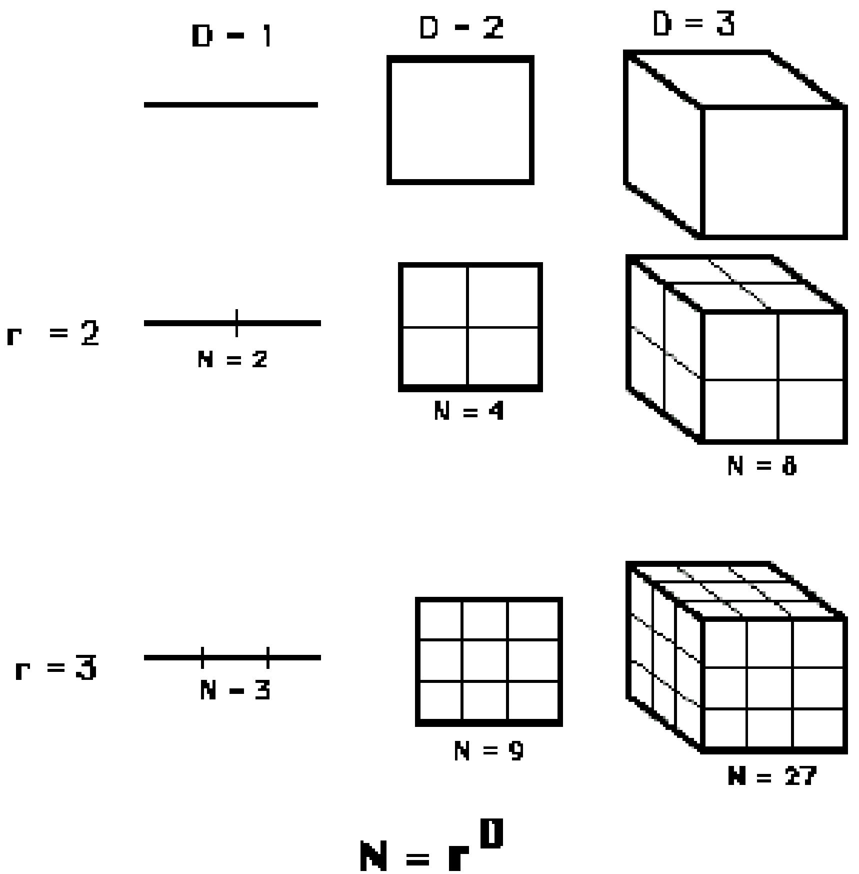
How long is the coast of Britain?
Statistical self-similarity and fractional dimension
Science: 156, 1967, 636-638
B. B. Mandelbrot

Seacoast shapes are examples of highly involved curves with the property that - in a statistical sense - each portion can be considered a reduced-scale image of the whole. This property will be referred to as "statistical self-similarity." The concept of "length" is usually meaningless for geographical curves. They can be considered superpositions of features of widely scattered characteristic sizes; as even finer features are taken into account, the total measured length increases, and there is usually no clear-cut gap or crossover, between the realm of geography and details with which geography need not be concerned.

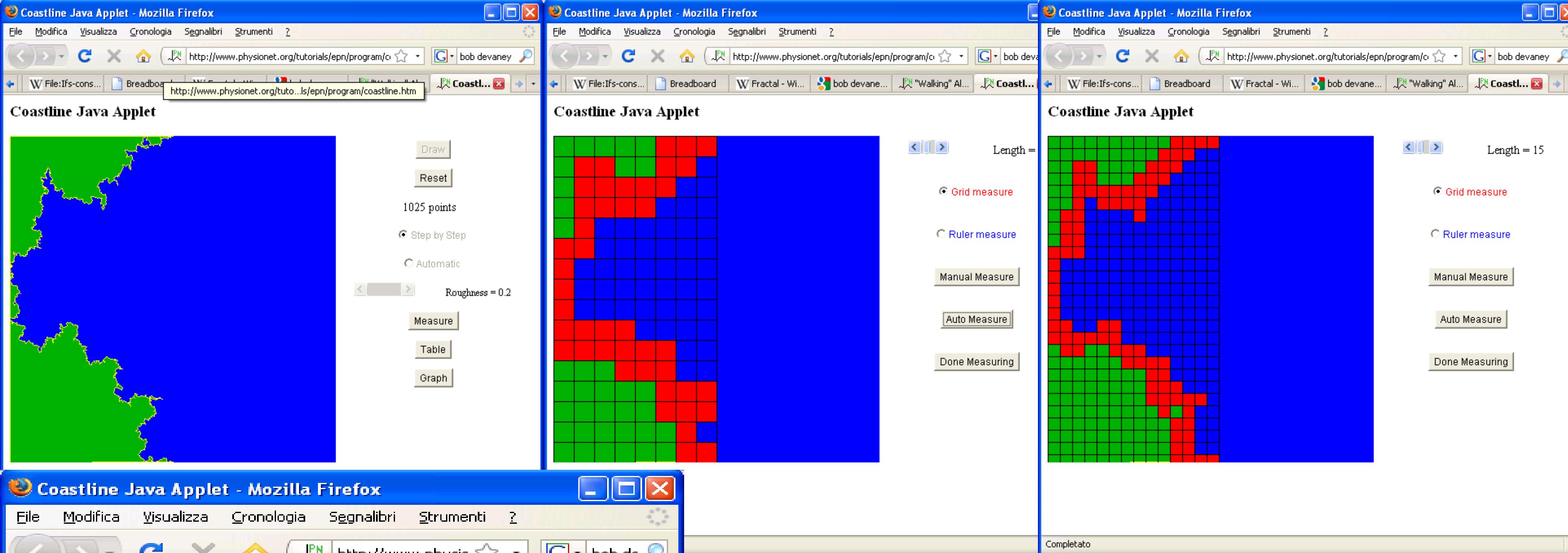
How long is the coast of Britain? Statistical self-similarity and fractional dimension Science: 156, 1967, 636-638 B. B. Mandelbrot

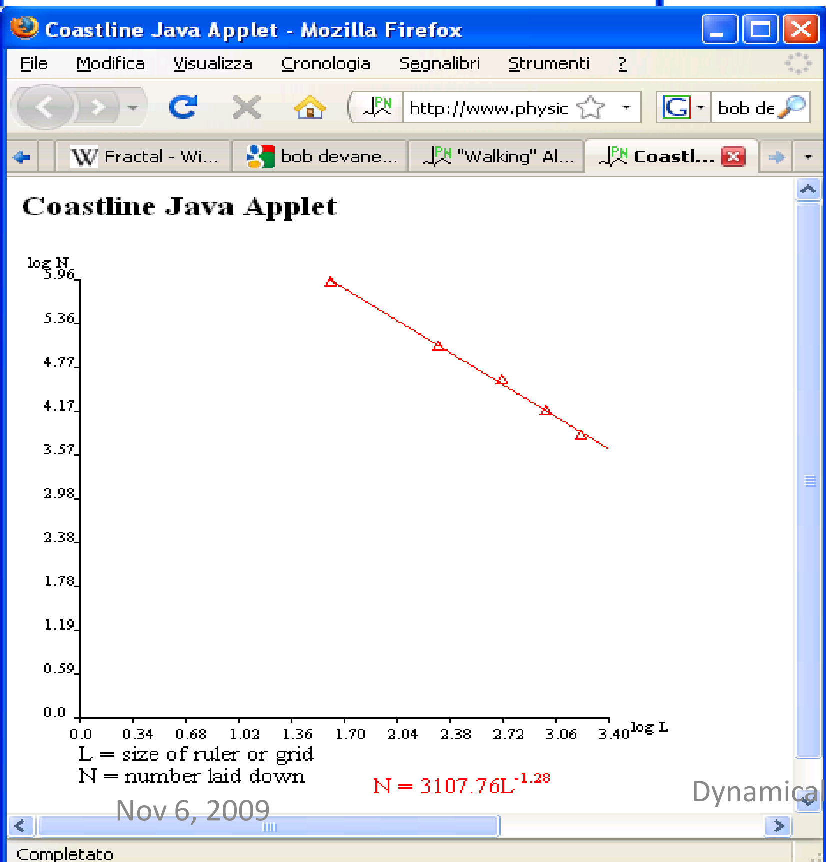
Quantities other than length are therefore needed to discriminate between various degrees of complication for a geographical curve. When a curve is self-similar, it is characterized by an exponent of similarity, D, which possesses many properties of a dimension, though it is usually a fraction greater that the dimension 1 commonly attributed to curves. I propose to reexamine in this light, some empirical observations in Richardson 1961 and interpret them as implying, for example, that the dimension of the west coast of Great Britain is D = 1.25. Thus, the so far esoteric concept of a "random figure of fractional dimension" is shown to have simple and concrete applications of great usefulness.

“Box counting” dimension



$$D = \lim_{s \to 0} \frac{\log N(s)}{\log(1/s)}$$





s	N(s)
25	47
20	67
15	100
10	159
5	386

$$\text{Log } N(s) = -D \log s + \text{cost}$$

http://www.physionet.org/tutorials/epn/program/coastline.htm

Box counting (Minkowski) dimension

Let E be a non-empty bounded subset of $\mathbf{R}^n$ and let $N_r(E)$ be the smallest number of sets of diameter r needed to cover E

- Lower dimension $\dim_B E = \liminf_{r\to 0} \log N_r(E) / -\log r$
- Upper dimension $\dim^B E = \limsup_{r\to 0} \log N_r(E) / -\log r$

- Box-counting dimension: if the lower and upper dimension agree then we define

$$\dim E = \lim_{r\to 0} \log N_r(E) / -\log r$$

The value of these limits remains unaltered if $N_r(E)$ is taken to be the smallest number of balls of radius r (cubes of side r) needed to cover E, or the number of r-mesh cubes that intersect E

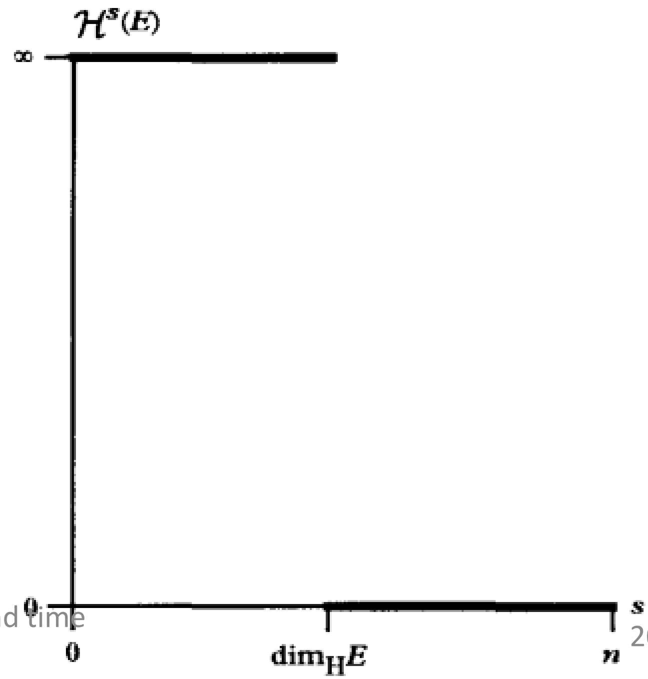
Hausdorff dimension

A finite or countable collection of subsets $\{U_i\}$ of $\mathbf{R}^n$ is a δ-cover of a set E if $|U_i| < \delta$ for all i and E is contained in $\bigcup_i U_i$

$H^s_\delta(E) = \inf\{\Sigma_i |U_i|^s, \{U_i\}$ is a a δ-cover of E$\}$

s-dimensional Hausdorff measure of E: $H^s(E) = \lim_{\delta\to 0} H^s_\delta(E)$

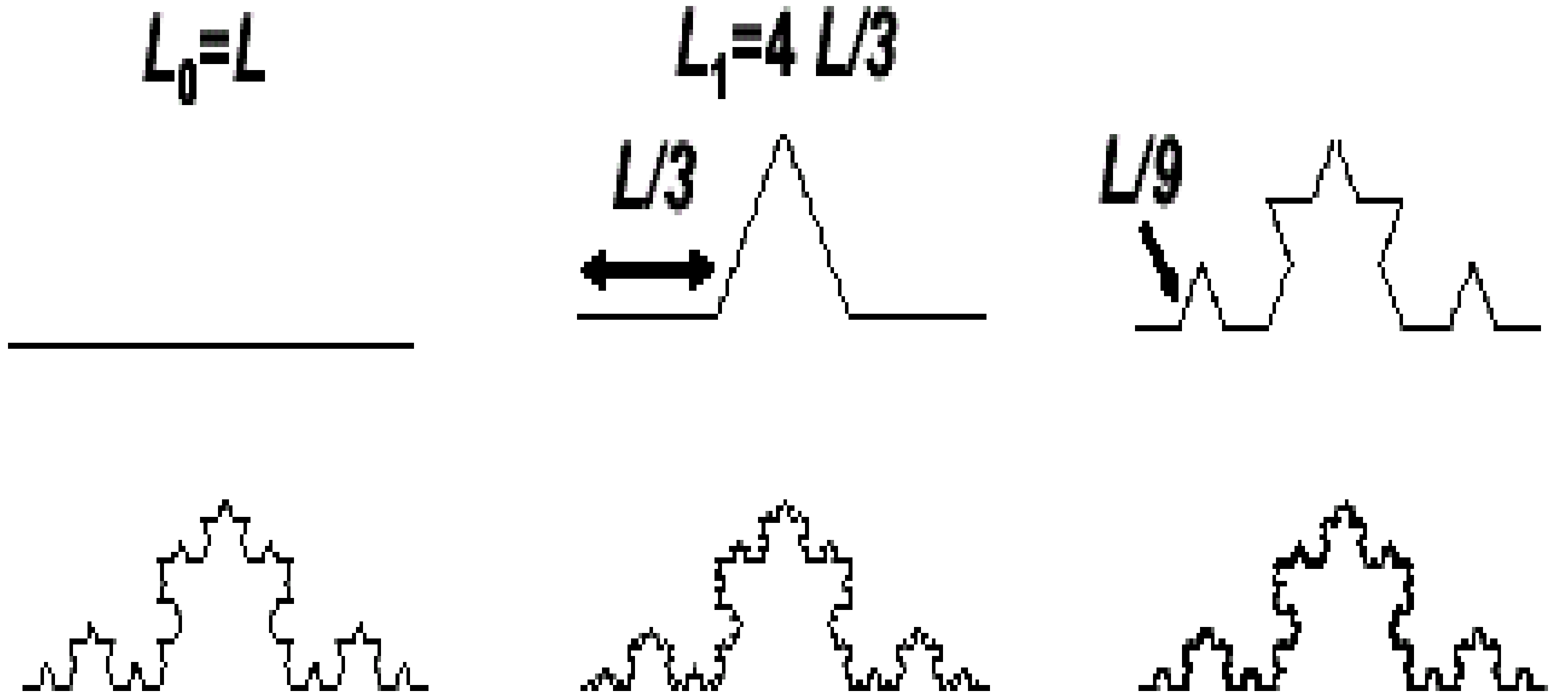
It is a Borel regular measure on $\mathbf{R}^n$, it behaves well under similarities and Lipschitz maps



The Hausdorff dimension $\dim_H E$ is the number at which the Hausdorff measure $H^s(E)$ jumps from ∞ to 0

$\dim_H E \le \dim_B E \le \dim^B E$

Von Koch curve (1904)



D=log4/log3=1.261859…

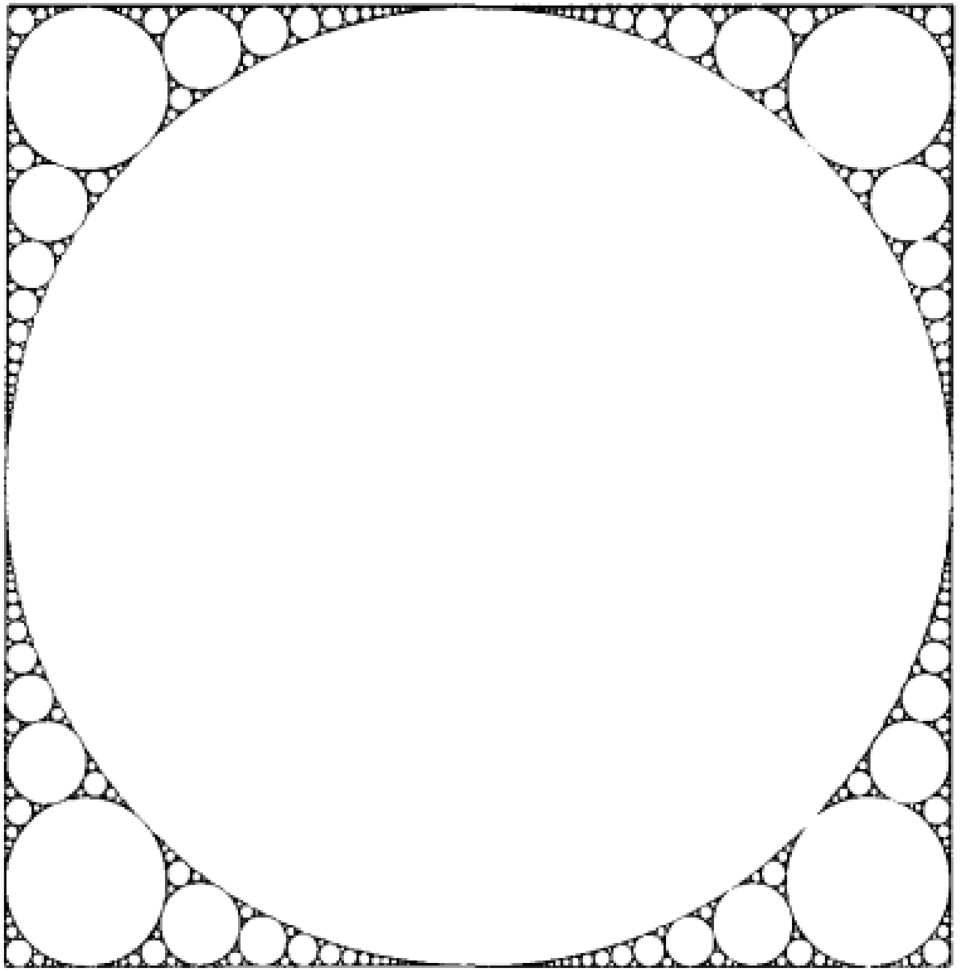
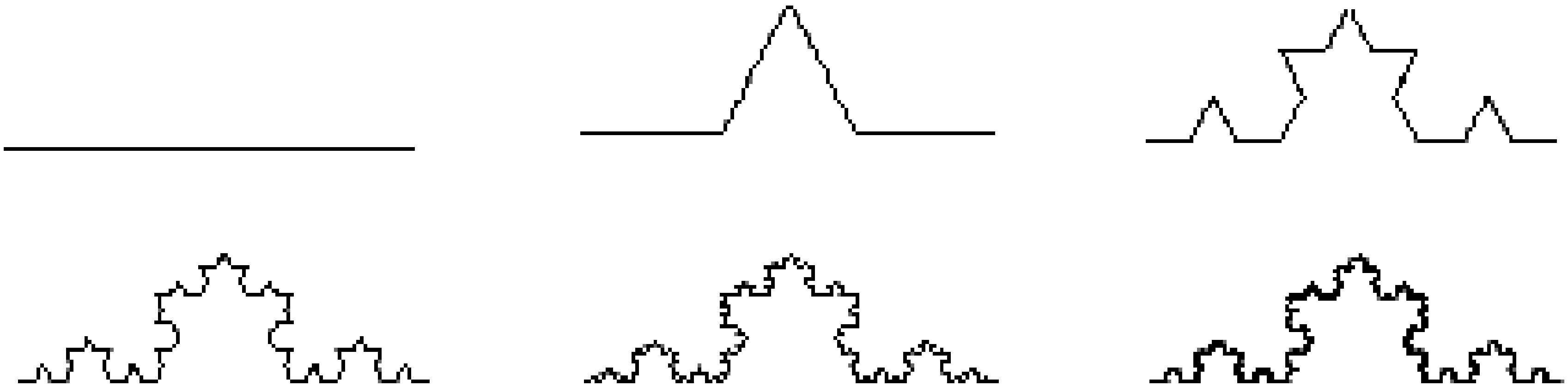


Figure 3.6 A cut-out set in the plane. Here, the largest possible disc is removed at each step. The family of discs removed is called the Apollonian packing of the square, and the cut-out set remaining is called the residual set, which has Hausdorff and box dimension about 1.31

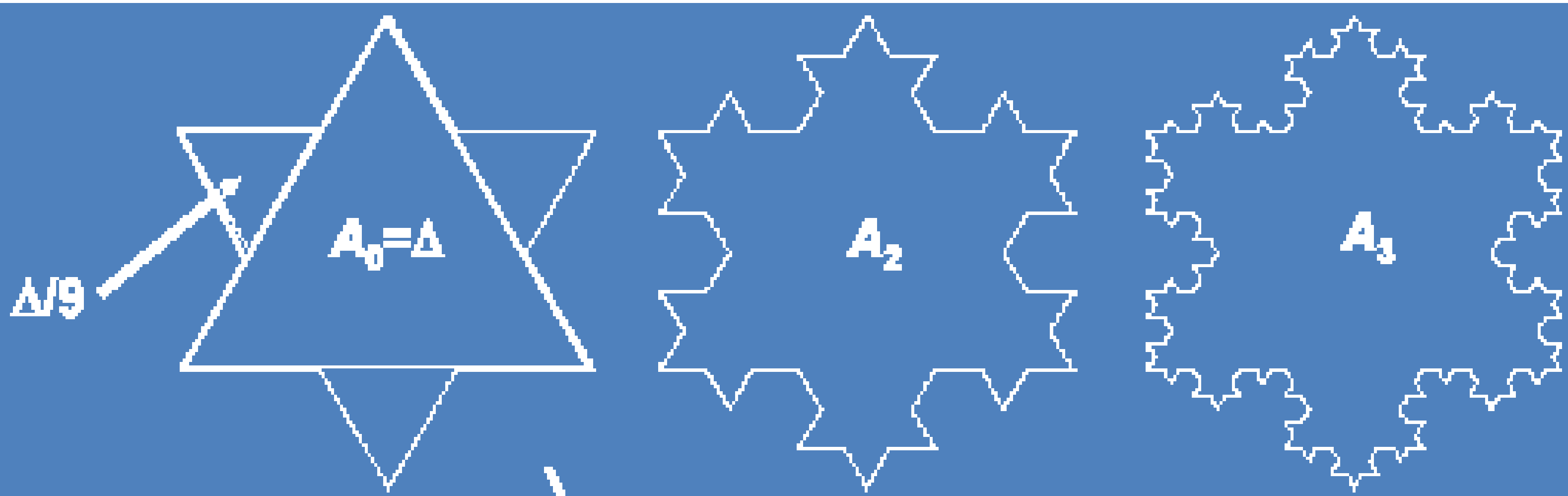
From: K. Falconer, Techniques in Fractal Geometry, Wiley 1997



$$L_0 = 1, \quad L_1 = 4/3, \quad L_2 = 4^2/3^2, \ \text{etc...} \ L_k \to \infty$$

$$s = 1/3^k, \ N(s) = 4^k \to D = \frac{\log 4^k}{\log 3^k} = \frac{\log 4}{\log 3}$$

Fractal snowflake



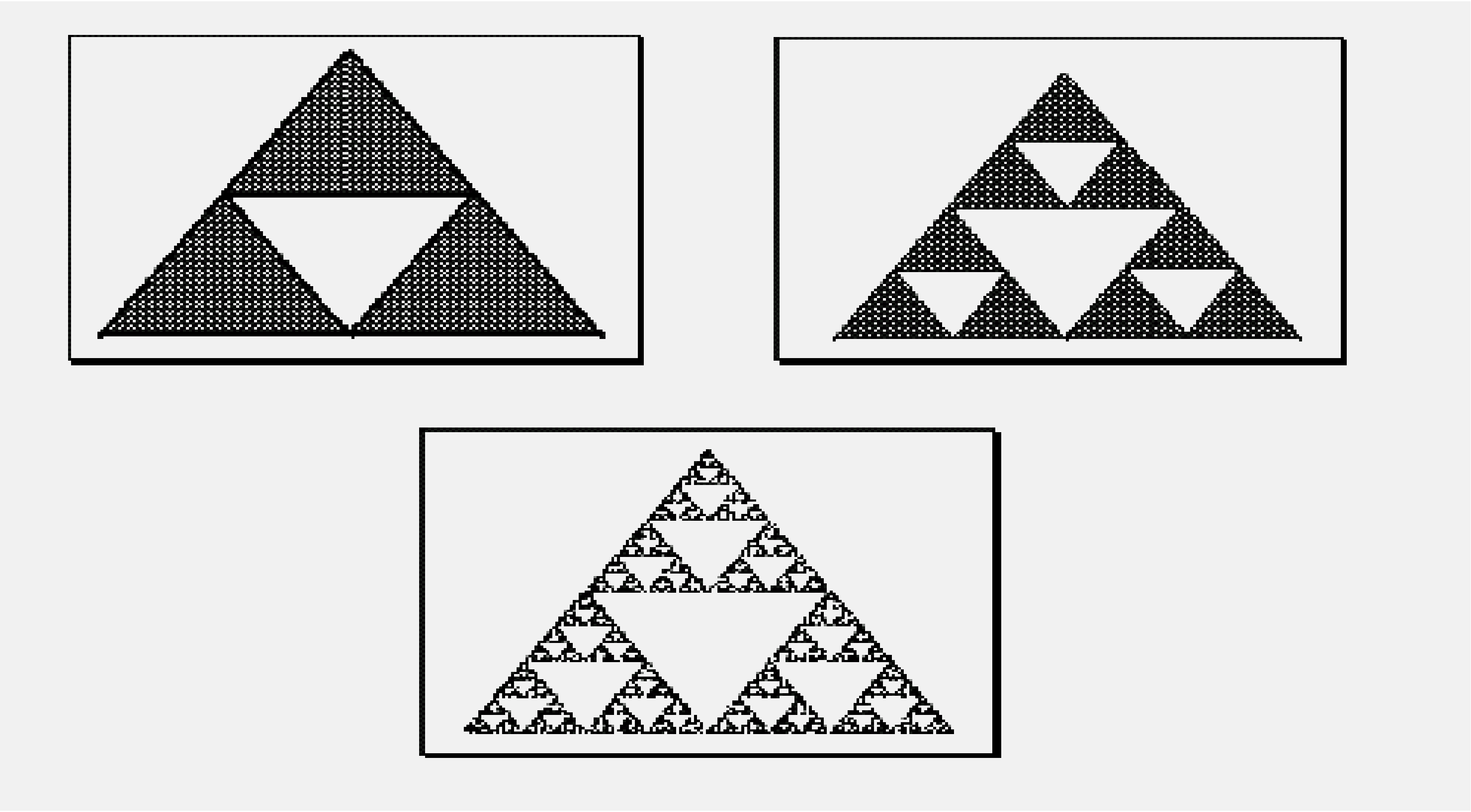
Area calculation:

$$A_0 = \Delta, \qquad A_1 = \Delta + \frac{3}{9}\Delta = \Delta\left(1+\frac{1}{3}\right), \qquad A_2 = \Delta + \frac{1}{3}\Delta + \frac{12}{81}\Delta = \Delta + \frac{1}{3}\Delta + \frac{1}{3}\frac{4}{9}\Delta$$

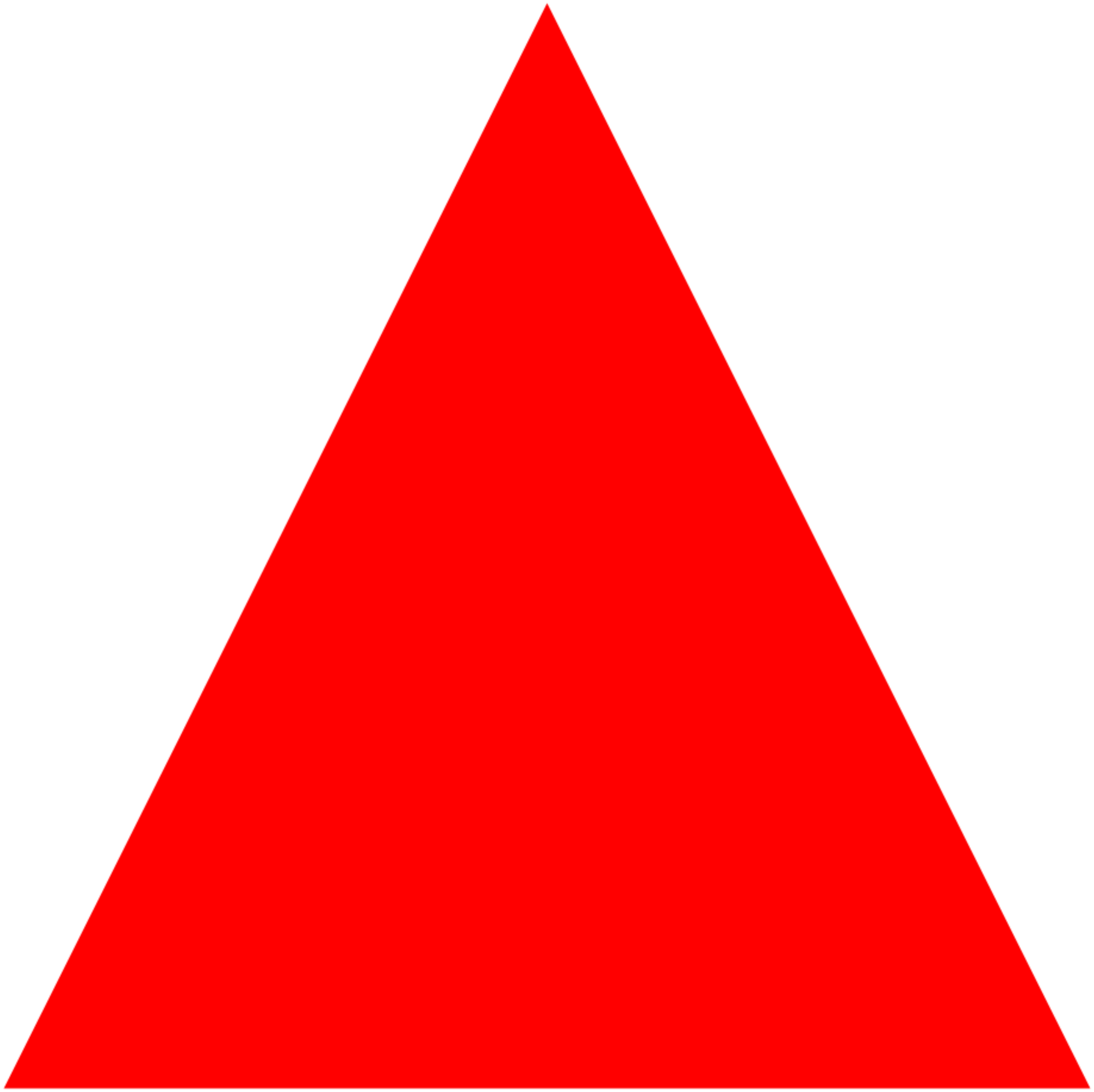
$$A_3 = \Delta + \frac{1}{3}\Delta + \frac{1}{3}\frac{4}{9}\Delta + \frac{1}{3}\frac{4}{9}\frac{4}{9}\Delta, \quad \text{etc...}$$

Infinite perimeter, finite area, D=log4/log3=1.261859...

Sierpinski triangle (1916)

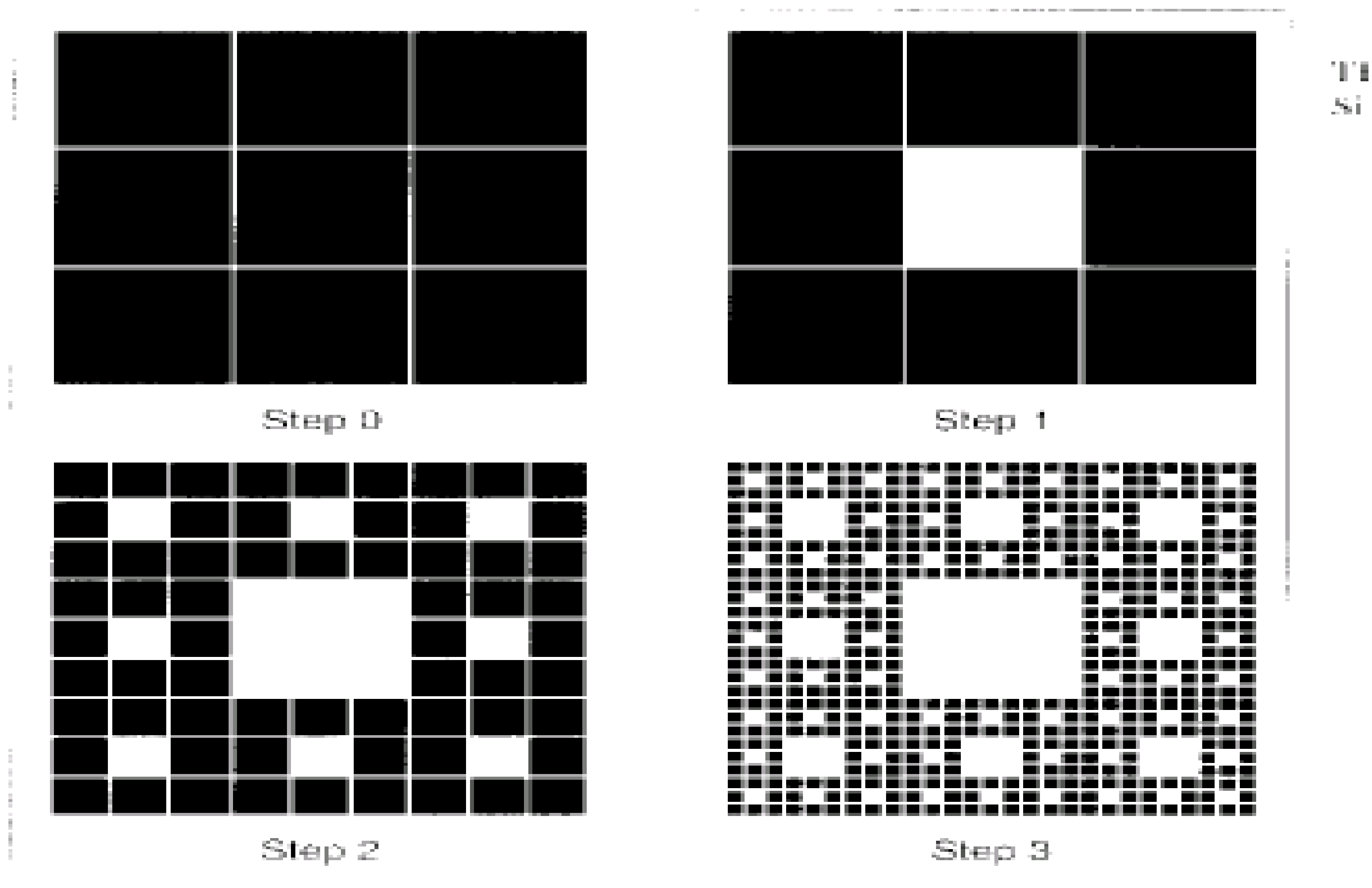


D=log3/log2=1.5849625...



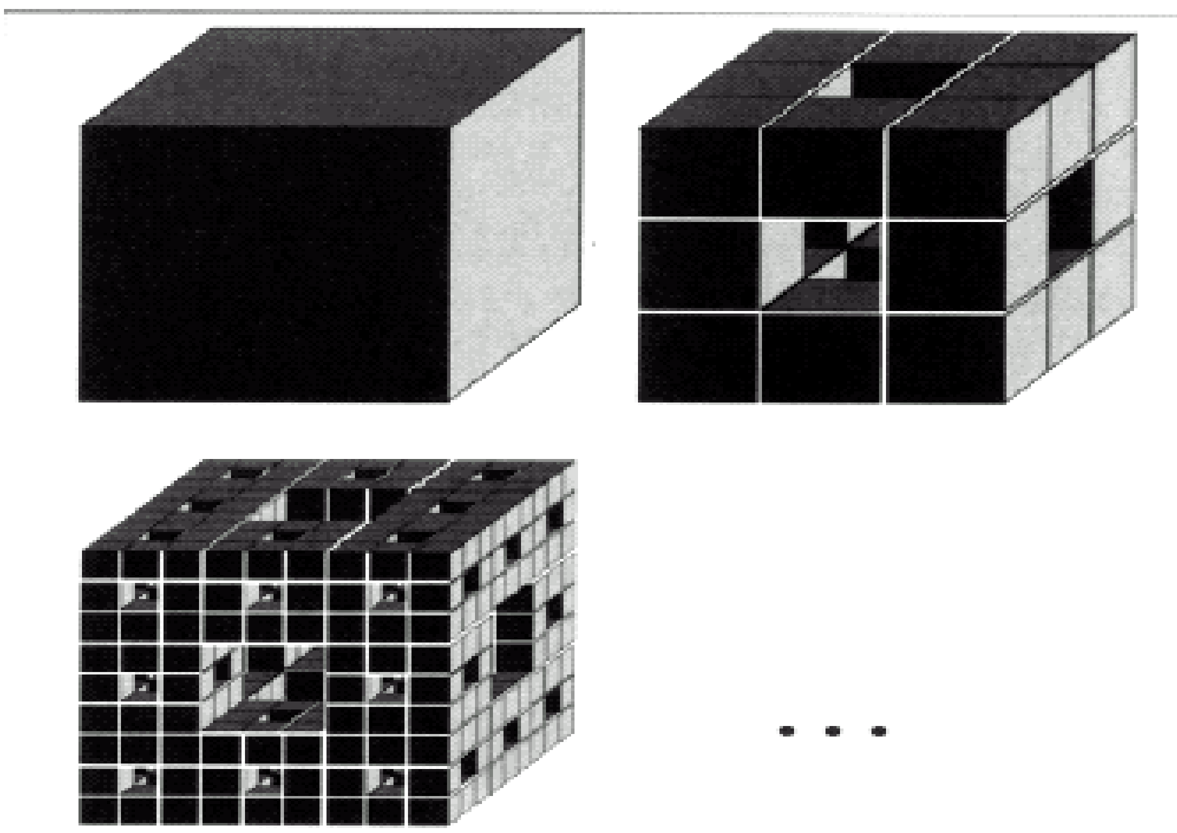
Source: Wikipedia

A fractal carpet (zero area)

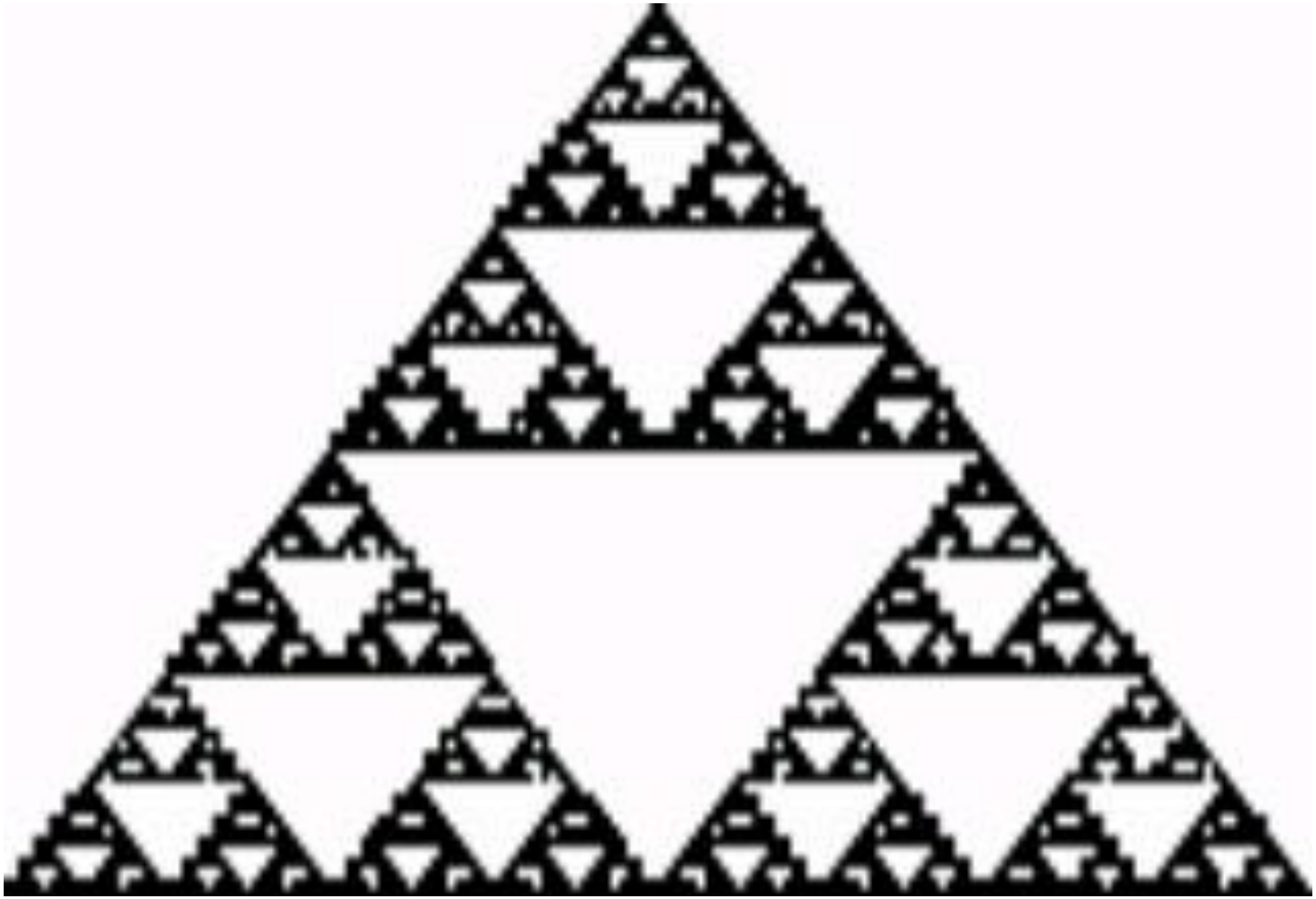


D=3log2/log3=1.892789...

A fractal sponge



Zooming in

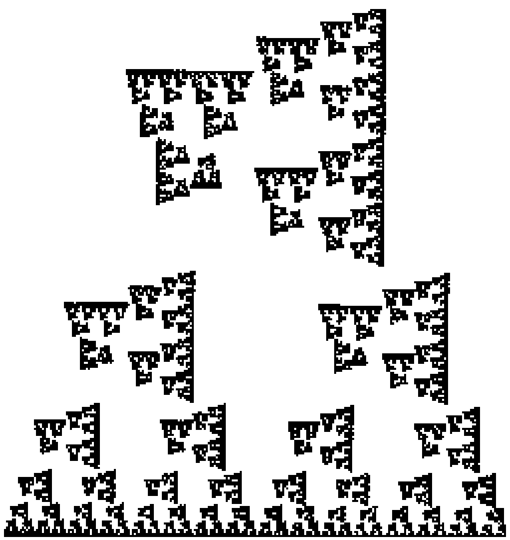


Changing parameters

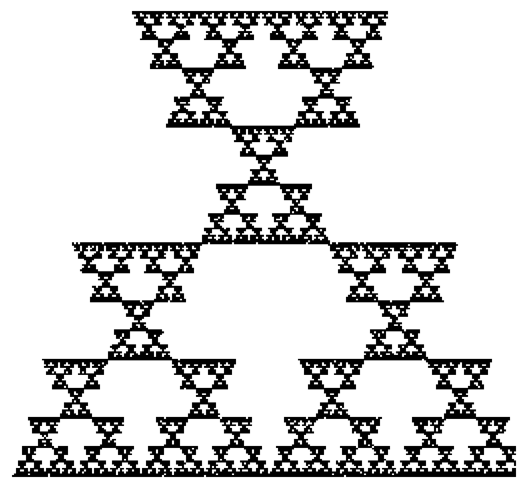
•The triangle of Sierpinski is the attractor of an iterated function system (i.f.s).

•The i.f.s. is made of three affine maps (each contracting by a factor ½ and leaving one of the initial vertices fixed)

•Combining the affine maps with rotations one can change the shape considerably



90 anticlockwise rotation about the top vertex



180° rotation about the same vertex

Hausdorff metric and compact sets

$X=[0,1]^2$

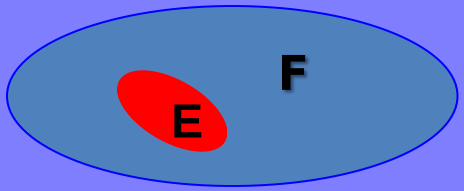
$d((x,y),(x',y'))=|x-x'|+|y-y'|$ Manhattan metric

$\mathscr{H}(X)=\{E$ compact nonempty subsets of $X\}$

$h(E,F)=\max(d(E,F),d(F,E))$

$d(E,F)=\max_{x\in E}\min_{y\in F} d(x,y)$ $d(E,F)\neq d(F,E)$

$d(E,F)>0$



$d(F,E)=0$

Theorem: $(\mathscr{H}(X),h)$ is a complete metric space

→ Cauchy sequences have a limit!

Contractions and Hausdorff metric

Proposition: if $w:X \to X$ is a contraction with Lipschitz constant s then w is also a contraction on $(\mathcal{H}(X),h)$ with Lipschitz constant s

To each family $\mathcal{F}$ of contractions on X one can associate a family of contractions on $(\mathcal{H}(X),h)$. By Banach-Caccioppoli to each such $\mathcal{F}$ will correspond a compact nonempty subset $\mathcal{A}$ of X: the attractor associated to $\mathcal{F}$

$$d(w(E),w(F)) = \max_{y \in E} \min_{z \in F} d(y,z) = \max_{e \in E} \min_{f \in F} d(w(e),w(f))$$

$$\leq s \max_{e \in E} \min_{f \in F} d(e,f) = s\, d(E,F)$$

Iterated function systems

$\mathcal{F} = \{w_1, \ldots, w_N\}$ each $w_i : X \to X$ is a contraction of constant s_i, $0 \leq s_i < 1$

Let $\mathcal{W} : \mathcal{H}(X) \to X$

$$\mathcal{W}(E) = \bigcup_{1 \leq i \leq N} w_i(E)$$

Then $\mathcal{W}$ contracts the Hausdorff metric h with Lipschitz constant $s = \max_{1 \leq i \leq N} s_i$. We denote by $\mathcal{A}$ the corresponding attractor

Given any subset E of X, the iterates $\mathcal{W}^n(E) \to \mathcal{A}$ exponentially fast, in fact $h(\mathcal{W}^n(E), \mathcal{A}) \approx s^n$ as $n \to \infty$

Self similarity and fractal dimension

If the contractions of the i.f.s. $\mathscr{F} = \{w_1, \ldots, w_N\}$ are

- Similarities the attractor $\mathscr{A}$ will be said self-similar
- Affine maps the attractor $\mathscr{A}$ will be said self-affine
- Conformal maps (i.e. their derivative is a similarity) then the attractor $\mathscr{A}$ will be said self-conformal

If the open set condition is verified, i.e. there exists an open set U such that $w_i(U) \cap w_j(U) = \emptyset$ if $i \neq j$ and $\cup_i w_i(U)$ is an open subset of U then the dimension d of the attractor $\mathscr{A}$ is the unique positive solution of $s_1^d + s_2^d + \ldots + s_N^d = 1$

Inverse problem

Inverse problem: given $\varepsilon>0$ and a target (fractal) set $\mathscr{T}$ can one find an i.f.s $\mathscr{F}$ such that the corresponding attractor $\mathscr{A}$ is ε-close to $\mathscr{T}$ w.r.t. the Hausdorff distance h?

Collage Theorem (Barnsley 1985) Let $\varepsilon>0$ and let $\mathscr{T}\in\mathscr{H}(X)$ be given. If the i.f.s. $\mathscr{F} = \{w_1, \ldots, w_N\}$ is such that

$$h(\cup_{1\leq i\leq N} w_i(\mathscr{T}) , \mathscr{T}) < \varepsilon$$

then

$$h(\mathscr{T}, \mathscr{A}) < \varepsilon / (1-s)$$

where s is the Lipschitz constant of $\mathscr{F}$

Fractal image compression ?

The Collage Theorem tells us that to find an i.f.s. whose attractor "looks like" a give set one must find a set of contracting maps such that the union (collage) of the images of the given set under these maps is near (w.r.t. Hausdorff metric) to the original set.

The collage theorem sometimes allows incredible compression rates of images (of course with loss). It can be especially useful when the information contained in details is not considered very very important

Fractal image compression !

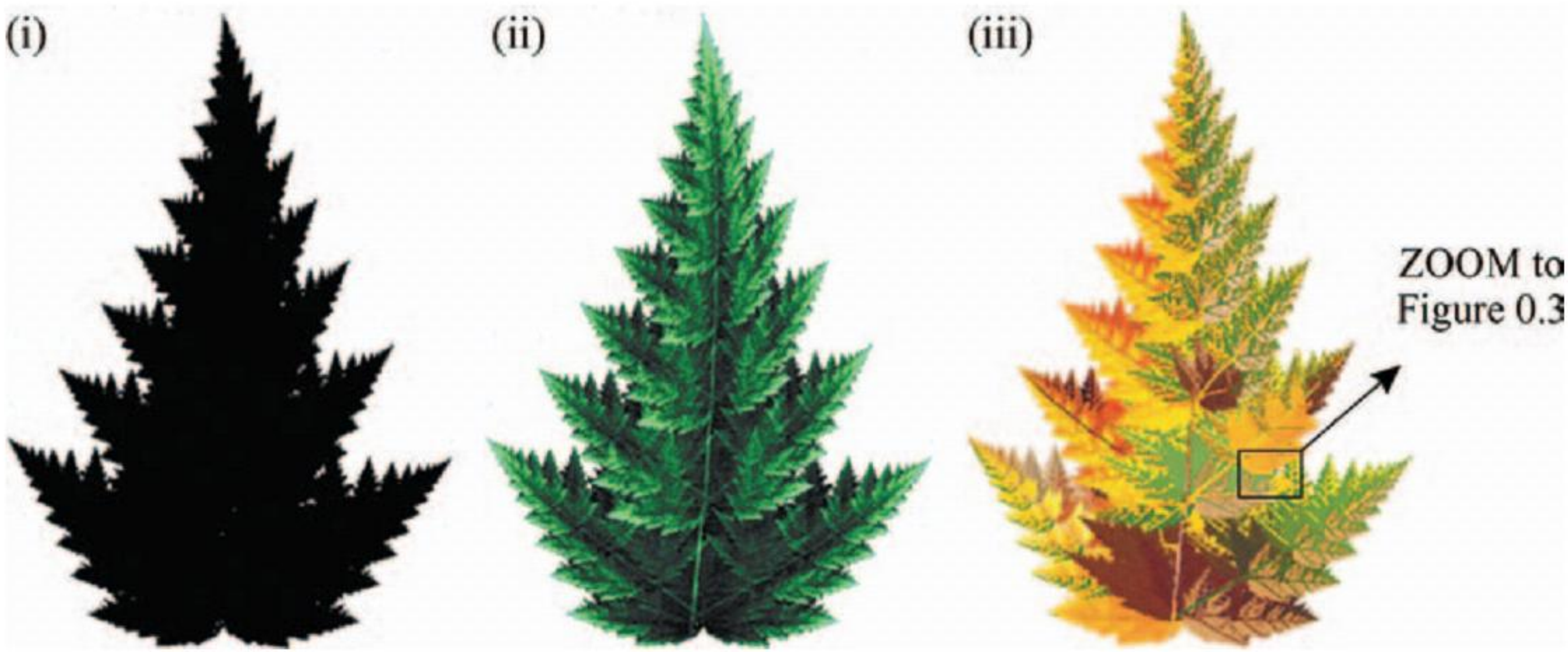
The top-selling multimedia encyclopedia Encarta, published by Microsoft Corporation, includes on one CD-ROM seven thousand color photographs which may be viewed interactively on a computer screen. The images are diverse; they are of buildings, musical instruments, people's faces, baseball bats, ferns, etc. What most users do not know is that all of these photographs are based on fractals and that they represent a (seemingly magical) practical success of mathematics.

JUNE 1996 NOTICES OF THE AMS 657

Fractal Image Compression by Michael F. Barnsley

e.g: Barnsley's fern: can be encoded with 160 bytes= 4*10*4

4 maps 10 parameters (each parameter using 4 bytes)

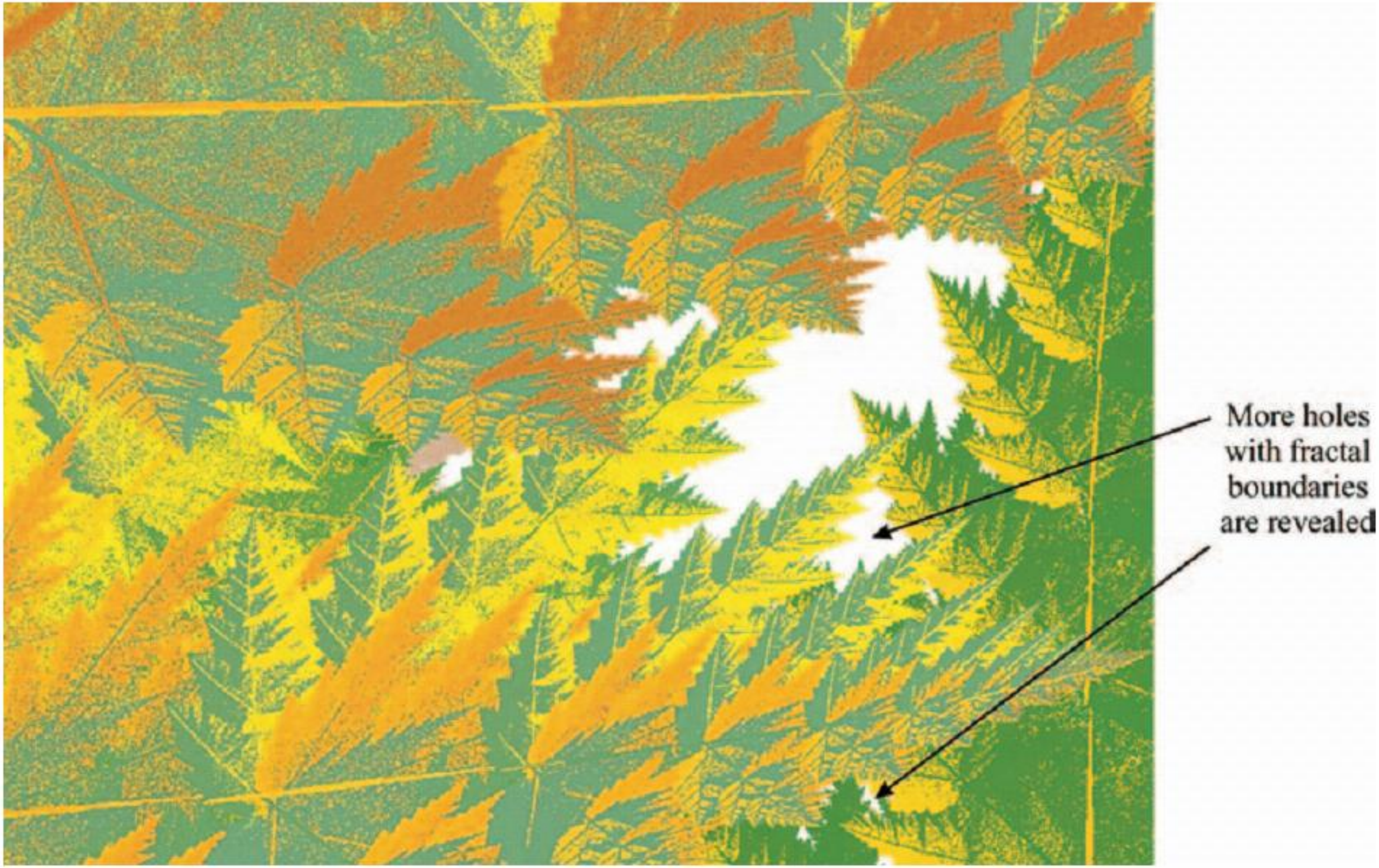


the measure attractor and (iii) the

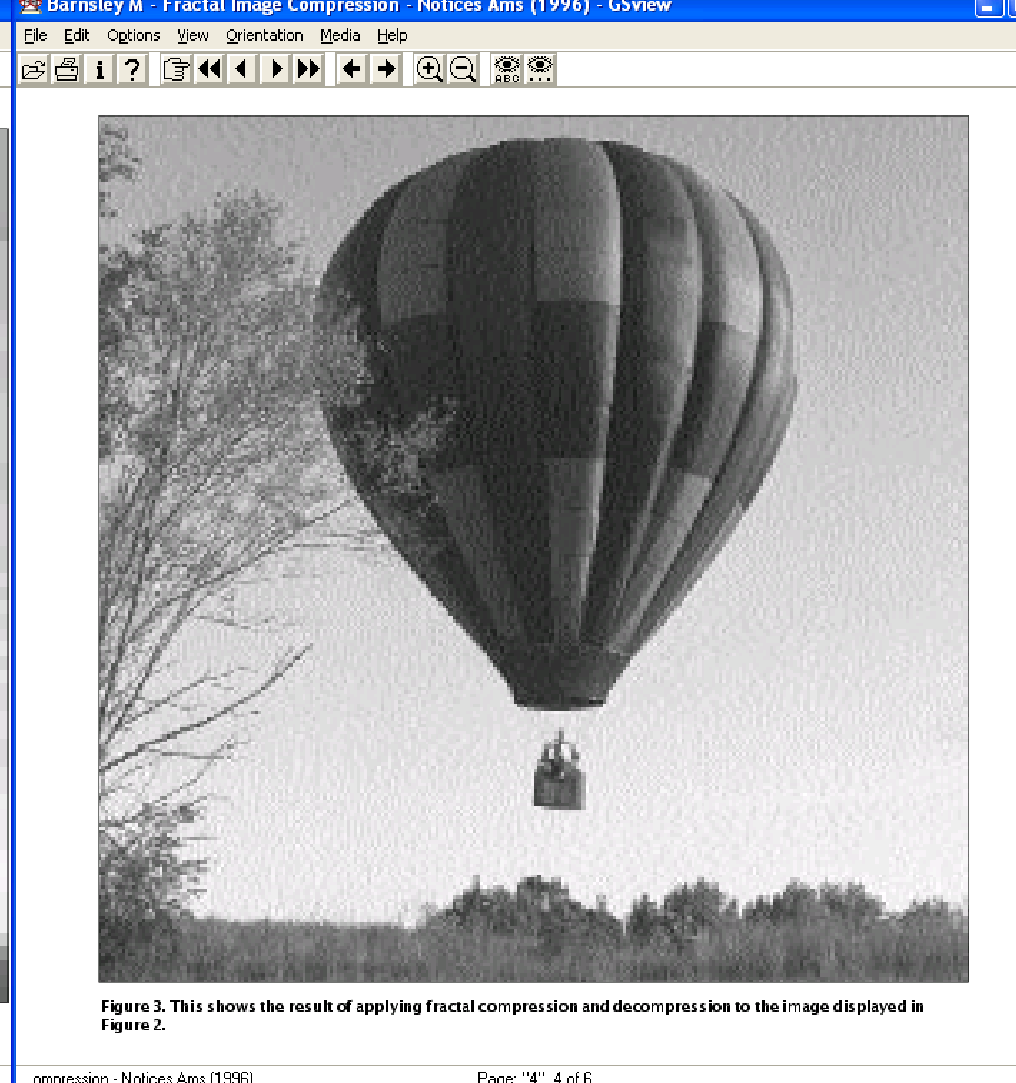
$$f_n(x, y) = \left(\frac{a_n x + b_n y + c_n}{g_n x + h_n y + j_n}, \frac{d_n x + e_n y + k_n}{g_n x + h_n y + j_n} \right)$$

n	a_n	b_n	c_n	d_n	e_n	k_n	g_n	h_n	j_n	p_n
1	19.05	0.72	1.86	−0.15	16.9	−0.28	5.63	2.01	20.0	$\frac{60}{100}$
2	0.2	4.4	7.5	−0.3	−4.4	−10.4	0.2	8.8	15.4	$\frac{1}{100}$
3	96.5	35.2	5.8	−131.4	−6.5	19.1	134.8	30.7	7.5	$\frac{20}{100}$
4	−32.5	5.81	−2.9	122.9	−0.1	−19.9	−128.1	−24.3	−5.8	$\frac{19}{100}$

From M. Barnsely SUPERFRACTALS
Cambridge University Press
2006



From M. Barnsely
SUPERFRACTALS
Cambridge University Press
2006



LEFT: the original digital image of Balloon, 512 pixels by 512 pixels, with 256 gray levels at each pixel. RIGHT: shows the same image after fractal compression. The fractal transform file is approximately one fifth the size of the original.

JUNE 1996 NOTICES OF THE AMS 657 Fractal Image Compression by Michael F. Barnsley

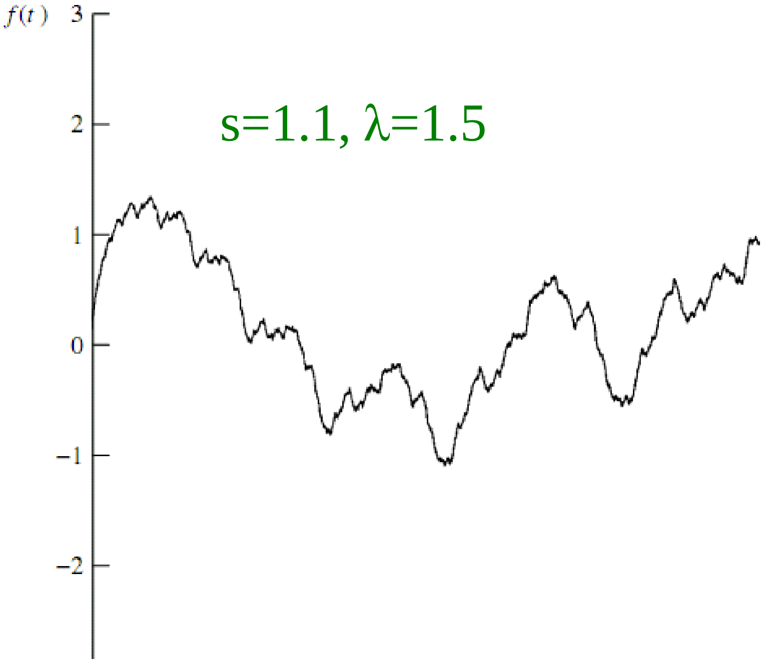
Fractal graphs of functions

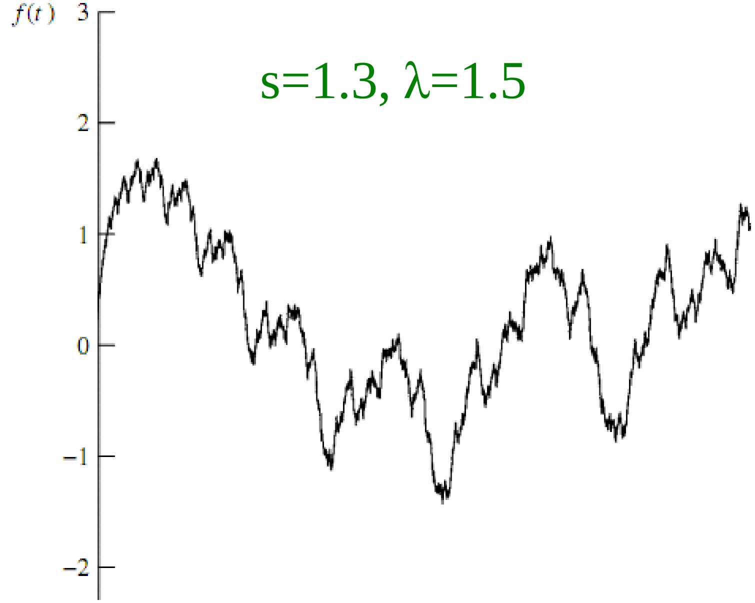
Many interesting fractals, both of theoretical and practical importance, occur as graphs of functions. Indeed many time series have fractal features, at least when recorded over fairly long time spans: examples include wind speed, levels of reservoirs, population data and some financial time series market (the famous Mandelbrot cotton graphs)

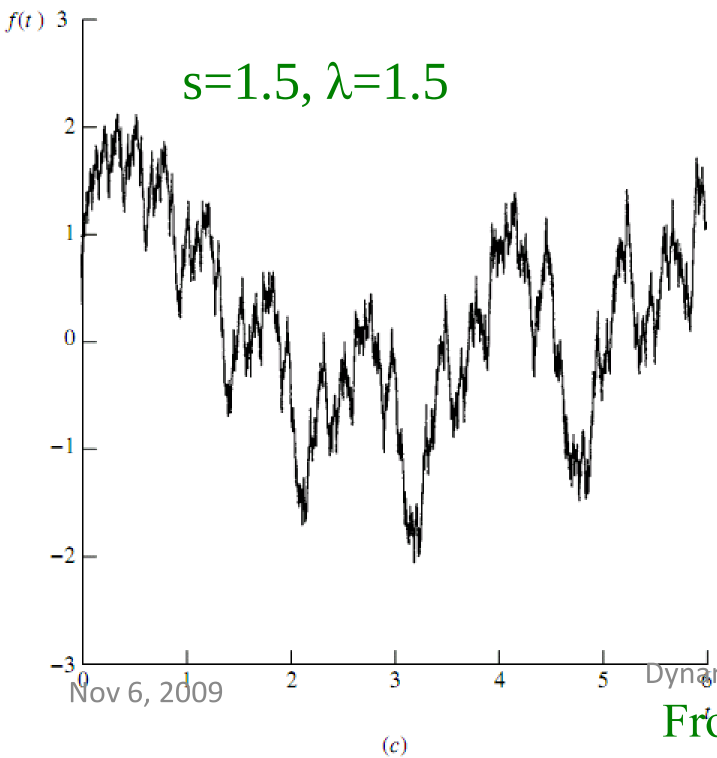
Weierstrass nowhere differentiable continuous function:

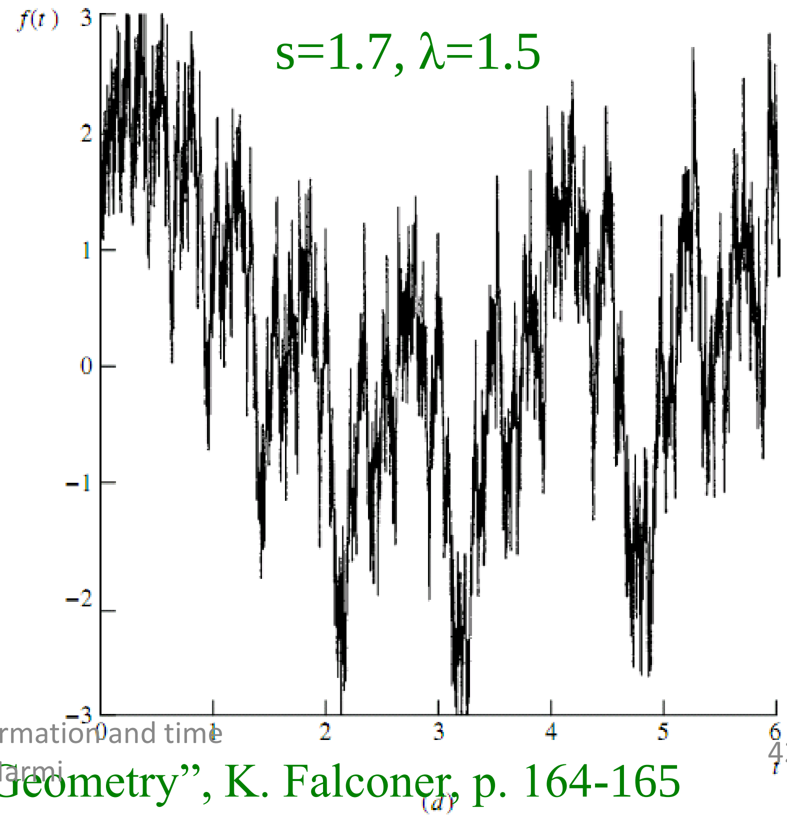
$$f(t)=\Sigma_{1\leq k\leq\infty}\ \lambda^{(s-2)k} \sin (\lambda^k\ t) \qquad 1<s<2,\ \lambda>2$$

The graph of f has box dimension s for λ large enough.









From “Fractal Geometry”, K. Falconer, p. 164-165

Fractal graph and i.f.s.

(from K. Falconer, Fractal Geometry, Wile (2003)

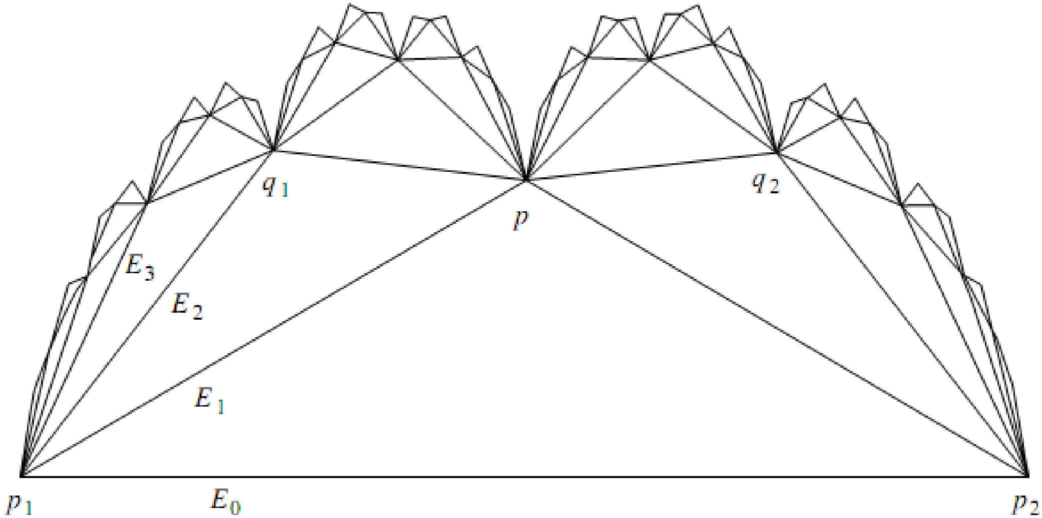


Figure 11.3 Stages in the construction of a self-affine curve F. The affine transformations S_1 and S_2 map the generating triangle p_1pp_2 onto the triangles p_1q_1p and pq_2p_2, respectively, and transform vertical lines to vertical lines. The rising sequence of polygonal curves $E_0, E_1, \ldots$ are given by $E_{k+1} = S_1(E_k) \cup S_2(E_k)$ and provide increasingly good approximations to F (shown in figure 11.4(*a*) for this case)

$$S_i(t, x) = (t/m + (i-1)/m,\ a_i t + c_i x + b_i).$$

Thus the S_i transform vertical lines to vertical lines, with the vertical strip $0 \leqslant t \leqslant 1$ mapped onto the strip $(i-1)/m \leqslant t \leqslant i/m$. We suppose that

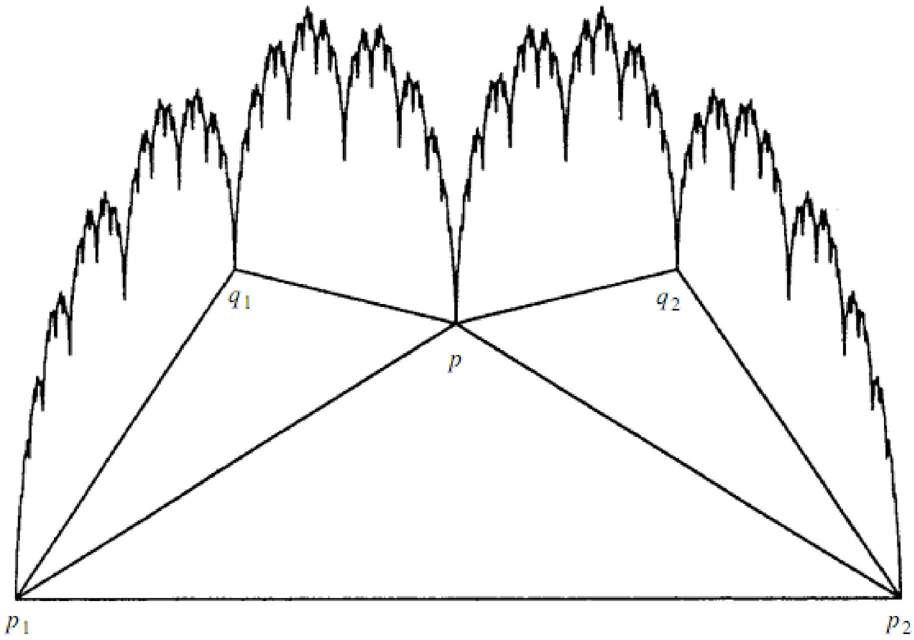
$$1/m < c_i < 1 \tag{11.9}$$

so that contraction in the t direction is stronger than in the x direction.

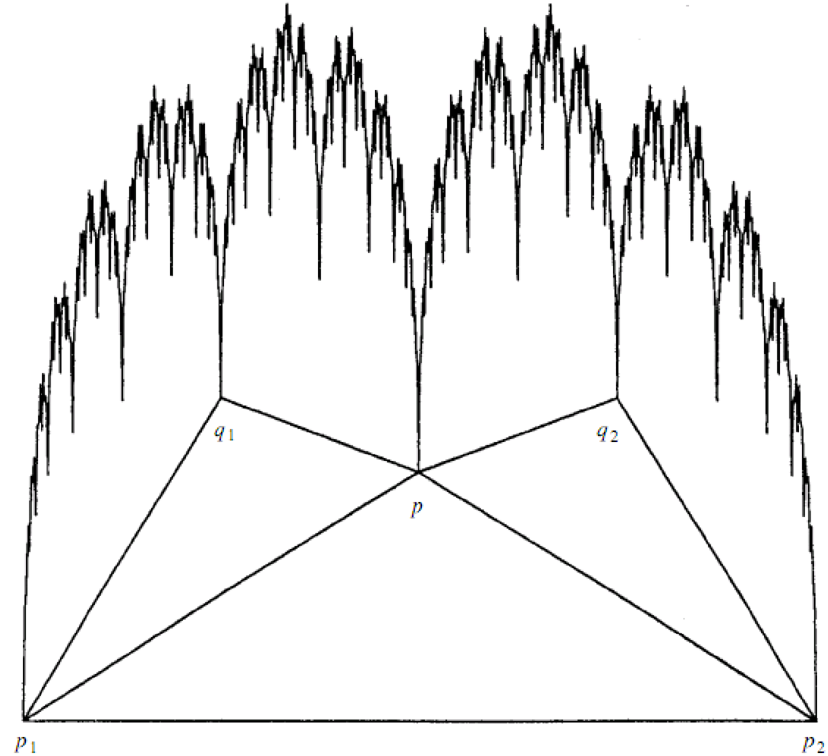
Let $p_1 = (0, b_1/(1-c_1))$ and $p_m = (1, (a_m + b_m)/(1-c_m))$ be the fixed points of S_1 and S_m. We assume that the matrix entries have been chosen so that

$$S_i(p_m) = S_{i+1}(p_1) \qquad (1 \leqslant i \leqslant m-1) \tag{11.10}$$

so that the segments $[S_i(p_1), S_i(p_m)]$ join up to form a polygonal curve E_1. To



(a)



Self-affine curves defined by the two affine transformations that map the triangle p_1pp_2 onto p_1q_1p and pq_2p_2 respectively. In (a) the vertical contraction of both transformations is 0.7 giving dim graph f = 1.49, and in (b) the vertical contraction of both transformations is 0.8, giving dim graph f = 1.68

from K. Falconer, Fractal Geometry, Wiley (2003)

Probabilistic i.f.s.

$\mathcal{F} = \{w_1, \ldots, w_N\}$, $w_i : X \to X$ contraction of constant s_i, $0 \le s_i < 1$

$(p_1,\ldots,p_N)$ probability vector $0 \le p_i \le 1$, $p_1+\ldots+p_N = 1$

Iteration: at each step with probability p_i one applies w_i

i.f.s.: k iterates of a point $\to N^k$ points $\mathcal{W} : \mathcal{H}(X) \to X$

$\mathcal{W}(E) = \cup_1 w_i(E)$

Probabilistic i.f.s.: k iterates of a point $\to$ k points

Theorem: each probabilistic i.f.s. has a unique Borel probability invariant measure μ with support = $\mathcal{A}$

Invariance: $\mu(E)= \Sigma_{1\le i\le N}\, p_i \mu(w_i^{-1}(E))$ for all Borel sets E, equivalently

$\int_X g(x)d\mu(x)= \Sigma_{1\le i\le N}\, p_i \int_X g(w_i(x))d\mu(x)$ for all continuous functions g

Probabilistic i.f.s.

If $\mathscr{M}$ denotes the space of Borel probability measures on X endowed with the metric

$$d(\nu_1,\nu_2)=\sup\{| \int_X g(x)d\nu_1 (x)-\int_X g(x)d\nu_2 (x)|, \text{ g Lipschitz, Lip}(g) \le 1\}$$

Then a probabilistic i.f.s. acts on measures as follows

$$L_{p,w}\, \nu= \Sigma\, p_i\nu\ w_i^{-1}$$

And by duality acts con continuos functions $g:X\to \mathbf{R}$

$$\int_X g(x)d(L_{p,w}\, \nu)(x)= \Sigma_{1\le i\le N}\, p_i \int_X g(w_i(x))d\nu(x)$$

It is easy to verify that

$$d(L_{p,w}\, \nu_1 , L_{p,w}\, \nu_2) \le s\ d(\nu_1,\nu_2)$$

from which the previous theorem follows

Multifractal analysis of measures

Local dimension (local Hölder exponent) of a measure μ at a point x:

$\dim_{loc} \mu(x)=\lim_{r\to 0} \log \mu(B(x,r))/\log r$ (when the limit exists)

$\alpha>0$, $E_\alpha =\{x \in X, \dim_{loc} \mu(x)= \alpha\}$

For certain measures μ the sets E_α may be non-empty over a range of values of α: multifractal measures

multifractal spectrum (singularity spectrum) of the multifractal measure μ: is the function $\alpha \to f(\alpha)=\dim E_\alpha$

With equal probabilities, the Random Algorithm for the IFS with these rules

$T_3(x, y) = (x/2, y/2) + (0, 1/2)$	$T_4(x, y) = (x/2, y/2) + (1/2, 1/2)$
$T_1(x, y) = (x/2, y/2)$	$T_2(x, y) = (x/2, y/2) + (1/2, 0)$

fills in the unit square uniformly.

The pictures below were generated with these probabilities

$$p_1 = 0.1, \; p_2 = p_3 = p_4 = 0.3.$$

Successive pictures show increments of 25000 points. With enough patience, the whole square will fill in, but some regions fill in more quickly than others.