Dynamics and time series: theory and applications

Stefano Marmi

Scuola Normale Superiore

Lecture 15, March 10, 2010

- Lecture 1: An introduction to dynamical systems and to time series. (Today, 2 pm - 4 pm Aula Dini)
- Lecture 2: Ergodicity. Uniform distribution of orbits. Return times. Kac inequality Mixing (Thu Jan 14, 2 pm - 4 pm Aula Fermi) by Giulio Tiozzo
- Lecture 3: Kolmogorov-Sinai entropy. Randomness and deterministic chaos. (Wen Jan 20, 2 pm - 4 pm Aula Bianchi) by Giulio Tiozzo
- Lecture 4: Introduction to financial markets and to financial time series (Thu Jan 21,

  2 pm - 4 pm Aula Bianchi **Lettere**)
- Lecture 5: Central limit theorems (Wen Jan 27, 2 pm - 4 pm Bianchi) by Giulio Tiozzo
- Lecture 6: Financial time series: stylized facts and models (Thu Jan 28, 2 pm - 4 pm Bianchi)
- Lecture 7: The Efficient Market Hypothesis (Wen Feb 10)
- Lecture 8: An introduction to market microstructure and to high frequency finance, by Fabrizio Lillo (Thu Feb 11, Aula Dini)
- Lecture 9 on Wen Feb 17More on the efficient market hypothesis
- Lecture 10 An introduction to autoregressive models and to mean-variance optimization, Wen Feb 24
- Lecture 11 On equity trading strategies by A. Carollo, Thu Feb 25

- Lecture 12 Volatility by Roberto Renò, Mar 2
- Lecture 13 An introduction to ARMA and GARCH processes by Fulvio Corsi, Mar 3
- Lecture 14 HAR models for realized volatility: extensions and applications, by Fulvio Corsi, Mar 4
- Lecture 15 Takens' Theorem and an introduction to fractals and multifractals, TODAY
- Lecture 16 Factor models for the analysis of large datasets with applications to economics and finance, by Massimiliano Marcellino (European University Institute), Thu Mar 18, Aula Dini

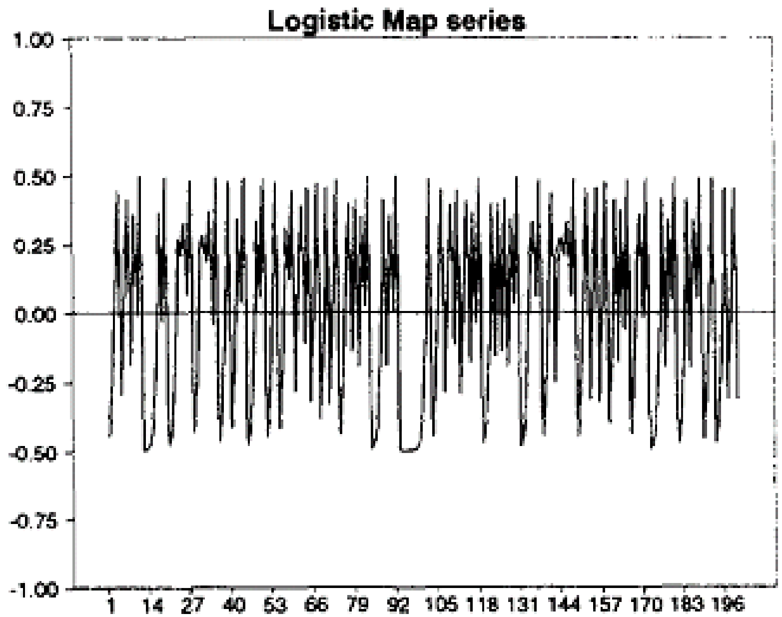
- Challenges and experiments:

  0. blog: http://theworldisatimeseries.wordpress.com

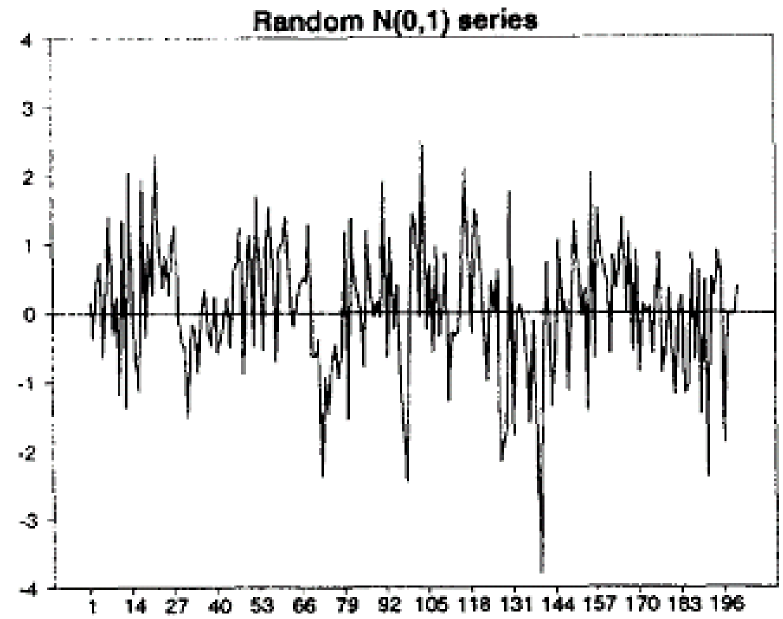
  1. statistical arbitrage in sports betting: collecting time series, etc..

  2. nonstationarity and volatility of financial series

Logistic map series (adjusted with mean)



Random N(0,1) series



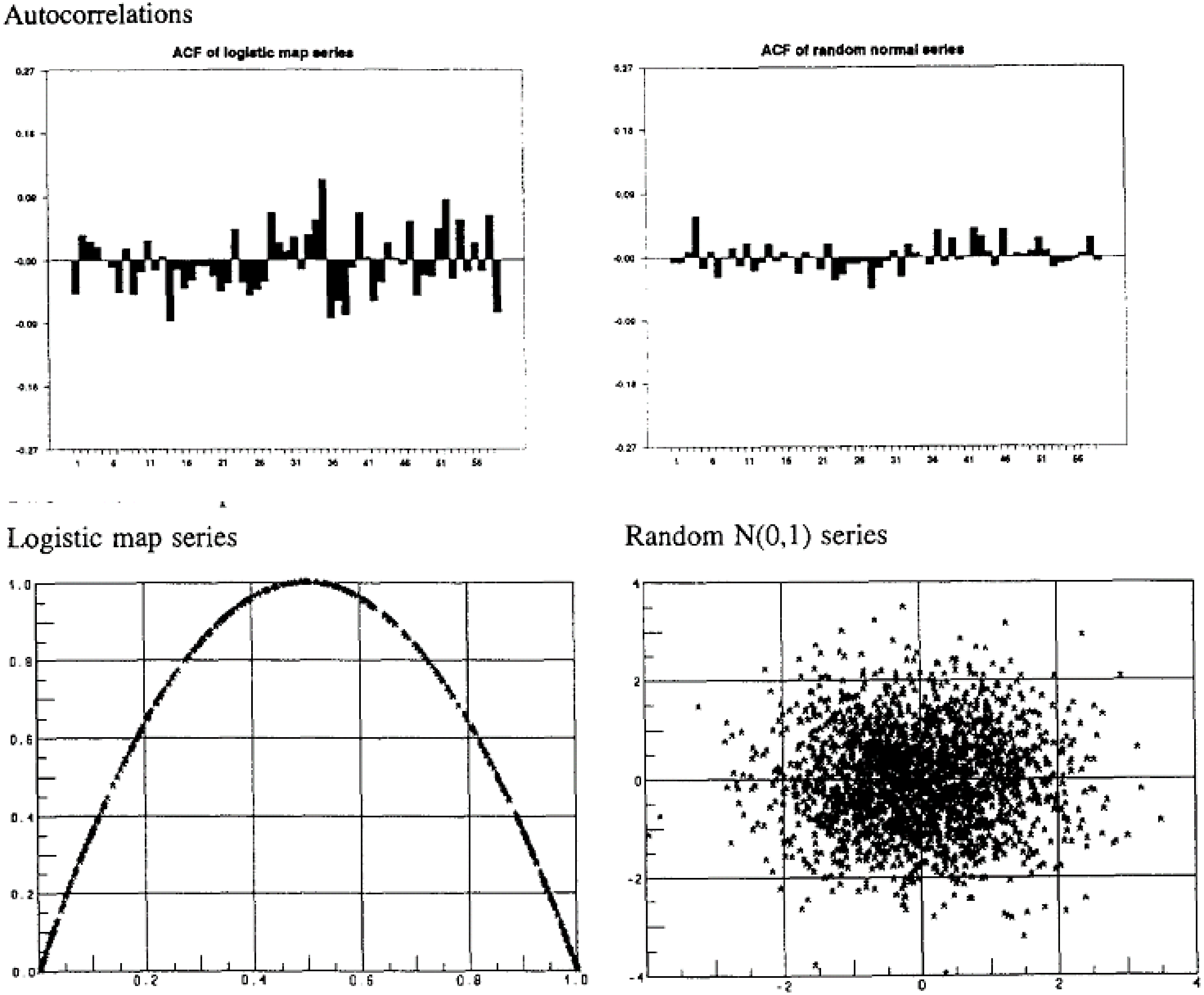


Fig. 2. Comparison of logistic map and random series.

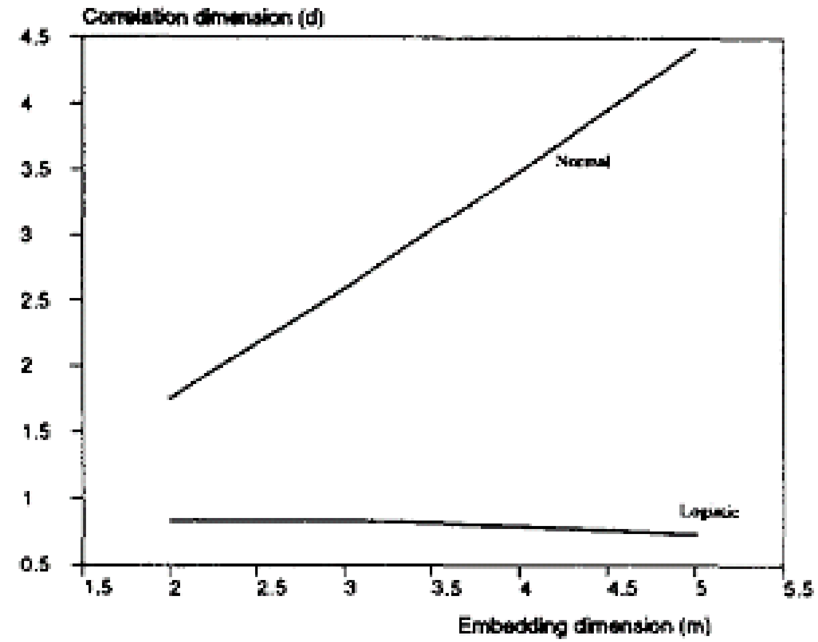
K. Takala, M. Virén / European Journal of Operational Research 93 (1996) 155–172

Embedding dimension = m

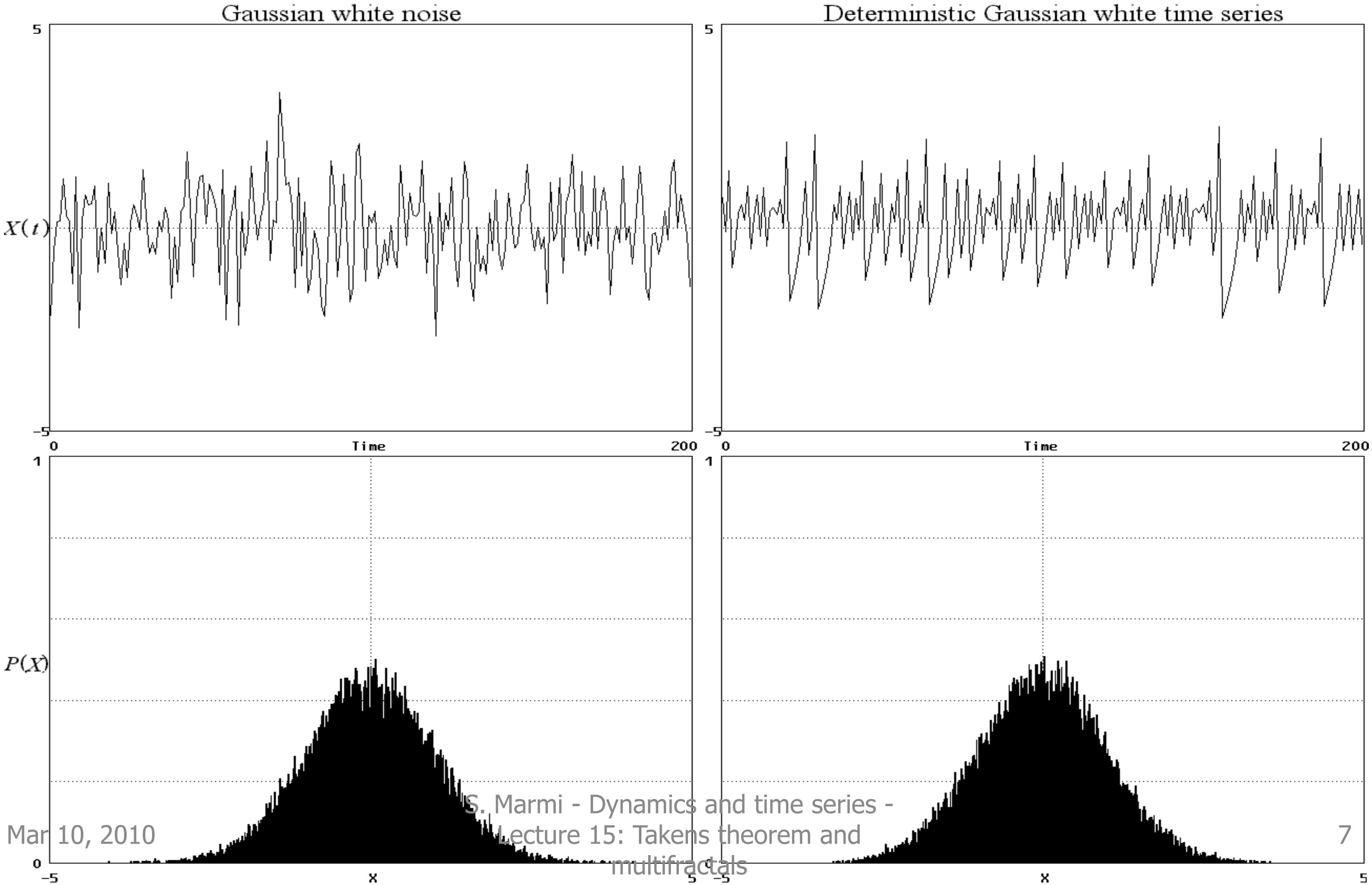
$$C_m(\varepsilon) = \lim_{N \to \infty} \frac{1}{N^2} \# \left\{ (x_{m,i}, x_{m,j}), \left\| x_{m,i} - x_{m,j} \right\| < \varepsilon \right\}$$

$$d(m) = \lim_{\varepsilon \to 0} \frac{\log C_m(\varepsilon)}{\log(\varepsilon)}$$

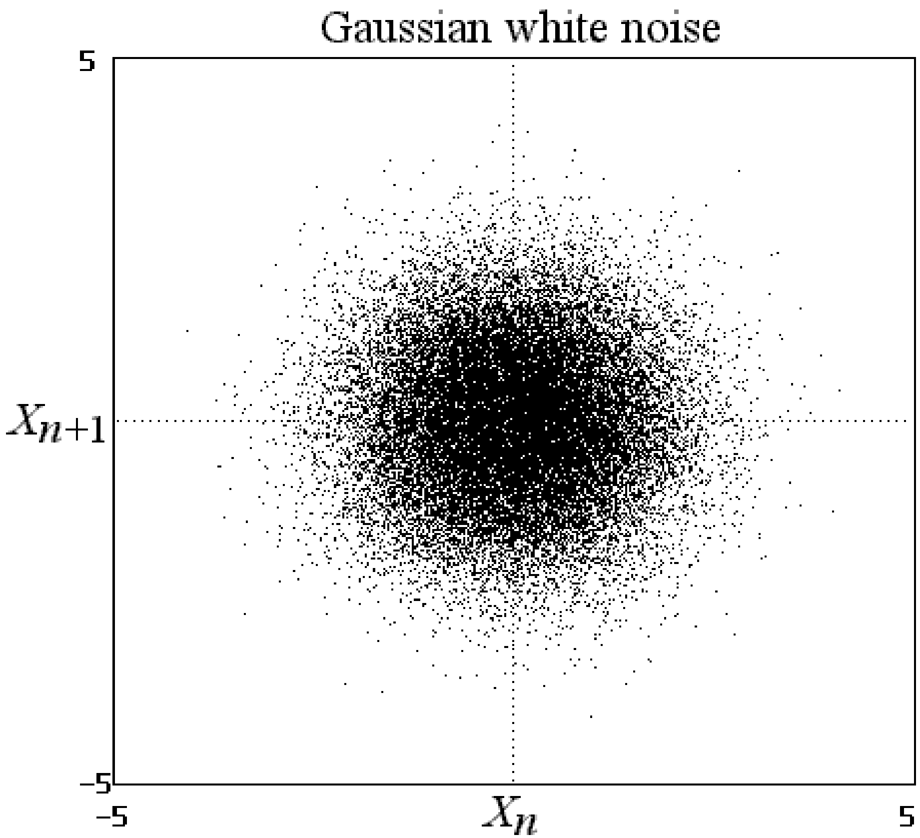
Correlation dimensions of logistic map and random normal processes

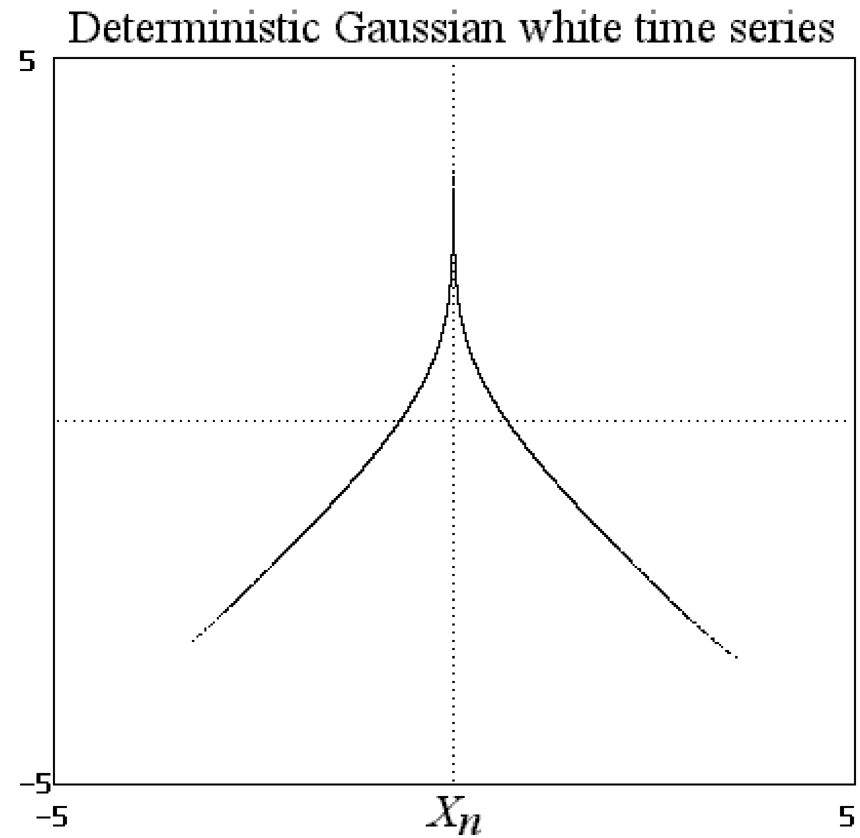


Deterministic or random? Appearance can be misleading…



Time delay map





Source: **sprott**.physics.wisc.edu/lectures/**tsa**.ppt

Logit and logistic

The logistic map $x \to L(x)=4x(1-x)$ preserves the probability measure $d\mu(x)=dx/(\pi\sqrt{x(1-x)})$

The transformation $h:[0,1] \to \mathbf{R}$, $h(x)=\ln x-\ln(1-x)$ conjugates L with a new map G

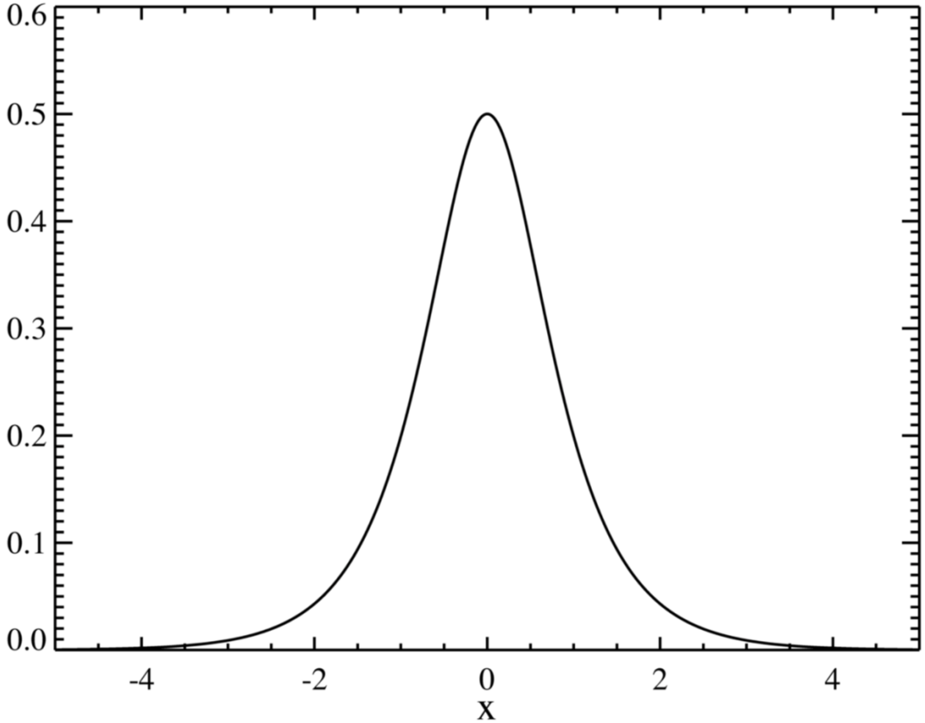
$$h \quad L=G \quad h$$

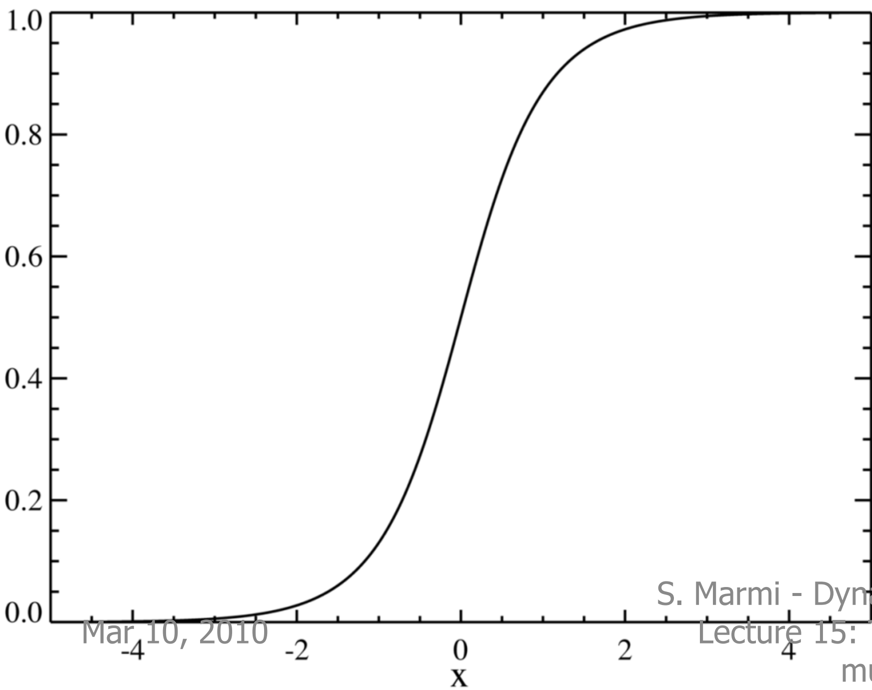
definined on **R.** The new invariant probability measure is $d\mu(x)=dx/[\pi(e^{x/2}+e^{-x/2})]$

G and L have the same dynamics (the only difference is a coordinates change)

Hyperbolic secant distribution

Source: wikipedia





Parameters	***none***
Support	x∈(-∞,+∞)
Probability density function (pdf)	$\frac{1}{2}\operatorname{sech}(\frac{1}{2}\pi x)$
Cumulative distribution function (cdf)	$\frac{2\arctan(\exp(\frac{1}{2}\pi x))}{\pi}$
Mean	0
Median	0
Mode	0
Variance	1
Skewness	0
Excess kurtosis	2
Entropy	$4/\pi\, G \approx 1.16624$

G = 0.915 965 594 177 219 015 054 603 514 932 384 110 774… Catalan's constant

Takens theorem

- $\phi : X \to X$ map, $f : X \to R$ smooth observable
- Time-delay map (reconstruction of the dynamics from periodic sampling):
- $F(f,\phi) : X \to R^n$ n is the number of delays
- $F(f,\phi)(x) = (f(x), f(\phi(x)), f(\phi\circ\phi(x)), ..., f(\phi^n (x)))^{-1}$
- Under mild assumptions if the dynamics has an attractor with dimension k and n>2k then for almost any choice of the observable the reconstruction map is injective

Immersions and embeddings

- A smooth map F on a compact smooth manifold A is an immersion if the derivative map DF(x) (represented by the Jacobian matrix of F at x) is one-to-one at every point $x \in A$. Since DF(x) is a linear map, this is equivalent to DF(x) having full rank on the tangent space. This can happen whether or not F is one-to-one. Under an immersion, no differential structure is lost in going from A to F(A).
- An embedding of A is a smooth diffeomorphism from A onto its image F(A), that is, a smooth one-to-one map which has a smooth inverse. For a compact manifold A, the map F is an embedding if and only if ,F is a one- to-one immersion.
- The set of embeddings is open in the set of smooth maps: arbitrarily small perturbations of an embedding will still be embeddings!

Embedology (Sauer, Yorke, Casdagli, J. Stat. Phys. 65 (1991)

Whitney showed that a generic smooth map ,F from a d-dimensional smooth compact manifold M to R^n , n>2d is actually a diffeomorphism on M. That is, M and F(M) are diffeomorphic. We generalize this in two ways:

- first, by replacing "generic" with "probability-one" (in a prescribed sense),
- second, by replacing the manifold M by a compact invariant set A

contained in some Rk that may have noninteger box-counting dimension (boxdim). In that case, we show that almost every smooth map from a neighborhood of A to R^n is one-to-one as long as n>2 * boxdim(A)

We also show that almost every smooth map is an embedding on compact subsets of smooth manifolds within l. This suggests that embedding techniques can be used to compute positive Lyapunov exponents (but not necessarily negative Lyapunov exponents). The positive Lyapunov exponents are usually carried by smooth unstable manifolds on attractors.

Embedology (Sauer, Yorke, Casdagli, J. Stat. Phys. 65 (1991))

Takens dealt with a restricted class of maps called delay-coordinate maps: these are time series of a single observed quantity from an experiment. He showed (F. Takens, Detecting strange attractors in turbulence, in Lecture Notes in Mathematics, No. 898 (Springer-Verlag, 1981) that if the dynamical system and the observed quantity are generic, then the delay-coordinate map from a d-dimensional smooth compact manifold M to R^n , n>2d is a diffeomorphism on M.

- we replace generic with probability-one
- and the manifold M by a possibly fractal set.

Thus, for a compact invariant subset A under mild conditions on the dynamical system, almost every delay-coordinate map to R^n is one-to-one on A provided that n>2.boxdim(A). Also, any manifold structure within I will be preserved in F(A).

- Only C^1 smoothness is needed.;
- For flows, the delay must be chosen so that there are no periodic orbits with period exactly equal to the time delay used or twice the delay

Embedding method

- Plot $x(t)$ vs. $x(t-\tau)$, $x(t-2\tau)$, $x(t-3\tau)$, …
- $x(t)$ can be any observable
- The embedding dimension is the # of delays
- The choice of τ and of the dimension are critical
- For a typical deterministic system, the orbit will be diffeomorphic to the attractor of the system (Takens theorem)

Choice of Embedding Parameters

Theoretically, a time delay coordinate map yields an valid embedding for any sufficiently large embedding dimension and for any time delay when the data are noise free and measured with infinite precision.

But, there are several problems:

(i) Data are not clean
(ii) Large embedding dimension are computationally expensive and unstable
(iii) Finite precision induces noise

Effectively, the solution is to search for:

(i) Optimal time delay τ
(ii) Minimum embedding dimension d

or

(i) Optimal time window τ_w

There is no one unique method solving all problems and neither there is an unique set of embedding parameters appropriate for all purposes.

Choice of Embedding Parameters

Theoretically, a time delay coordinate map yields an valid embedding for any sufficiently large embedding dimension and for any time delay when the data are noise free and measured with infinite precision.

But, there are several problems:

(i) Data are not clean
(ii) Large embedding dimension are computationally expensive and unstable
(iii) Finite precision induces noise

Effectively, the solution is to search for:

(i) Optimal time delay τ
(ii) Minimum embedding dimension d

or

(i) Optimal time window τ_w

There is no one unique method solving all problems and neither there is an unique set of embedding parameters appropriate for all purposes.

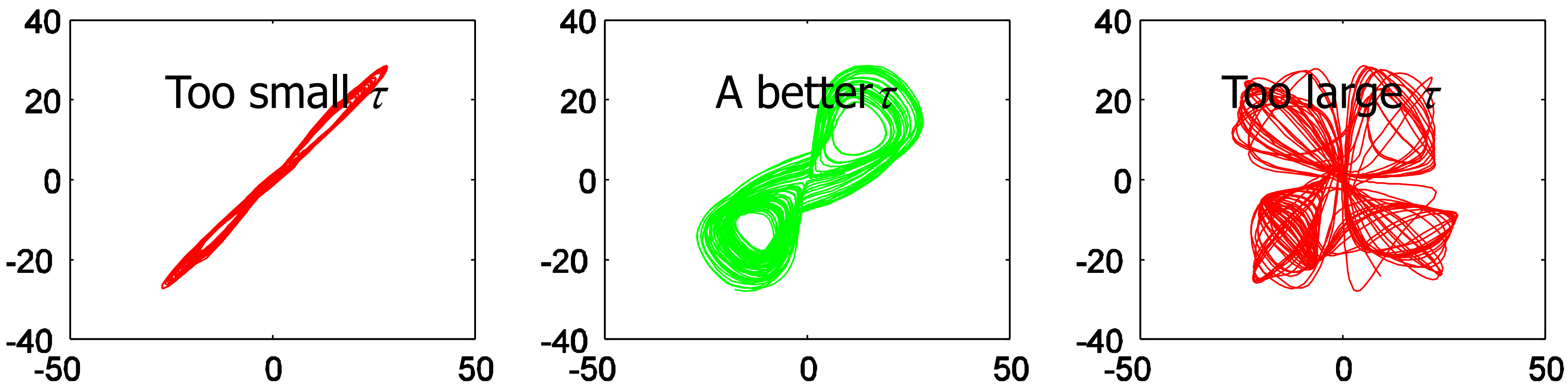
The Role of Time Delay τ

If τ is too small, $x(t)$ and $x(t-\tau)$ will be very close, then each reconstructed vector will consist of almost equal components → *Redundancy (τ_R)*

➡ The reconstructed state space will collapse into the main diagonal

If τ is too large, $x(t)$ and $x(t-\tau)$ will be completely unrelated, then each reconstructed vector will consist of irrelevant components → *Irrelevance (τ_I)*

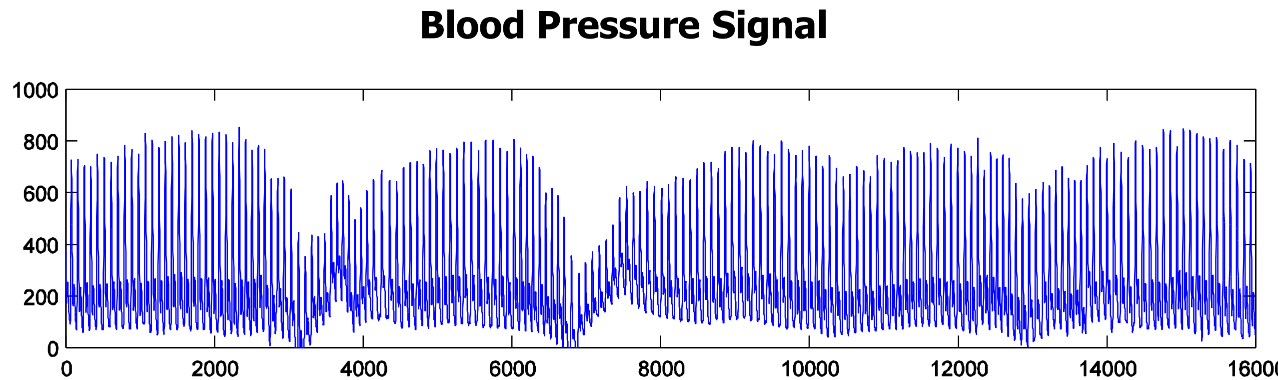
➡ The reconstructed state space will fill the entire state space.



http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

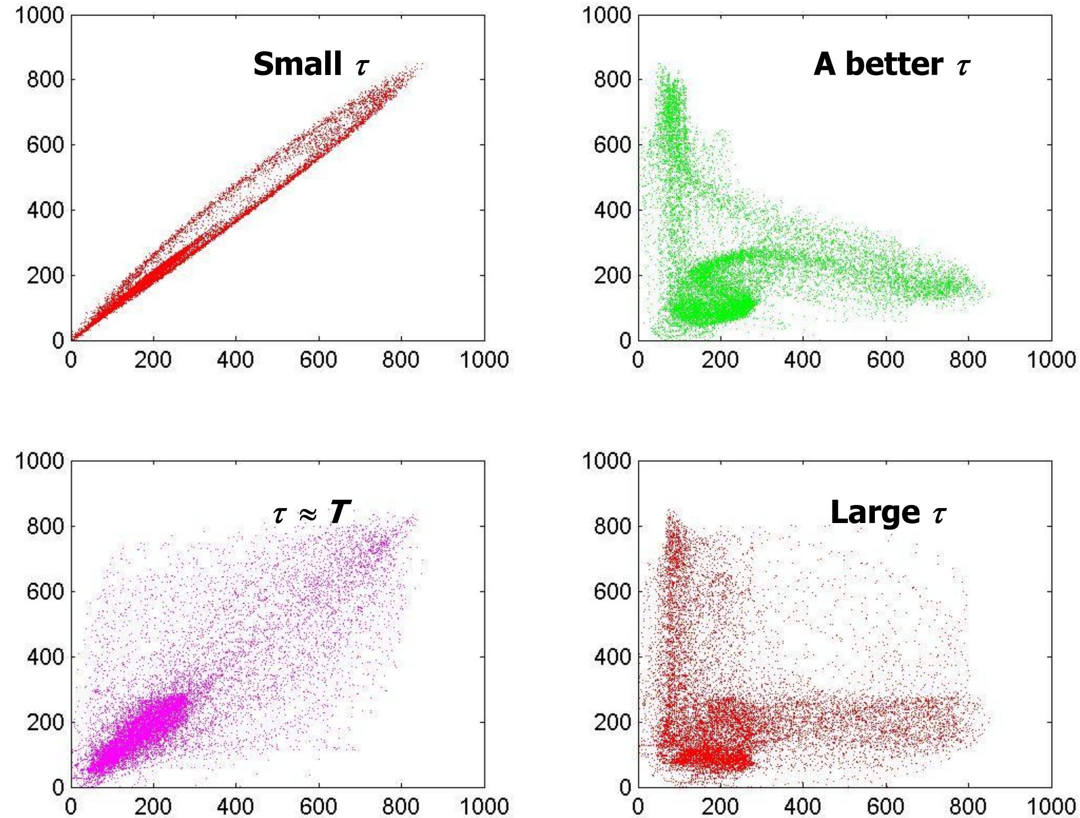
A better choice is:

$\tau_R < \tau_W < \tau_I$



Caution: τ should not be close to main period

Collapsing of state space



http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

Some Recipes to Choose τ

Based on Autocorrelation

Estimate autocorrelation function: $C(\tau) = \frac{1}{N-\tau-1}\sum_{t=0}^{N-\tau-1} x(t)x(t+\tau) = \langle x(t)x(t+\tau)\rangle$

Then, $\tau_{opt} \approx C(0)/e$

or

first zero crossing of $C(\tau)$

Modifications:

1. Consider minima of higher order autocorrelation functions, $<x(\tau)x(t+\tau)x(t+2\tau)>$ and then look for time when these minima for various orders coincide.

Albano et al. (1991) *Physica D*

2. Apply nonlinear autocorrelation functions: $<x^2(\tau)x^2(t+2\tau)>$

Billings, Tao (1991) *Int. J. Control.*

http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

Based on Time delayed Mutual Information

The information we have about the value of $x(t+\tau)$ if we know $x(t)$.

1. Generate the histogram for the probability distribution of the signal $x(t)$.
2. Let p_i is the probability that the signal will be inside the i-th bin and $p_{ij}(t)$ is the probability that $x(t)$ is in i-th bin and $x(t+\tau)$ is in j-th bin.
3. Then the mutual information for delay τ will be

$$I(\tau) = \sum_{i,j} p_{ij}(\tau) \log p_{ij}(\tau) - 2\sum_i p_i \log p_i$$

For $\tau \to 0$, $I(\tau) \to$ Shannon's Entropy

$\tau_{opt} \approx$ First minimum of $I(\tau)$

http://www.viskom.oeaw.ac.at/~joy/March22,%202004.ppt

Decay time of autocorrelation

$$\tau_d = \min\{\tau : C_{xx}(\tau) < \frac{1}{e}\}$$

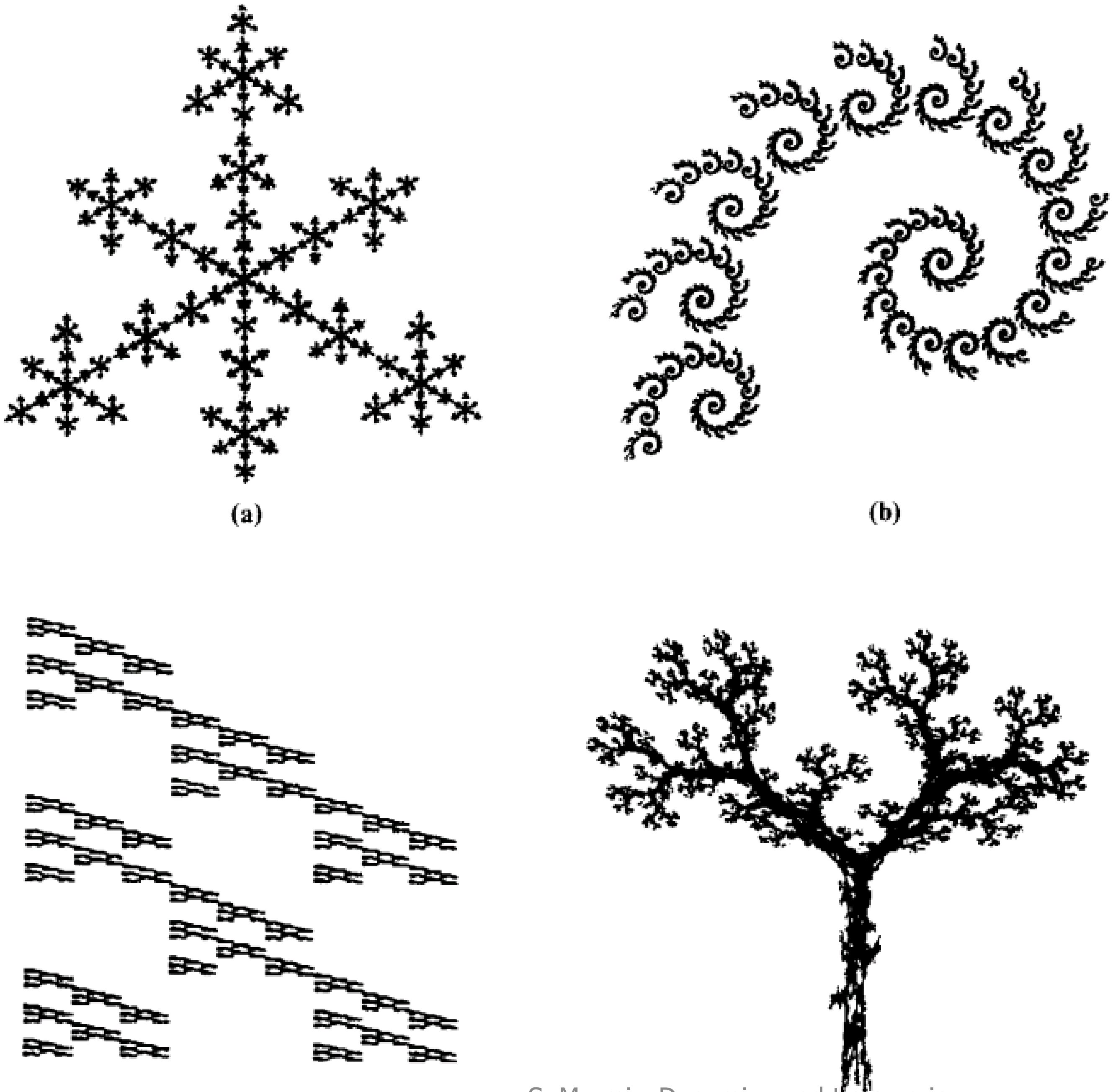
This is an important indicator of the strength of the autocorrelation of time series

It can be used to determine the time delay in embedology

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)



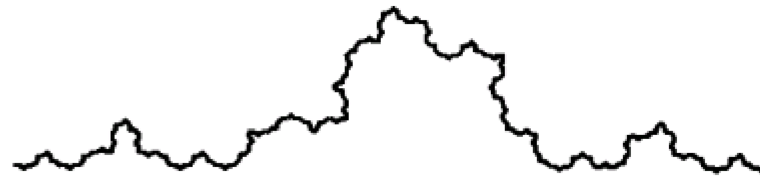
self-similar fractals

self-affine fractals

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

(e)

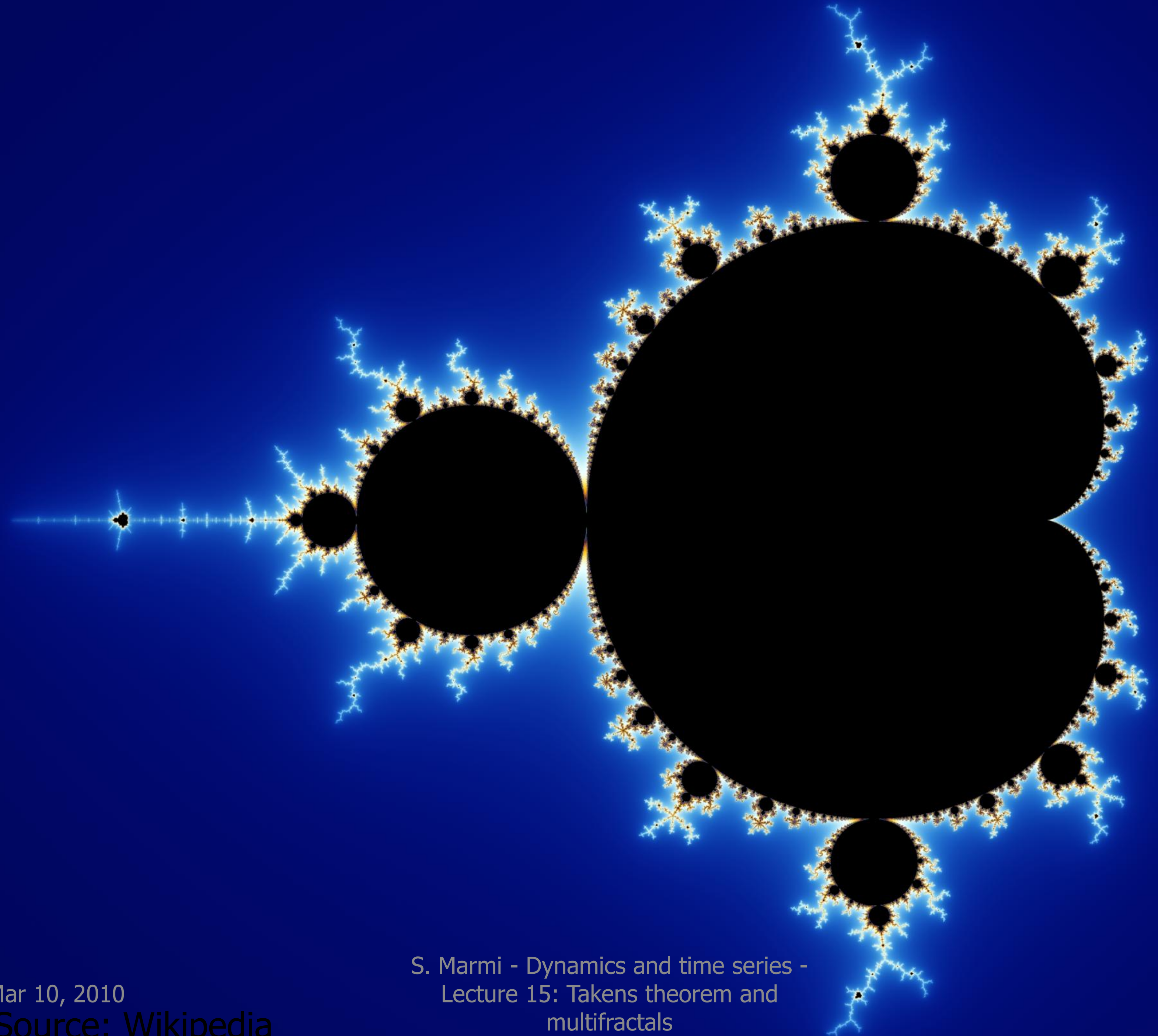
self-conformal
fractals



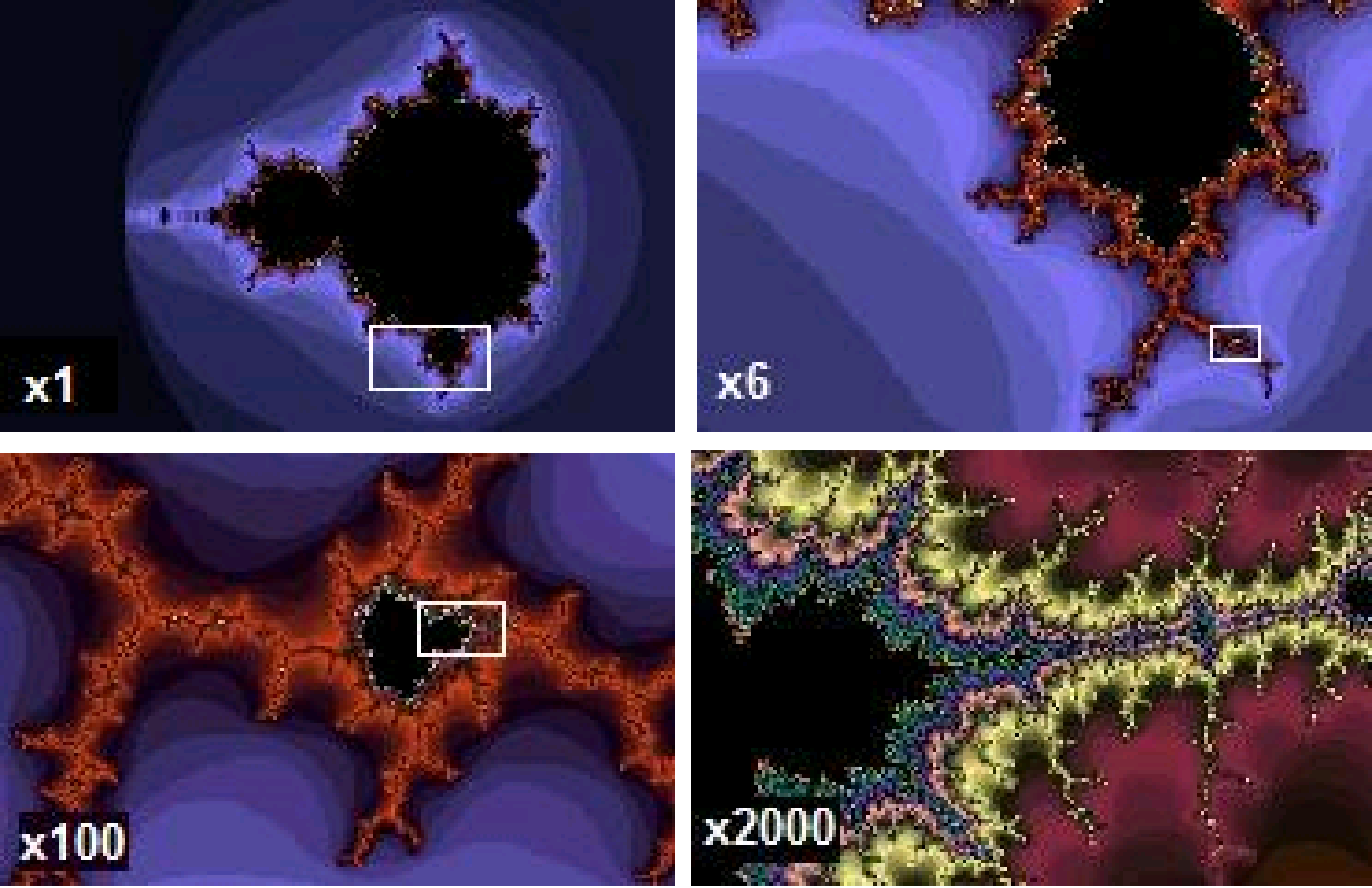
(f)

Statistically
self-similar
fractals

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

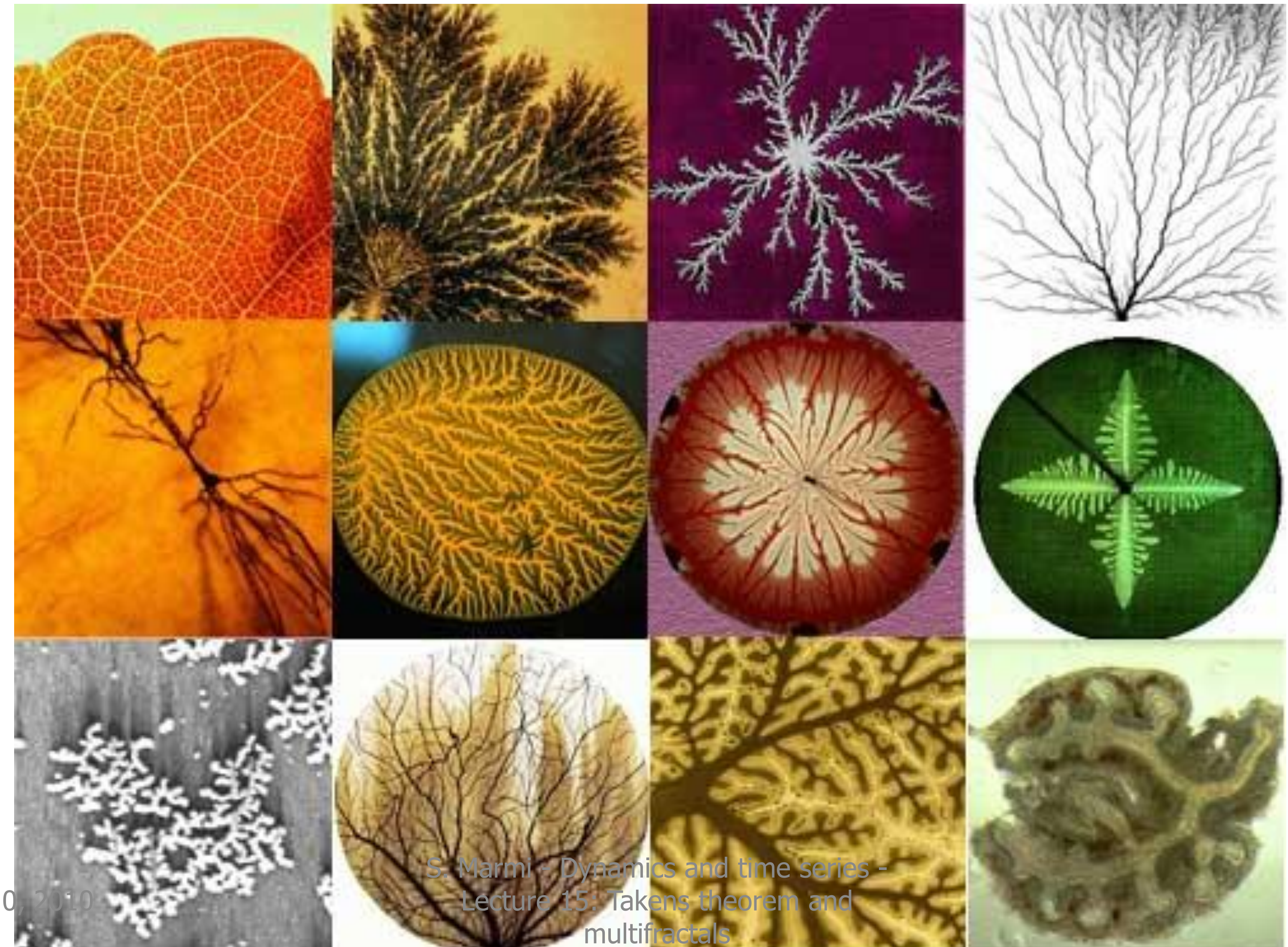


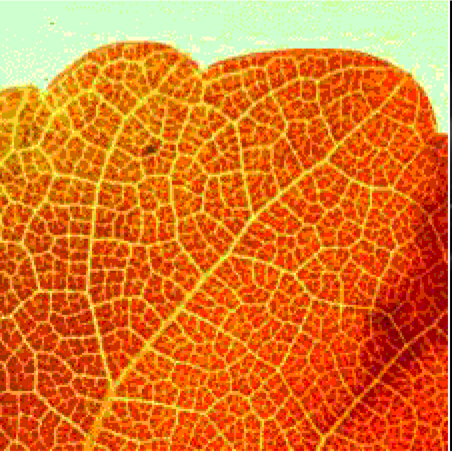
Source: Wikipedia



Source: Wikipedia

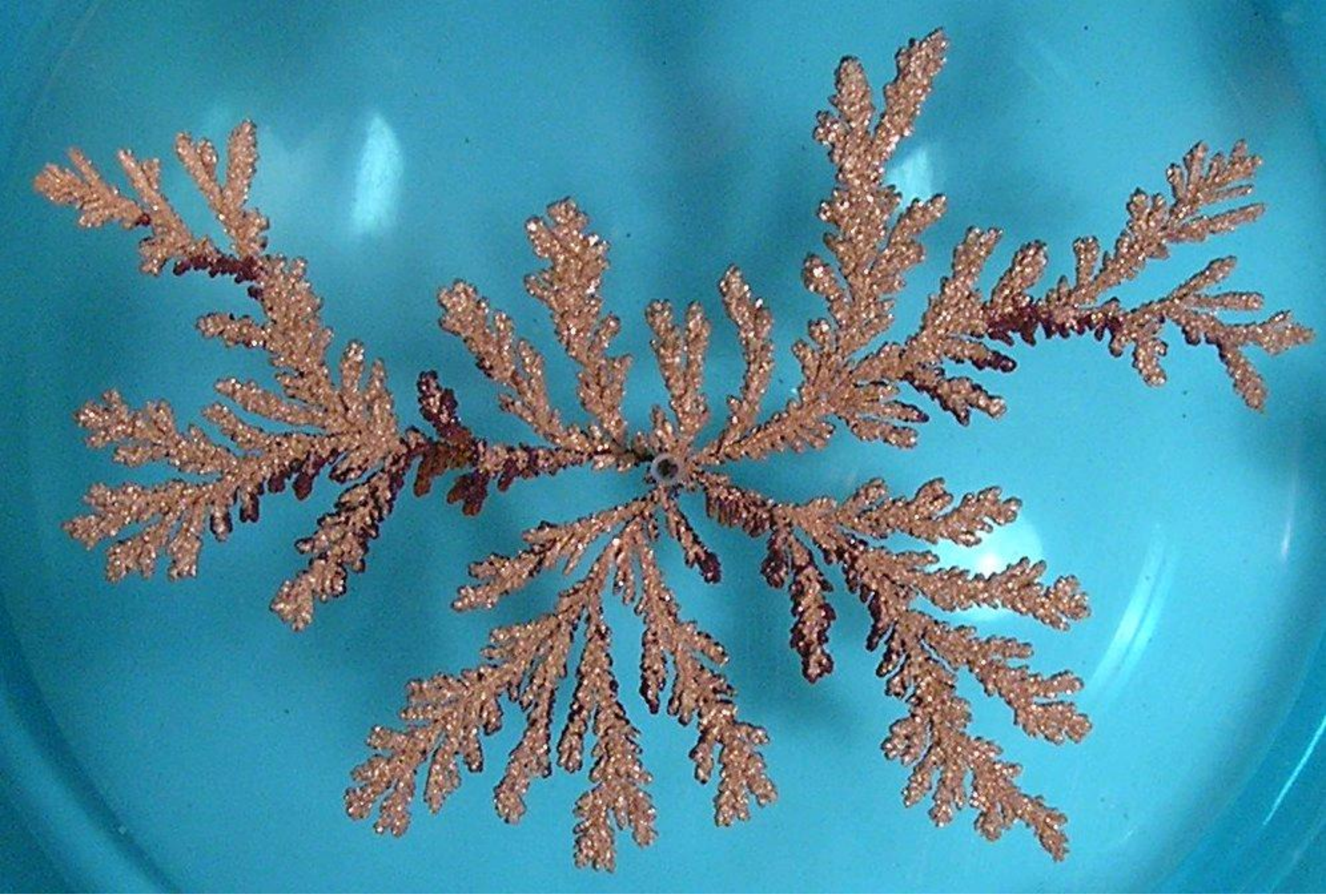
Mathematics, shapes and nature





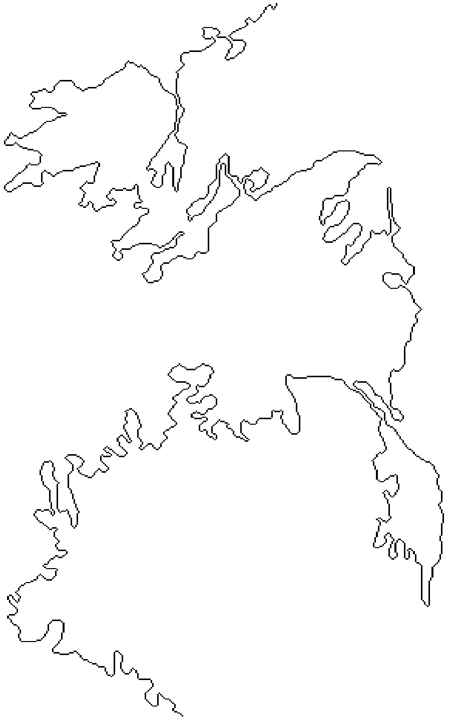
http://classes.yale.edu/fractals/Panorama/Nature/NatFracGallery/Gallery/Gallery.html

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

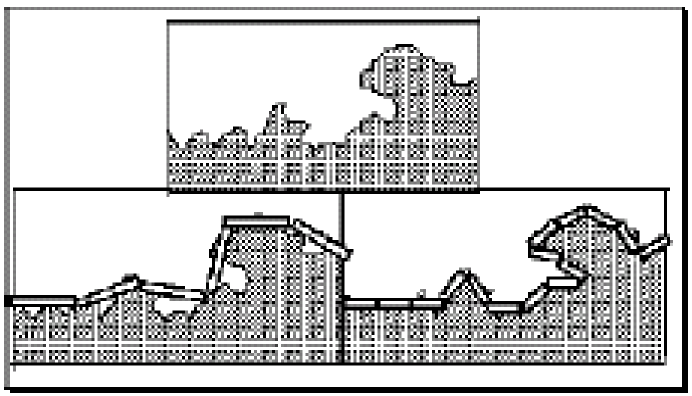
Greece

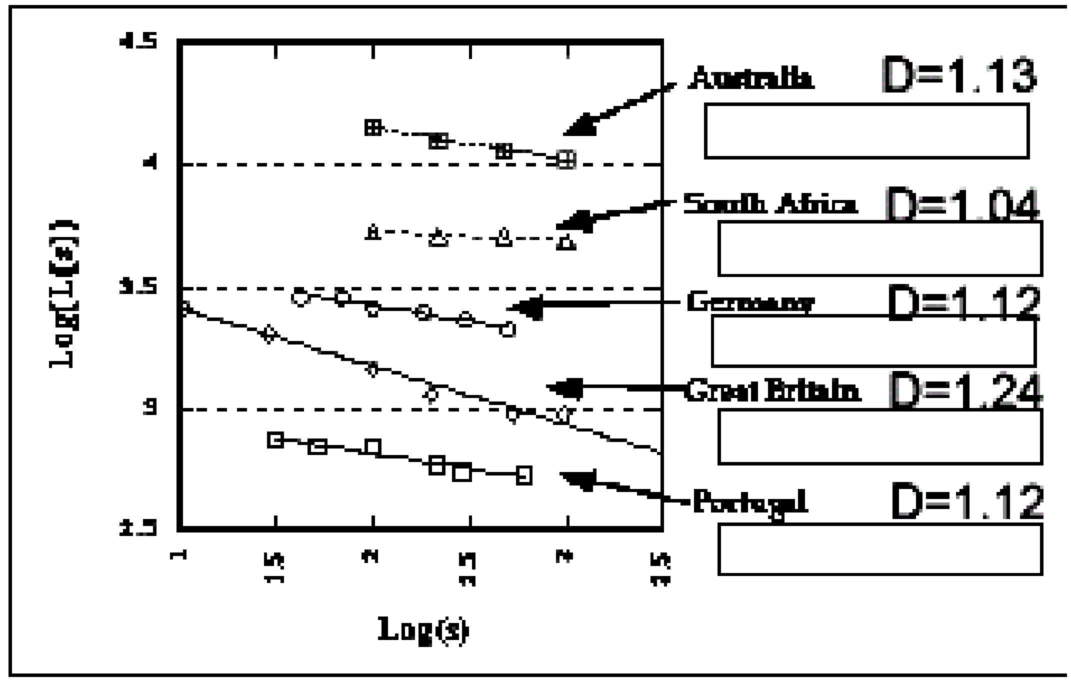
D=1.20

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

$$\text{Log } L(s) = (1-D) \log s + \text{cost}$$

(Richardson 1961, Mandelbrot Science 1967)

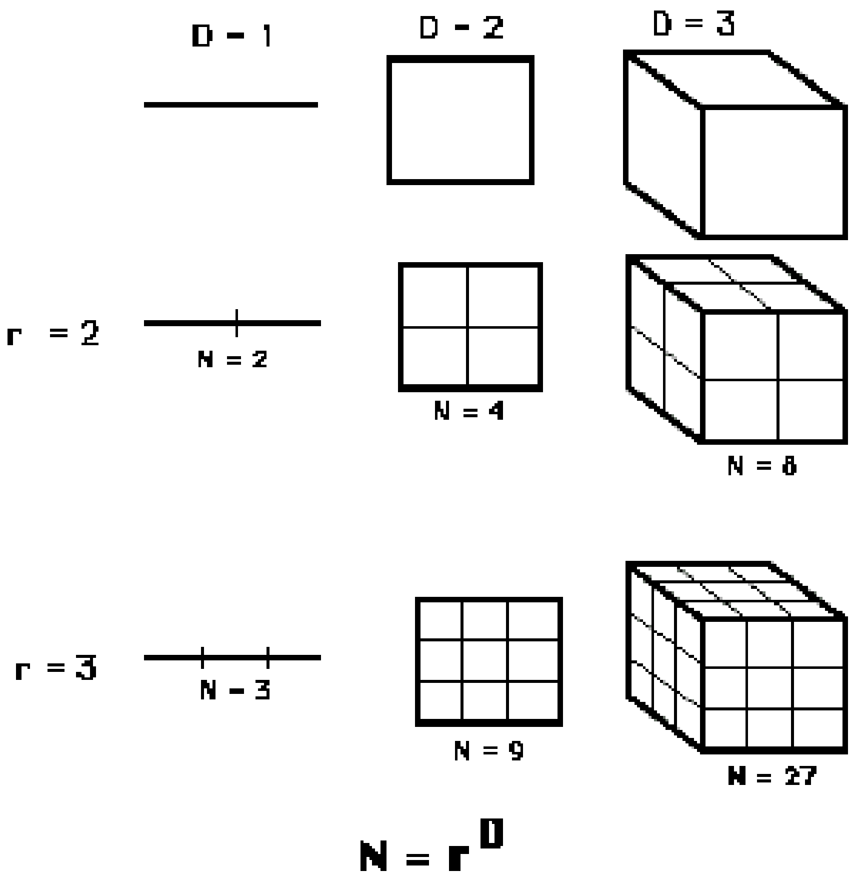
How long is the coast of Britain? Statistical self-similarity and fractional dimension Science: 156, 1967, 636-638 B. B. Mandelbrot

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.

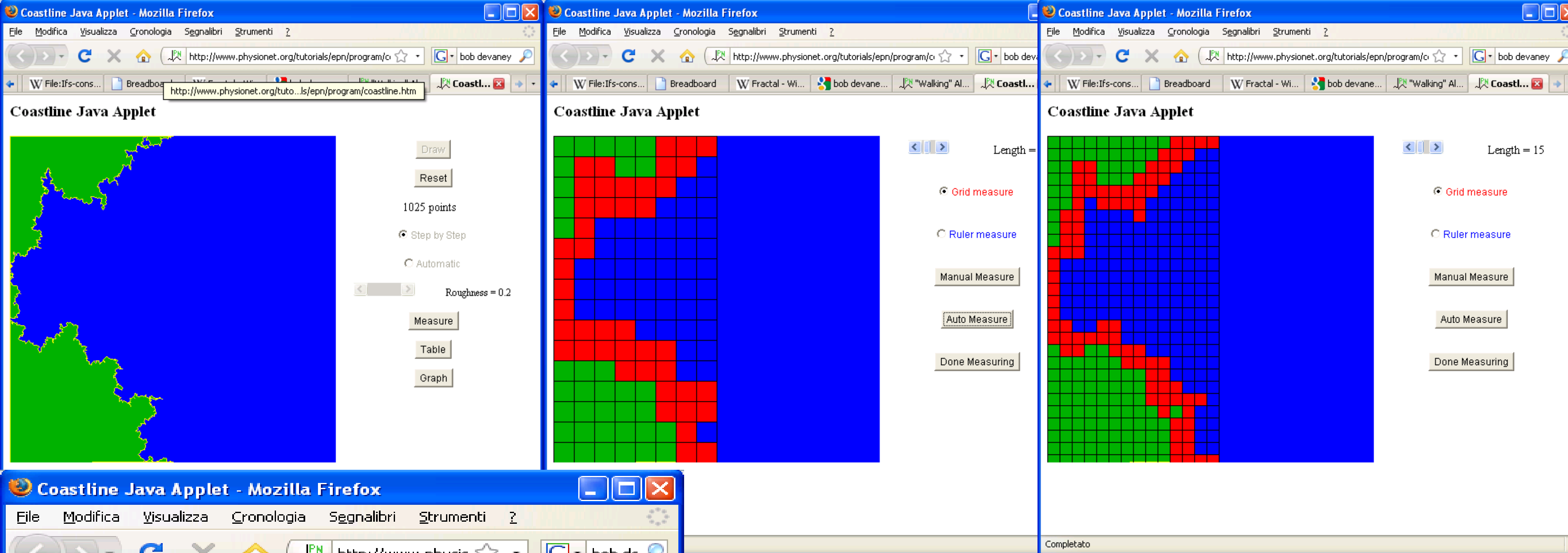
How long is the coast of Britain? Statistical self-similarity and fractional dimension Science: 156, 1967, 636-638 B. B. Mandelbrot

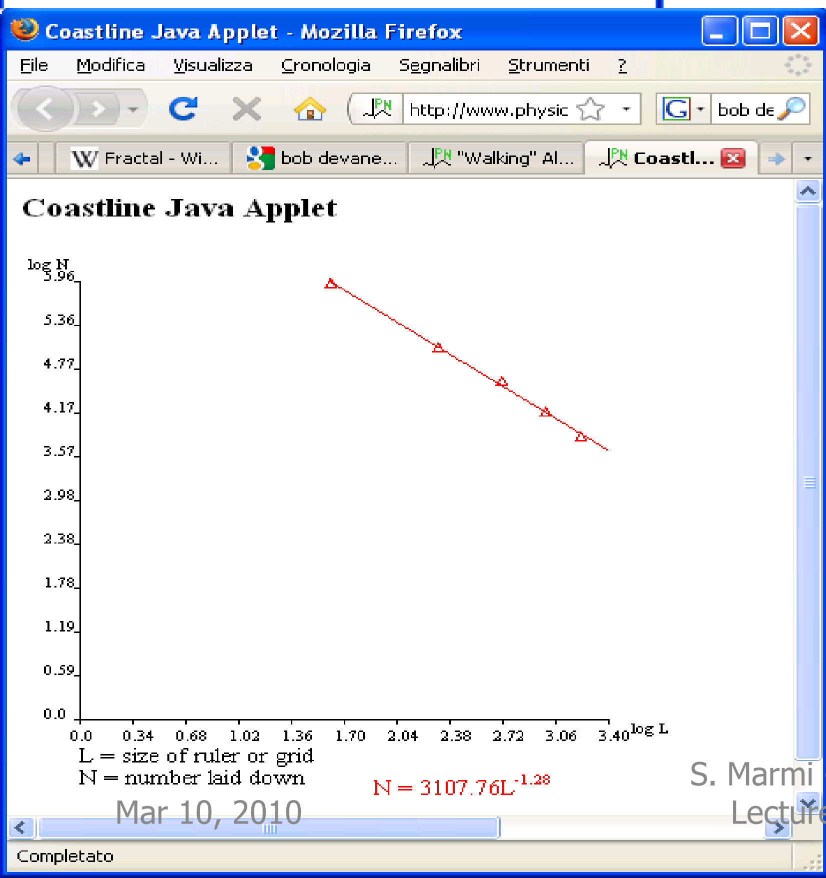
Quantities other than length are therefore needed to discriminate between various degrees of complication for a geographical curve. When a curve is self-similar, it is characterized by an exponent of similarity, D, which possesses many properties of a dimension, though it is usually a fraction greater 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

Log N(s) = -D log s + 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

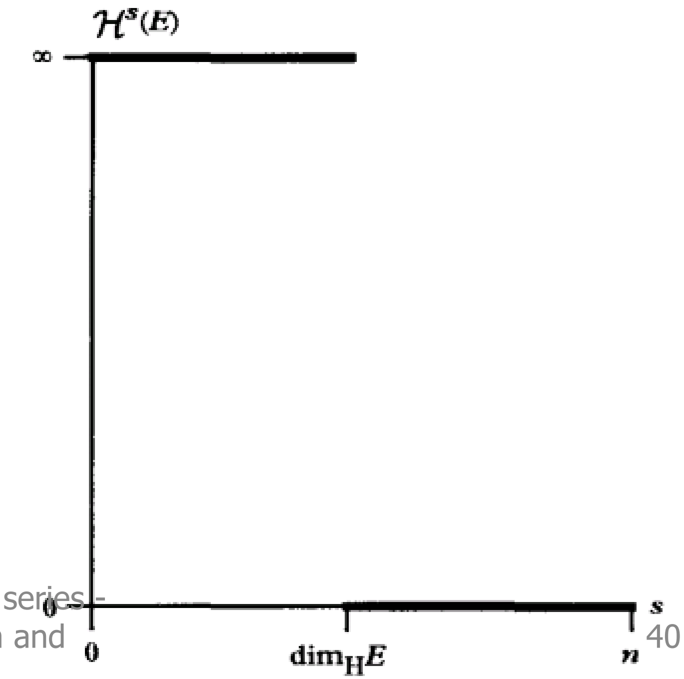
A finite or countable collection of subsets $\{U_i\}$ of $\mathbf{R^n}$ is a δ-cover of a set E if $|U_i| < \delta$ for all i and E is contained in $\bigcup_i U_i$

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

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

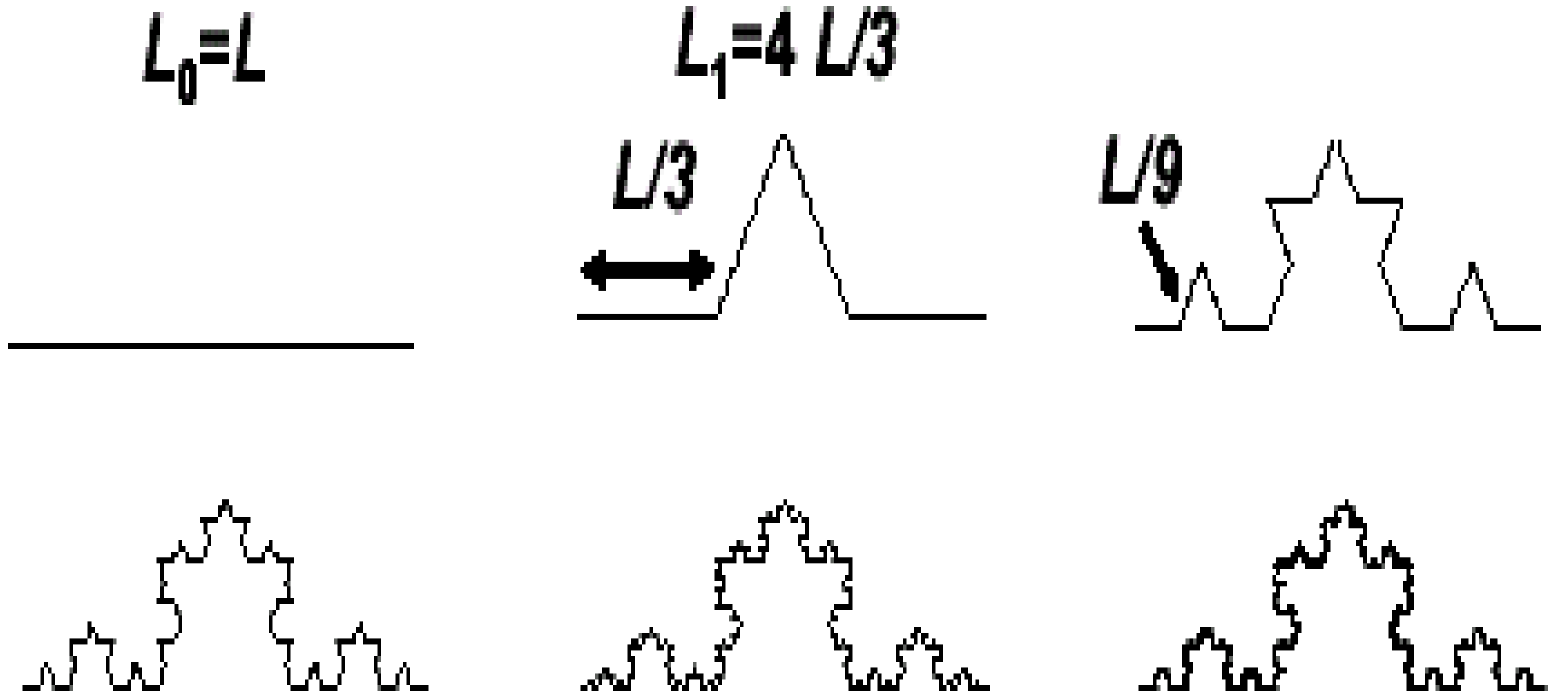
It is a Borel regular measure on $\mathbf{R^n}$, it behaves well under similarities and Lipschitz maps

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...

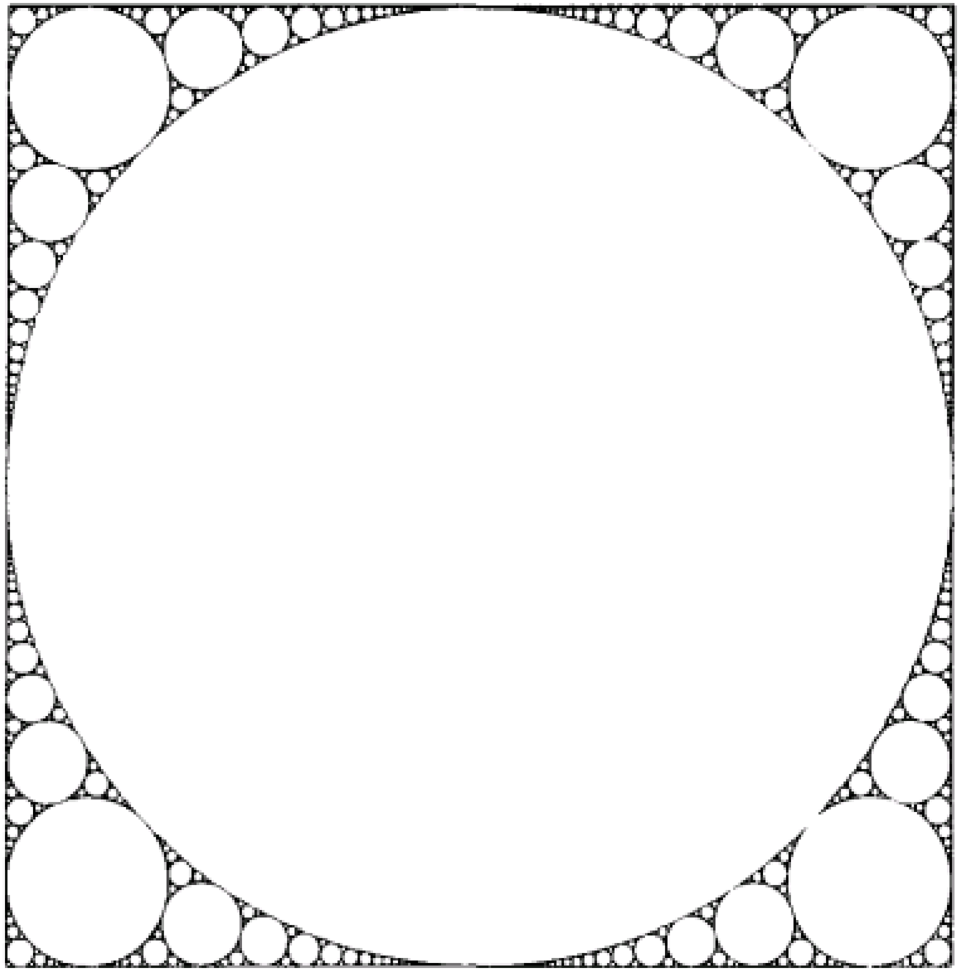
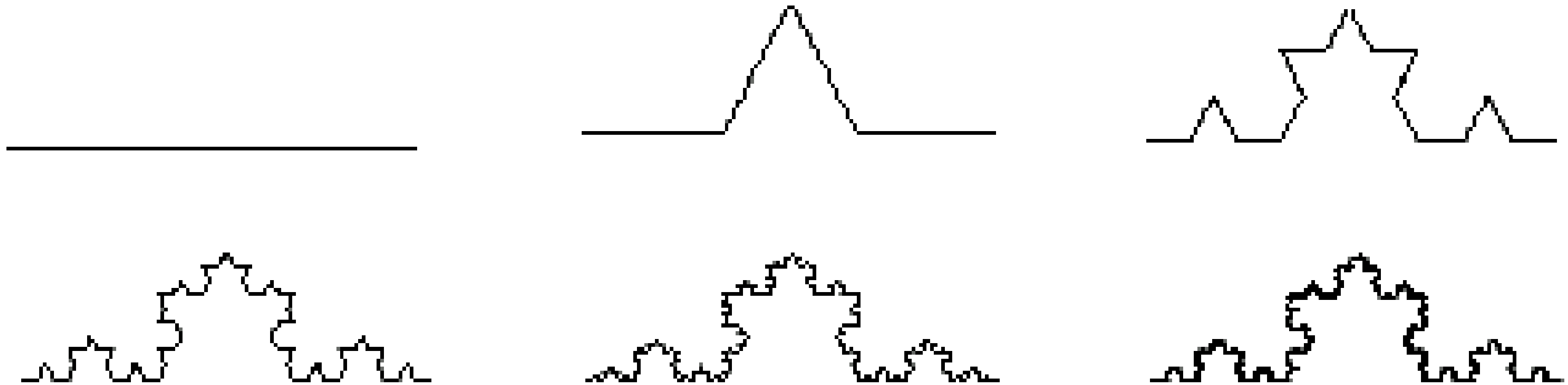


Figure 3.6 A cut-out set in the plane. Here, the largest possible disc is removed at each step. The family of discs removed is called the Apollonian packing of the square, and the cut-out set remaining is called the residual set, which has Hausdorff and box dimension about 1.31

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



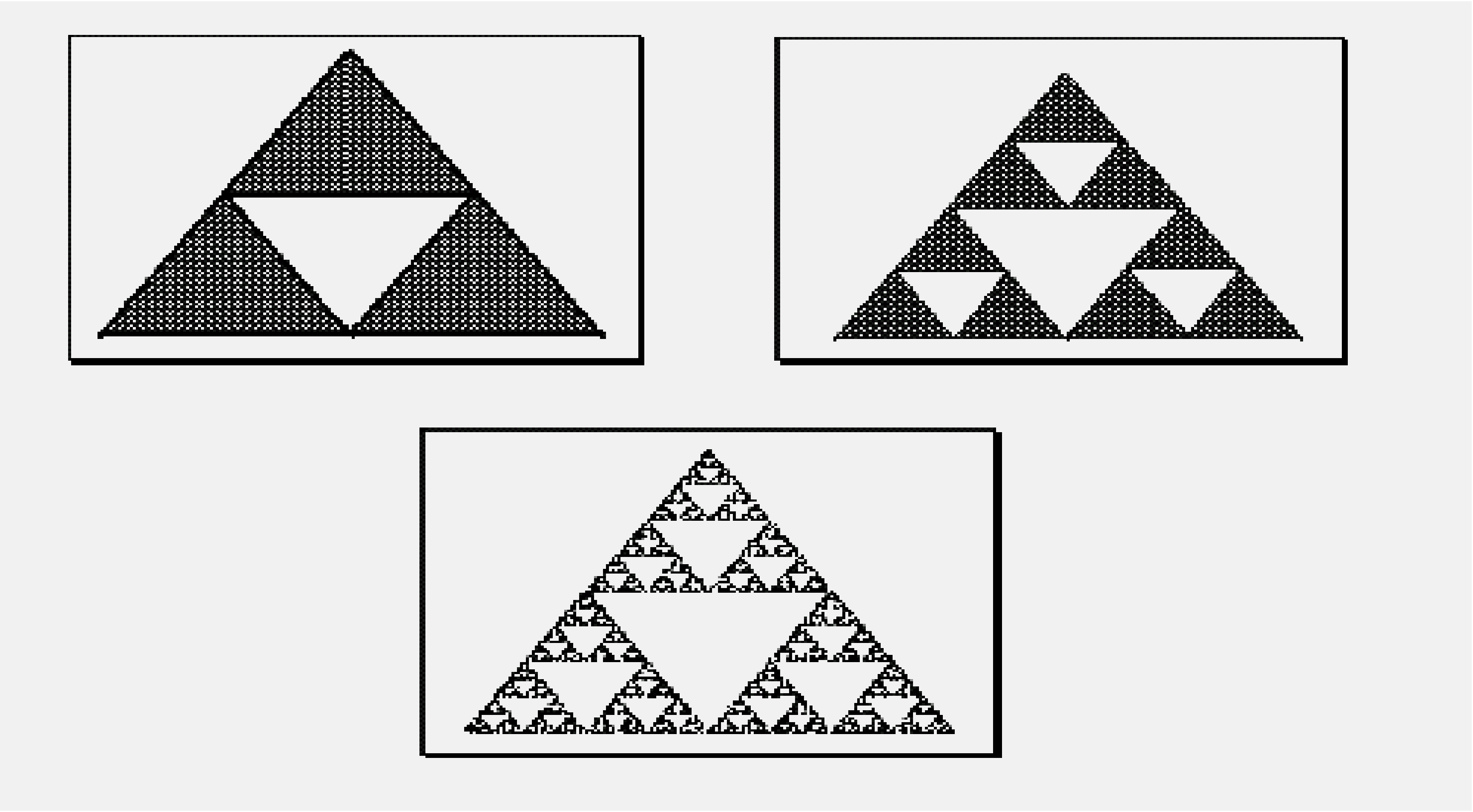
$$L_0 = 1, \quad L_1 = 4/3, \quad L_2 = 4^2/3^2, \text{ etc... } L_k \to \infty$$

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

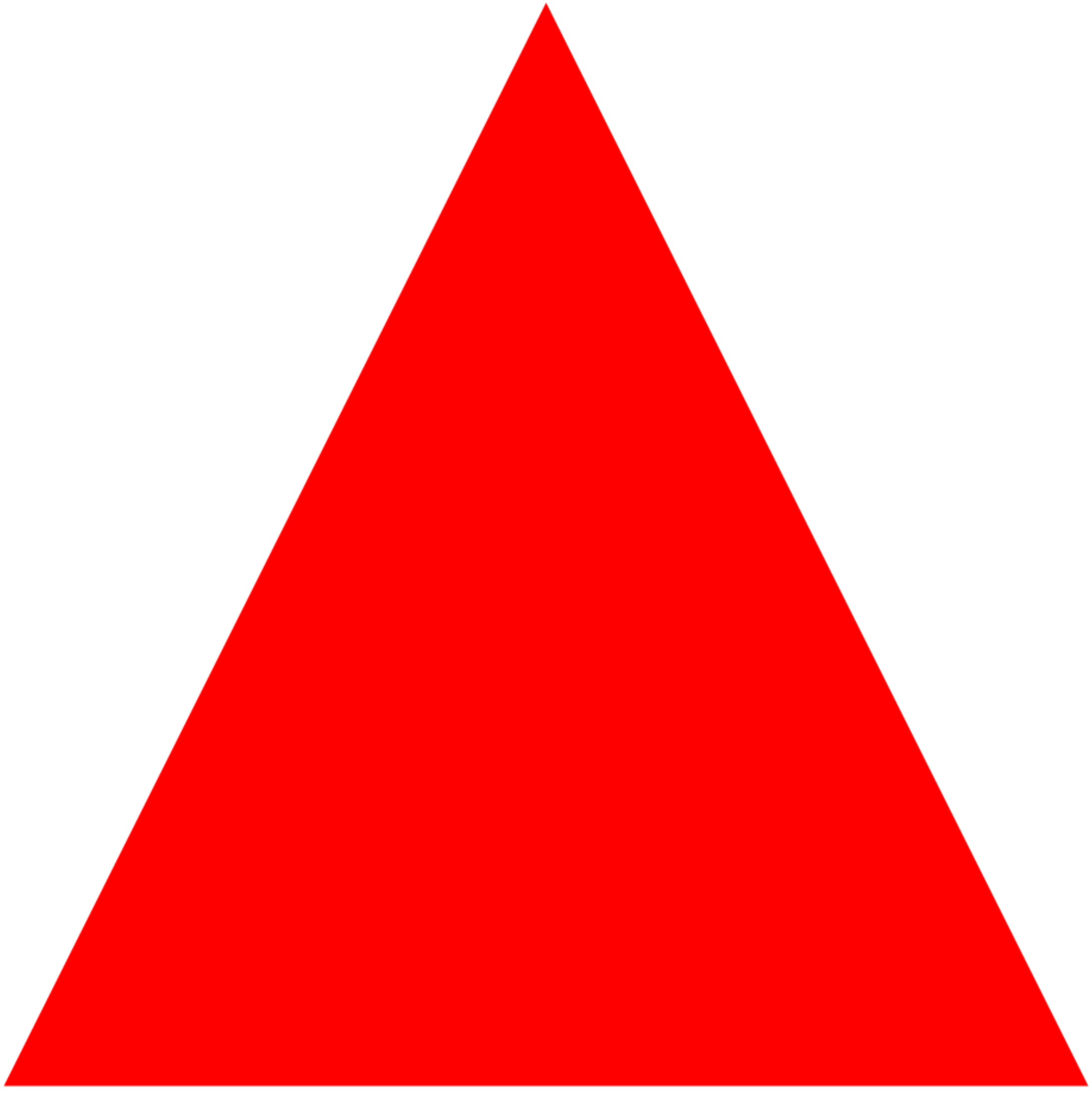
Fractal snowflake

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

Sierpinski triangle (1916)

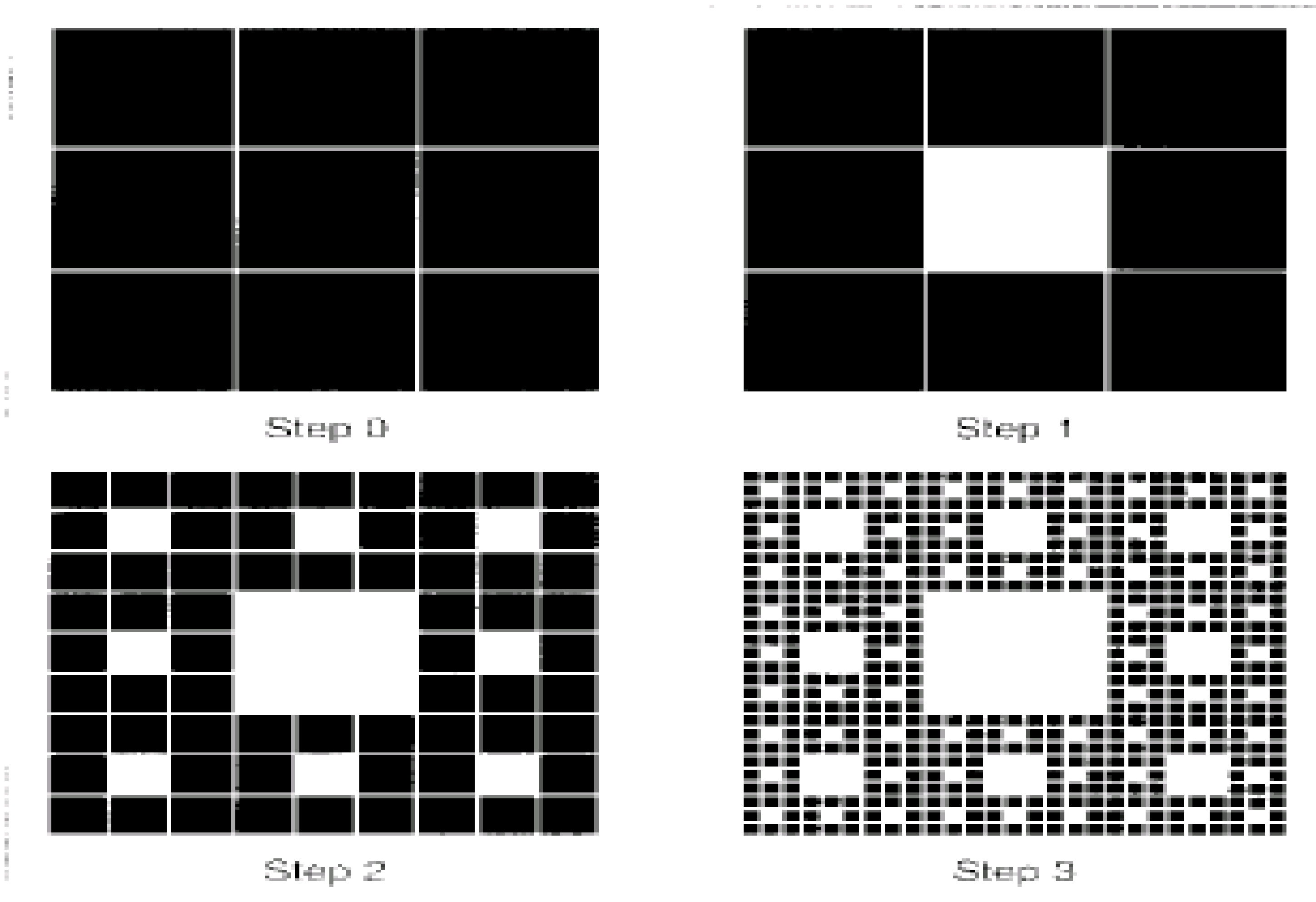


D=log3/log2=1.5849625...



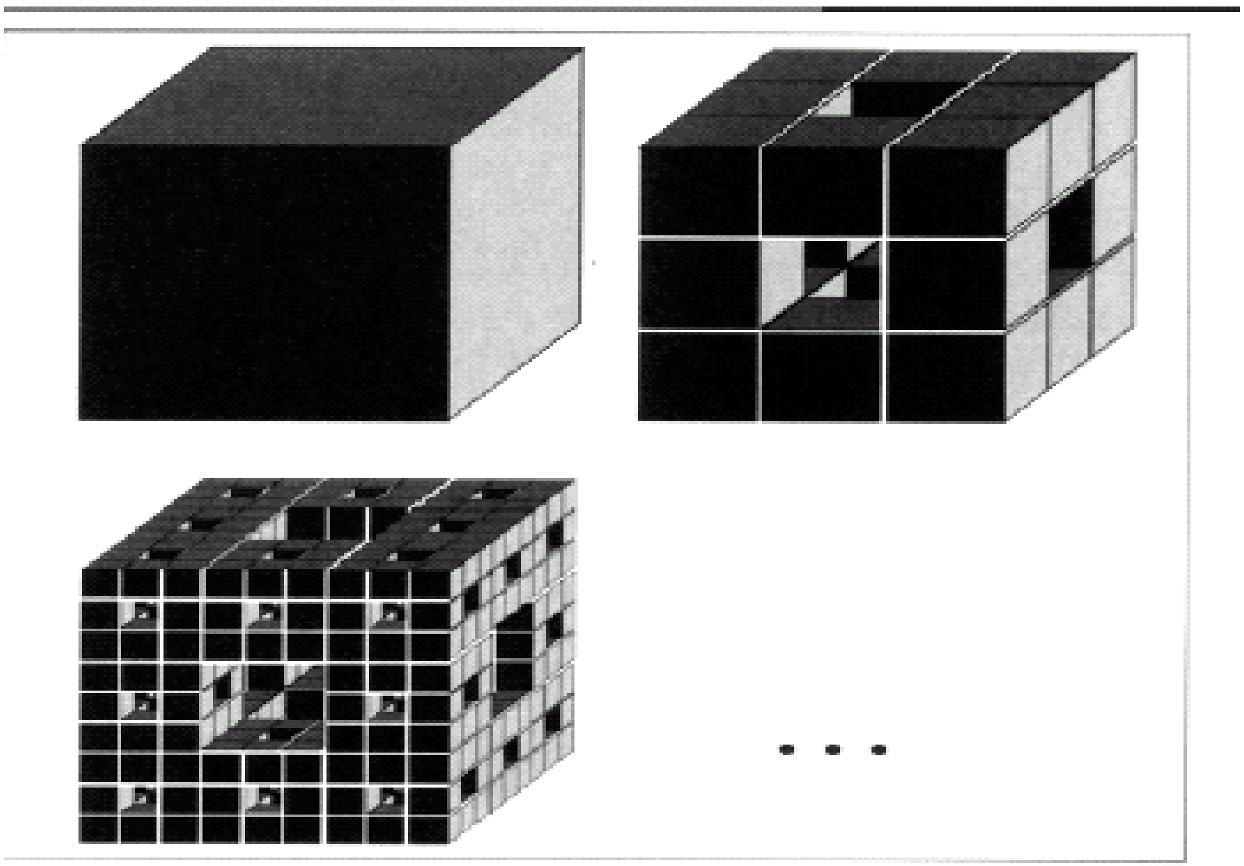
Source: Wikipedia

A fractal carpet (zero area)

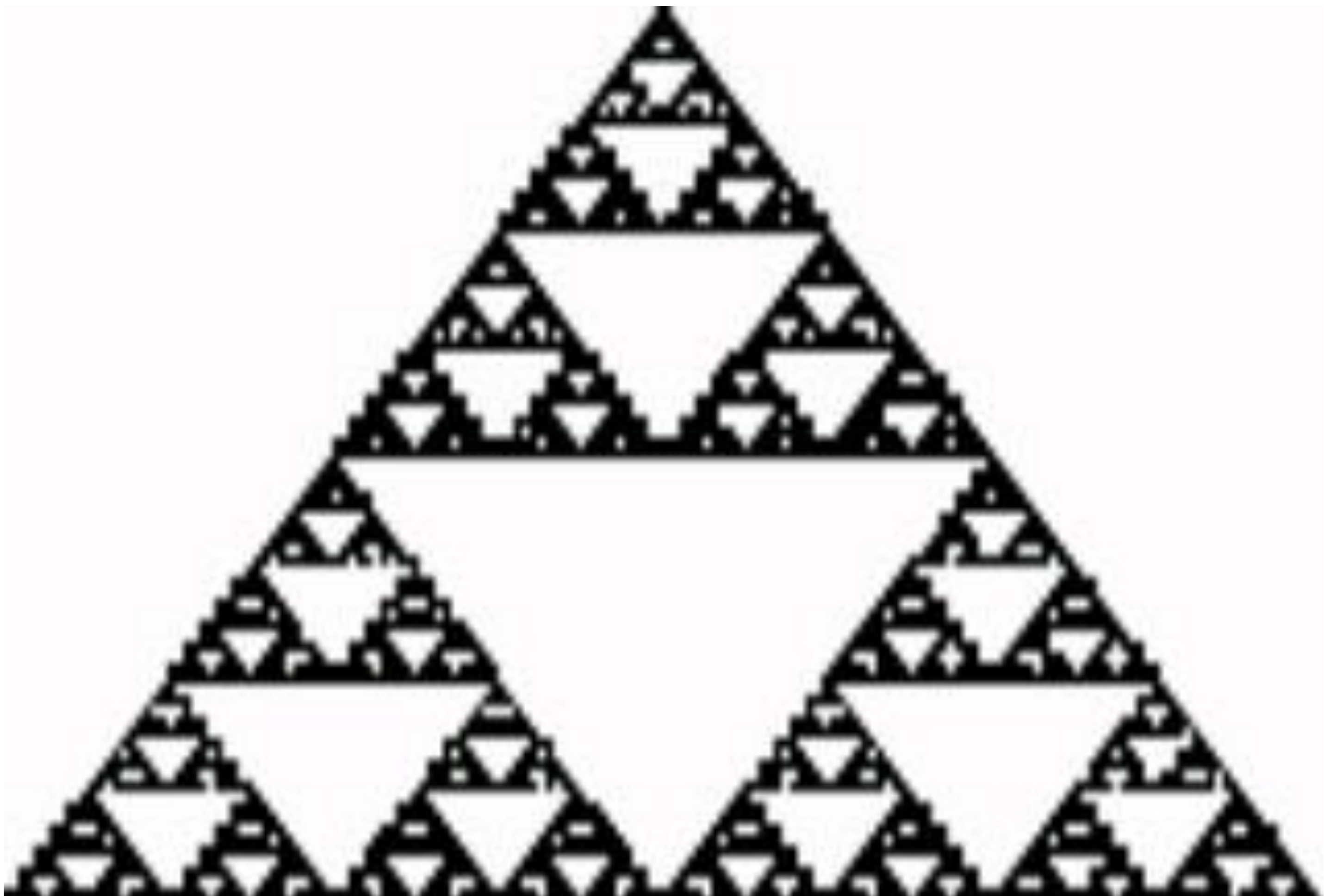


D=3log2/log3=1.892789...

A fractal sponge

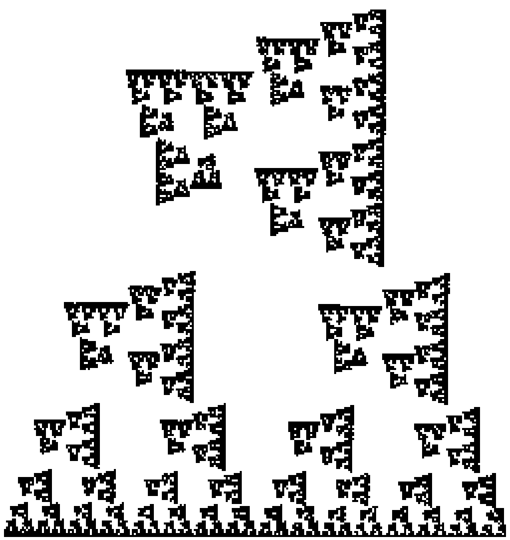


Zooming in

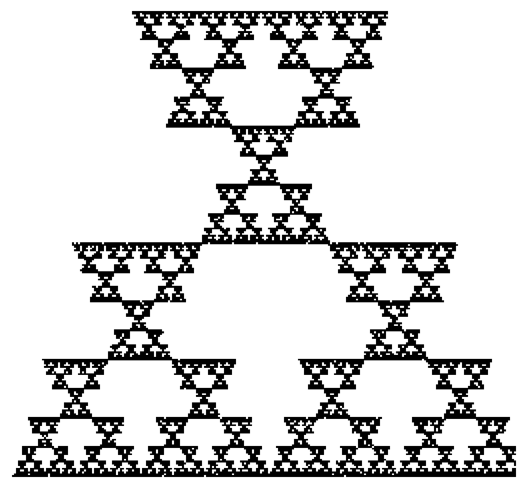


Changing parameters

- 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

$X=[0,1]^2$

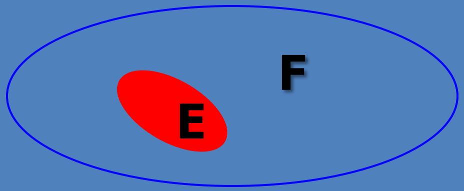
$d((x,y),(x',y'))=|x-x'|+|y-y'|$ Manhattan metric

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

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

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

$d(E,F)>0$



$d(F,E)=0$

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

$\rightarrow$ 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, $0 \le s_i < 1$

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

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

Then $\mathcal{W}$ contracts the Hausdorff metric h with Lipschitz constant $s = \max_{1 \le i \le N} s_i$. 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. $\mathcal{F} = \{w_1, \ldots, w_N\}$ are

- Similarities $\Rightarrow$ the attractor $\mathcal{A}$ will be said self-similar
- Affine maps $\Rightarrow$ the attractor $\mathcal{A}$ will be said self-affine
- Conformal maps (i.e. their derivative is a similarity) then the attractor $\mathcal{A}$ 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 $\bigcup_i w_i(U)$ is an open subset of U then the dimension d of the attractor $\mathcal{A}$ 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 $\mathcal{T}$ can one find an i.f.s $\mathcal{F}$ such that the corresponding attractor $\mathcal{A}$ is ε-close to $\mathcal{T}$ w.r.t. the Hausdorff distance h?

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

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

then

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

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

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 !

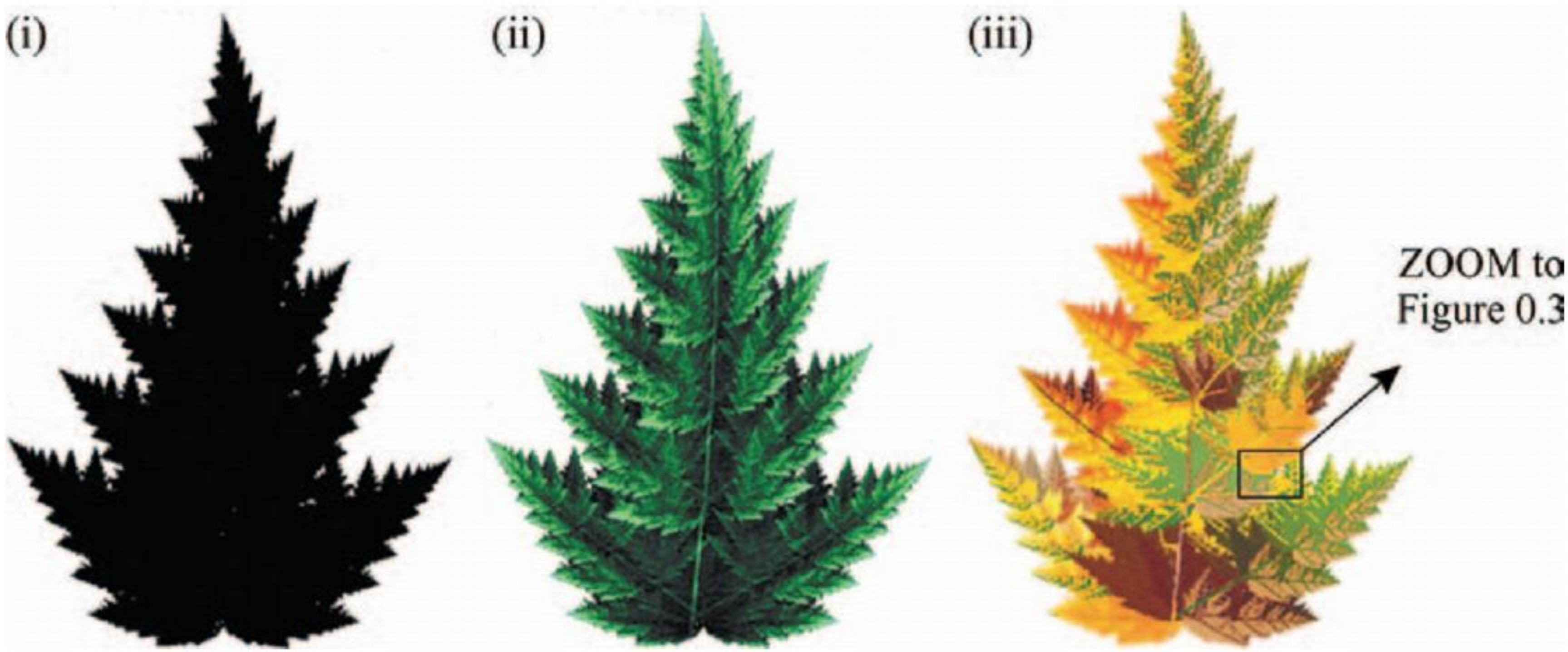
The top-selling multimedia encyclopedia Encarta, published by Microsoft Corporation, includes on one CD-ROM seven thousand color photographs which may be viewed interactively on a computer screen. The images are diverse; they are of buildings, musical instruments, people's faces, baseball bats, ferns, 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)

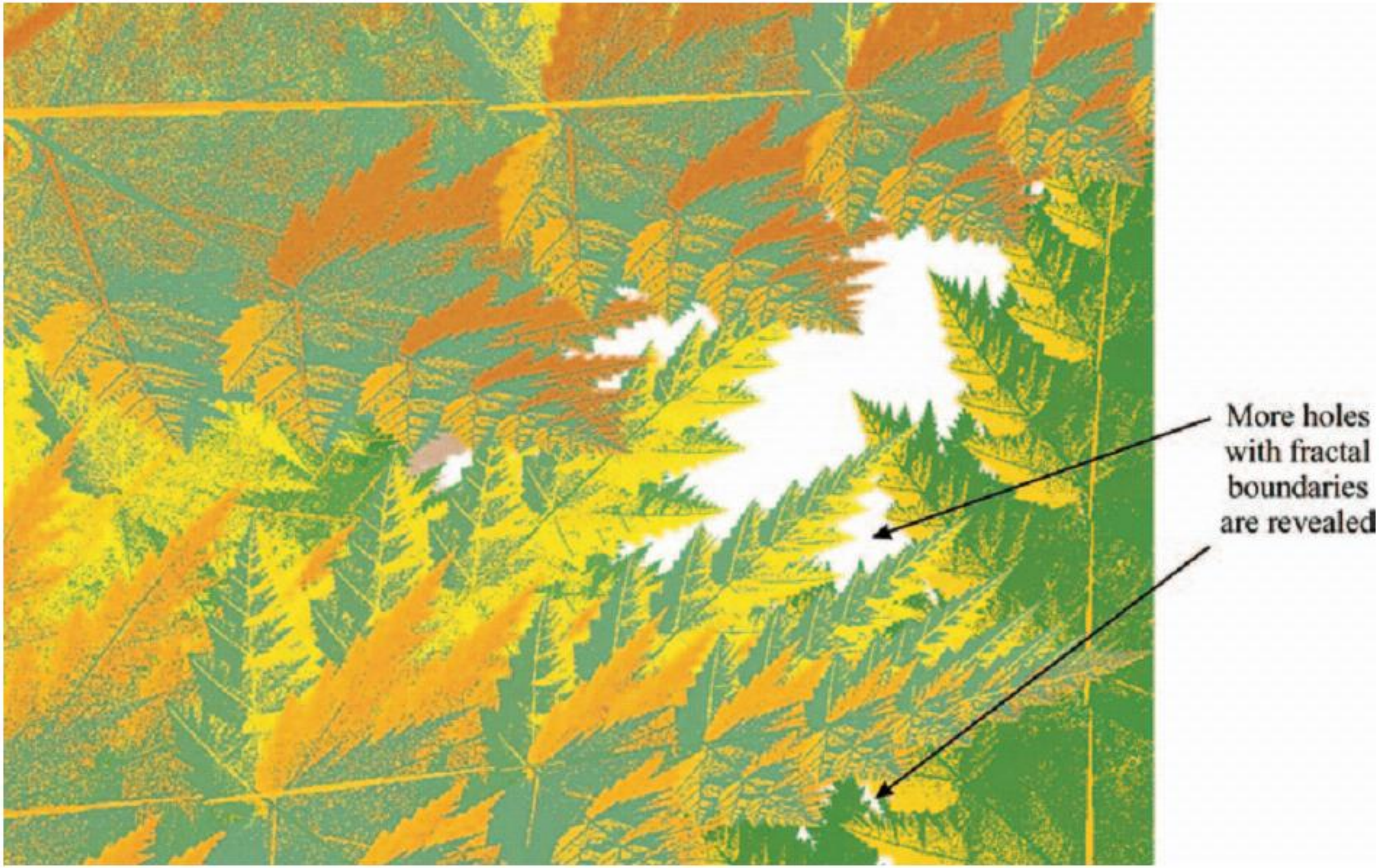


the measure attractor and (iii) the

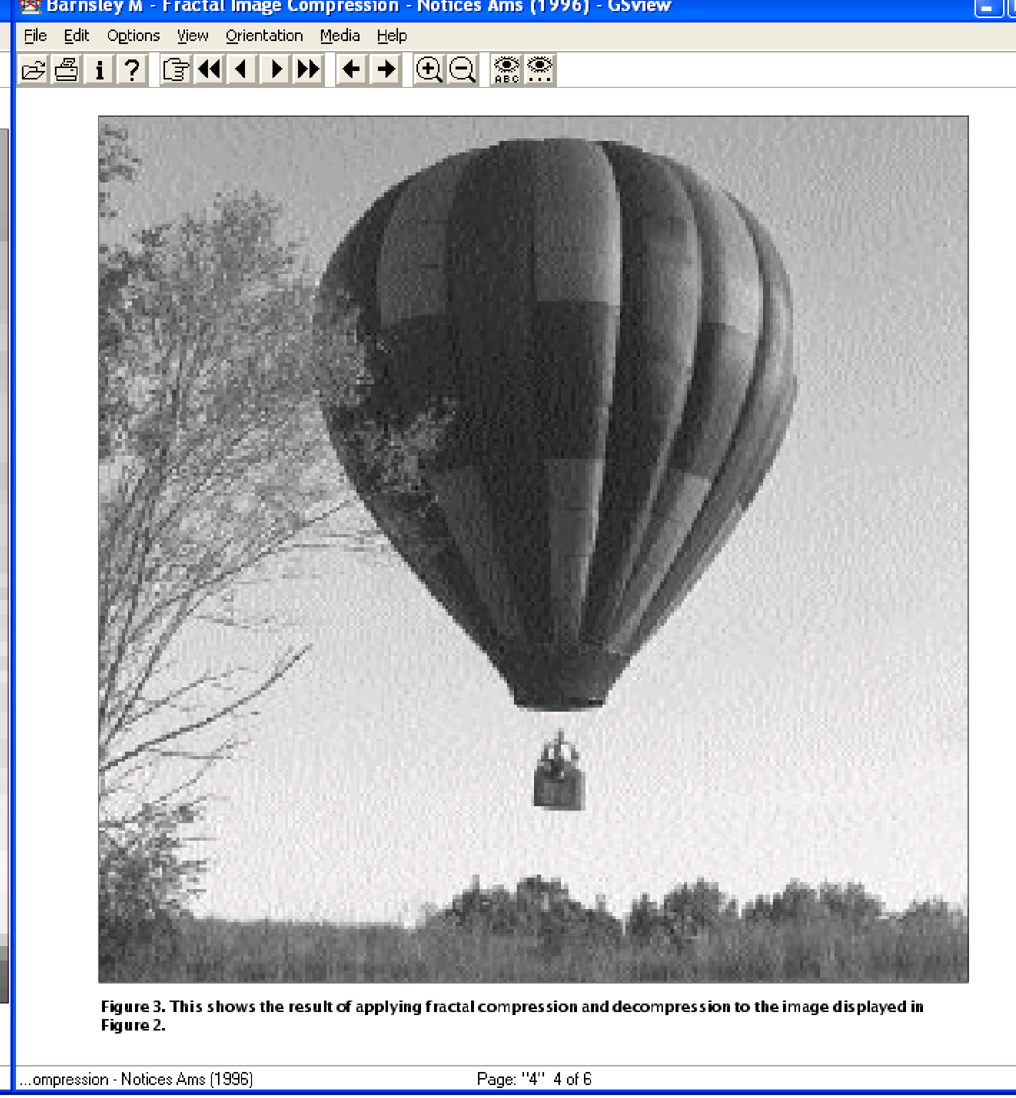
$$f_n(x, y) = \left(\frac{a_n x + b_n y + c_n}{g_n x + h_n y + j_n}, \frac{d_n x + e_n y + k_n}{g_n x + h_n y + j_n} \right)$$

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	−6.1	−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

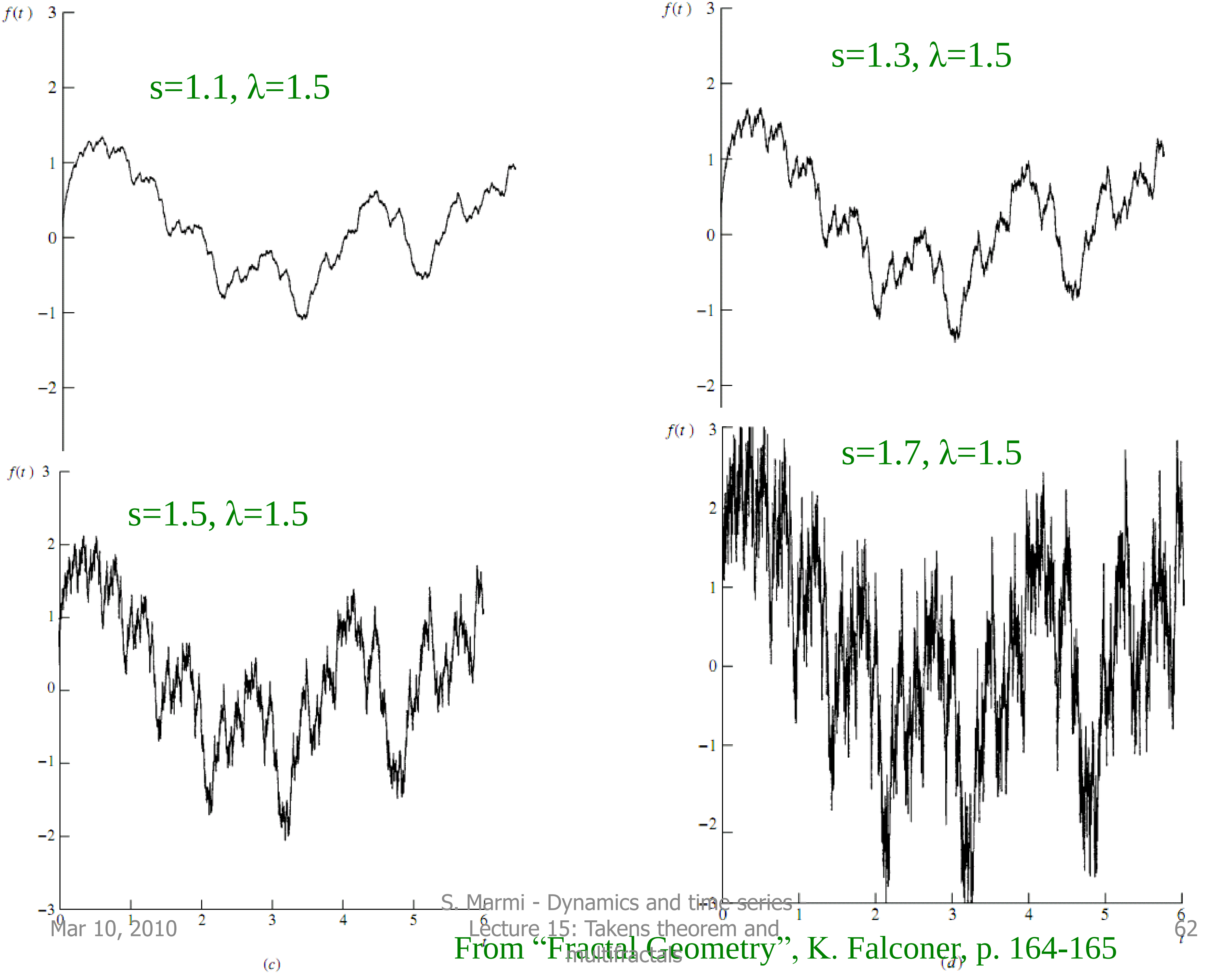
Fractal graphs of functions

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 graphs and i.f.s.

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

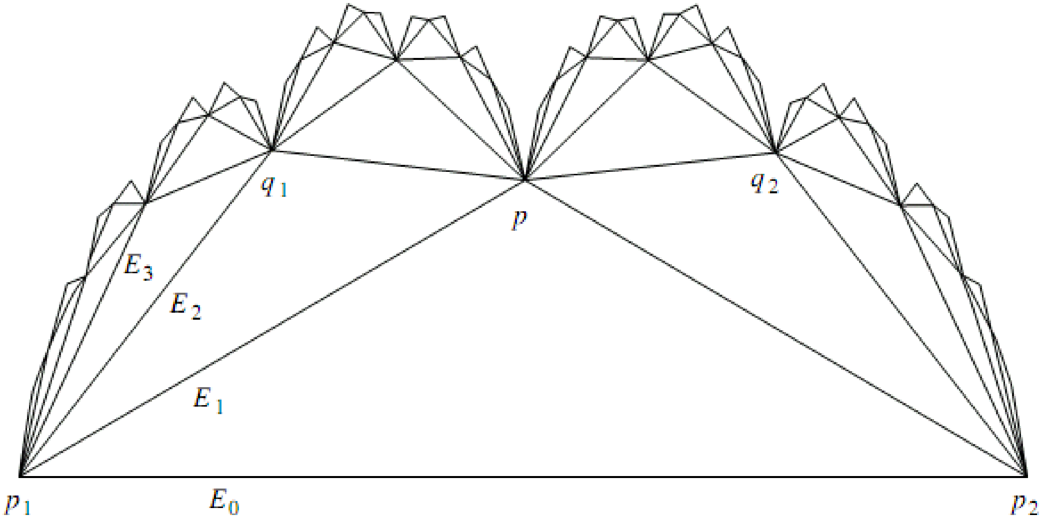


Figure 11.3 Stages in the construction of a self-affine curve F. The affine transformations S_1 and S_2 map the generating triangle p_1pp_2 onto the triangles p_1q_1p and pq_2p_2, respectively, and transform vertical lines to vertical lines. The rising sequence of polygonal curves $E_0, E_1, \ldots$ 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

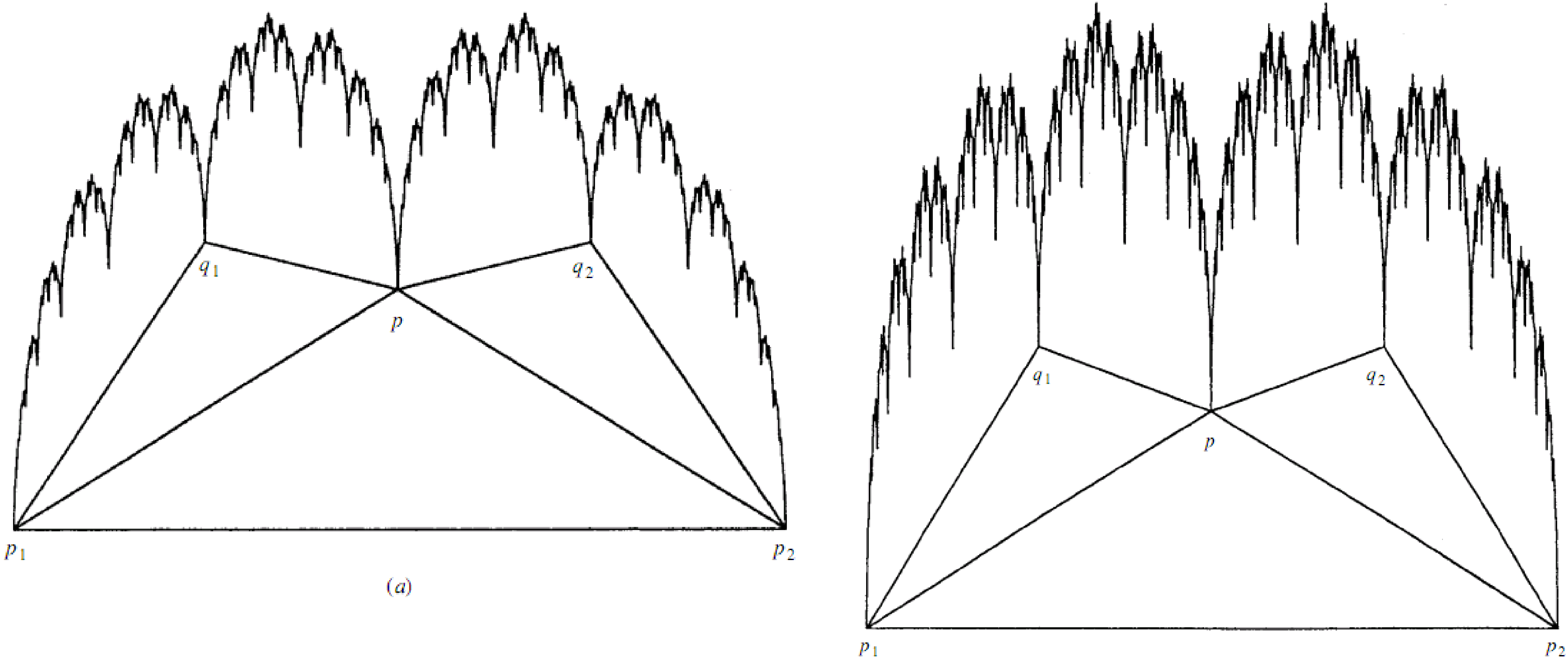
$$1/m < c_i < 1 \tag{11.9}$$

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

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

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



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

$\mathcal{W}(E) = \mathrm{U}_1\, w_i(E)$

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

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

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→ **R**

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

It is easy to verify that

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

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

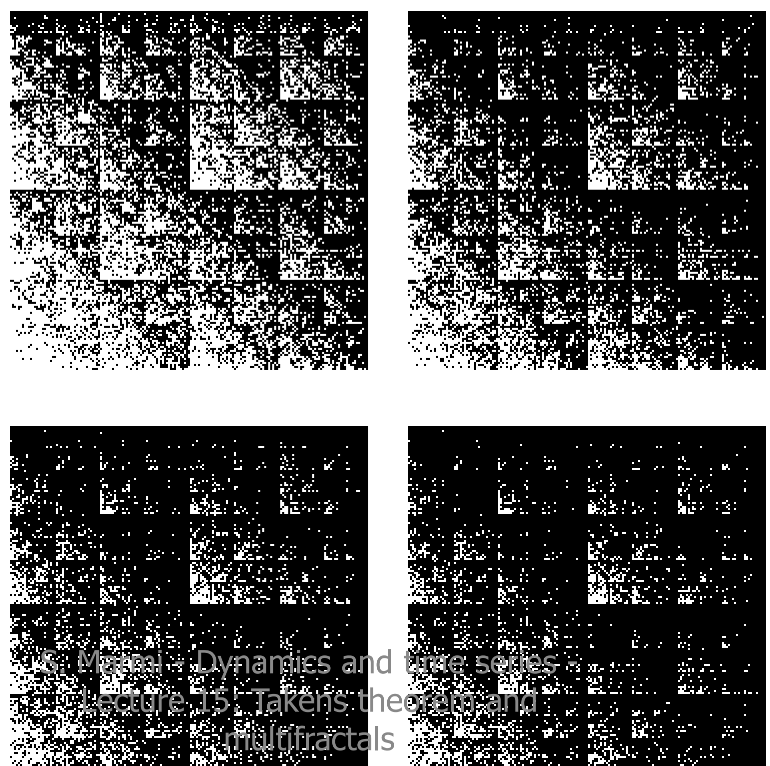
$T_3(x, y) = (x/2, y/2) + (0, 1/2)$	$T_4(x, y) = (x/2, y/2) + (1/2, 1/2)$
$T_1(x, y) = (x/2, y/2)$	$T_2(x, y) = (x/2, y/2) + (1/2, 0)$

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.



Multifractals

Variable Probability Histograms

The probabilities of applying each transformation represent the fraction of the total number of iterates in the region determined by the transformation.

With the IFS and probabilities of the last example, in a typical picture about 0.1 of the points will lie in the square with address 1, and about 0.3 of the points will lie in each of the squares with address 2, 3, and 4.

Arguing in the same way, about 0.01 = 0.1*0.1 of the points will lie in the square with address 11, about 0.03 = 0.1*0.3 of the points will lie in the square with address 12, and so on.

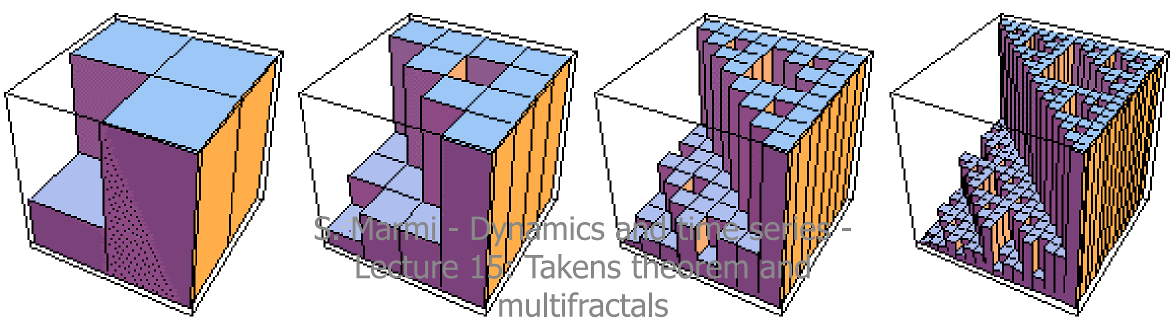
.3	.3
.1	.3

.09	.09	.09	.09
.03	.09	.03	.09
.03	.03	.09	.09
.01	.03	.03	.09

Higher iterates are easier to understand visually.

Here we show the first four generations, with the height of the box in a region representing the fraction of the points in that region.

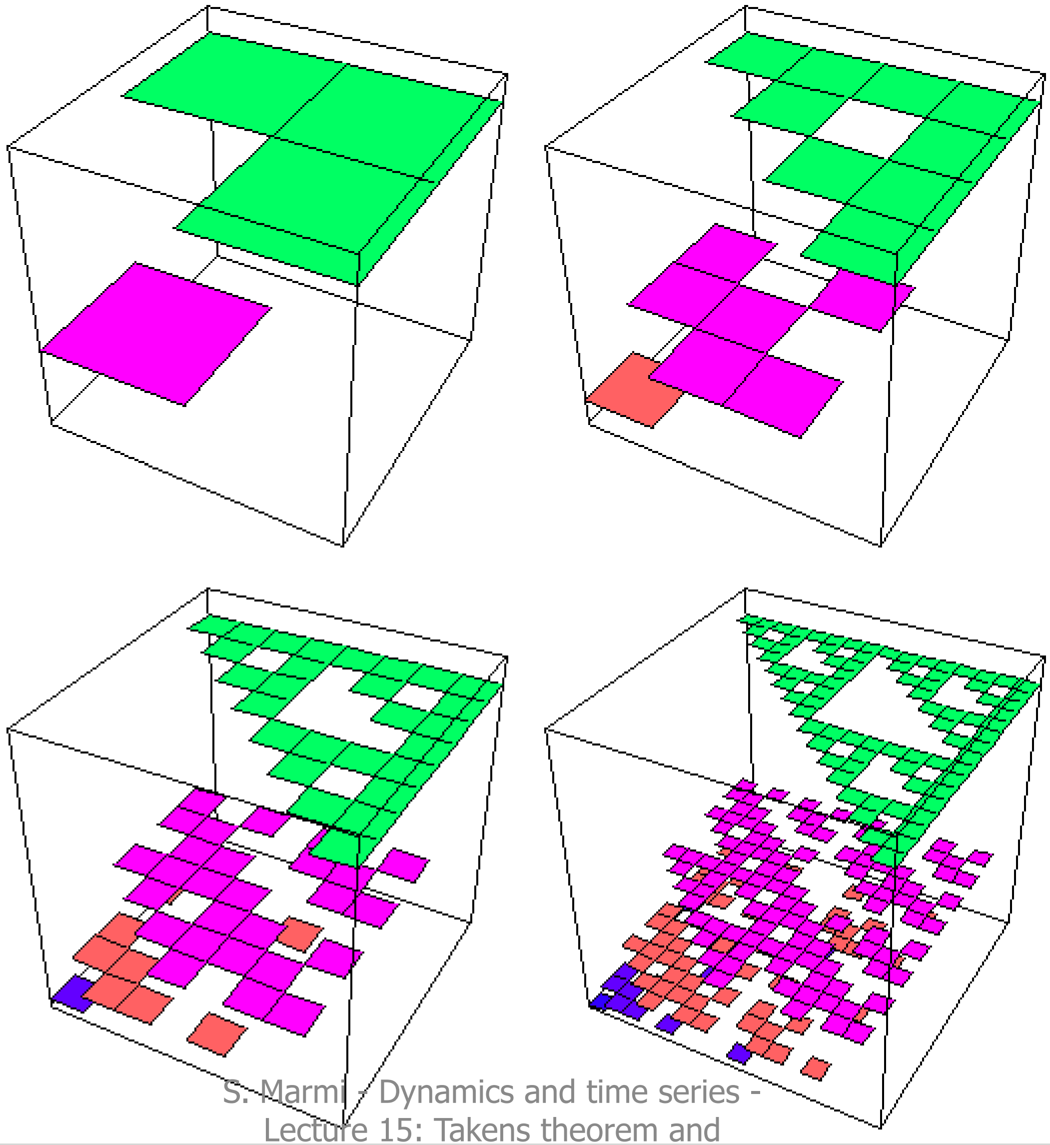
All the pictures have been adjusted to have the same height, whereas square 4 has 0.3 of the points, square 44 has 0.09 of the points, square 444 has 0.027 of the points, and so on.



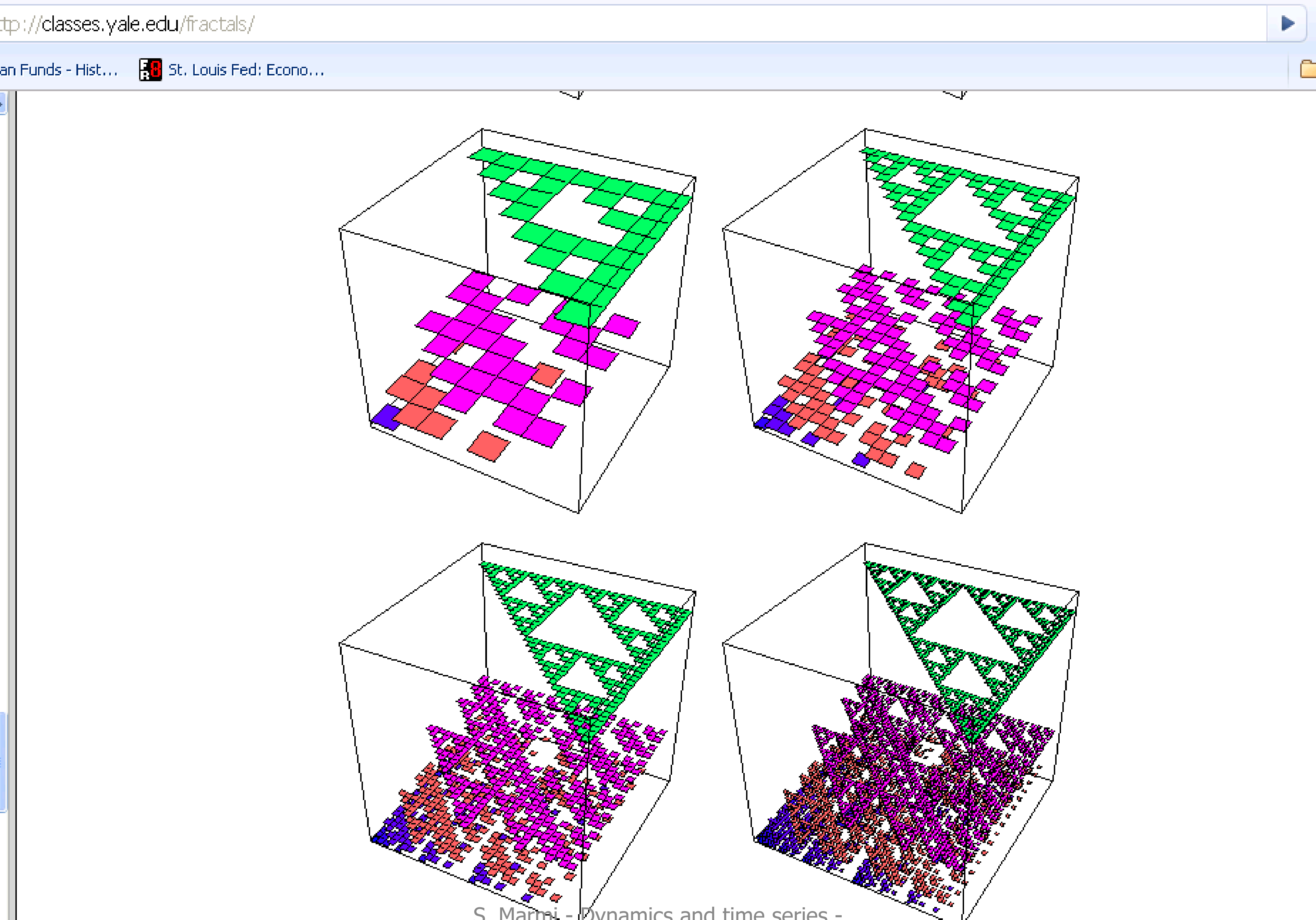
tp://classes.yale.edu/fractals/

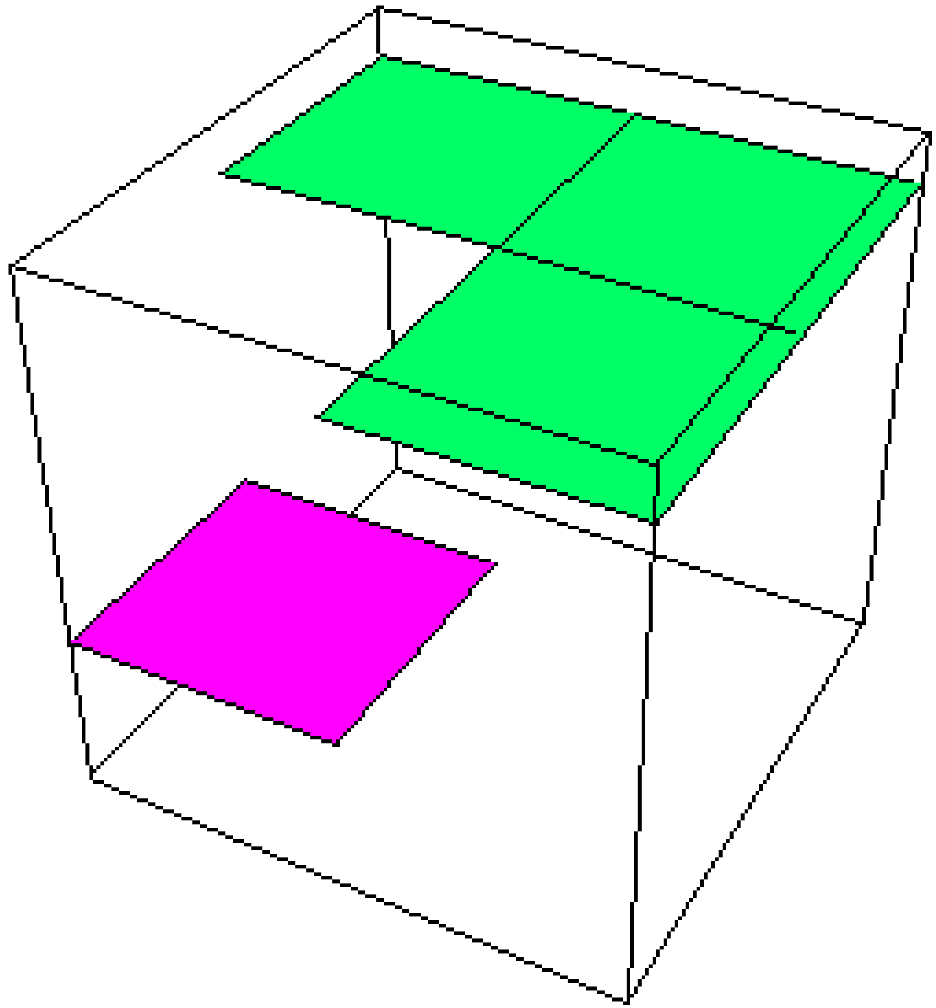
an Funds - Hist... St. Louis Fed: Econo... Altri P

So again the height represents the fraction of the points landing in that region.



ct....gif Mostra tutti i download...


tp://classes.yale.edu/fractals/
an Funds - Hist...
St. Louis Fed: Econo...



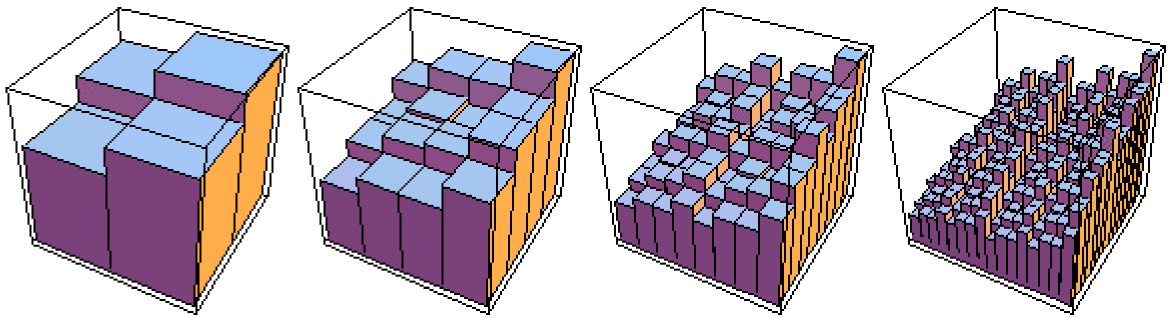
http://classes.yale.edu/fractals/MultiFractals/MFGaskSect/MFGaskSectMv.gif

Different Probabilities, Another Example

In this example, we introduce more variability in the probabilities:

$$p_1 = 0.2,\ p_2 = 0.25,\ p_3 = 0.25, \text{ and } p_4 = 0.3.$$

Among other things, the number of values of the probabilities of regions increases more rapidly.



Smaller regions have smaller probabilities; if these graphs weren't rescalled vertically they would appear to become closer and closer to a flat surface of heig 0. Click here for an animation of the first four iterates, all drawn to the same vertical scale.

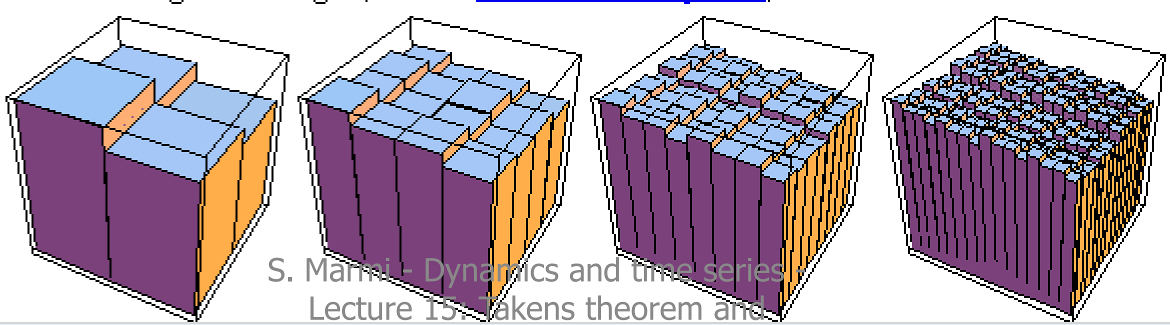
For each region we expect that

$$\text{prob scales as (side length)}^{\text{some power}}$$

So instead of letting the height of the graph represent the probability of the region, now we assign height Log(prob)/Log(side length) to the region.

Because the probability measures the fraction of the points that occupy a region, we think of this ratio as a dimension.

Being viewed at the resolution of the side length of the region, this is a coarse Holder exponent; it is also called the coarse dimension.



Multifractals

Local Holder Exponents

Taking limits as the side length of the regions go to zero, the coarse Holder exponent can be refined to the local Holder exponent (or *roughness*) at (x, y) is

$$d_{loc}(x,y) = \lim_{n \to \text{infinity}} \text{Log}(\text{Prob}(i_1 \ldots i_n))/\text{Log}(2^{-n})$$

where $\text{Prob}(i_1 \ldots i_n)$ is the probability $pr(i_1)^* \ldots {}^*pr(i_n)$, if (x,y) lies in the square with address $i_1 \ldots i_n$.

The value for a square of finite length address is called the coarse Holder exponent. So the local Holder exponent of a point (x, y) is the limit as N -> infinity of the coarse Holder exponents of the length N address squares containing (x, y).

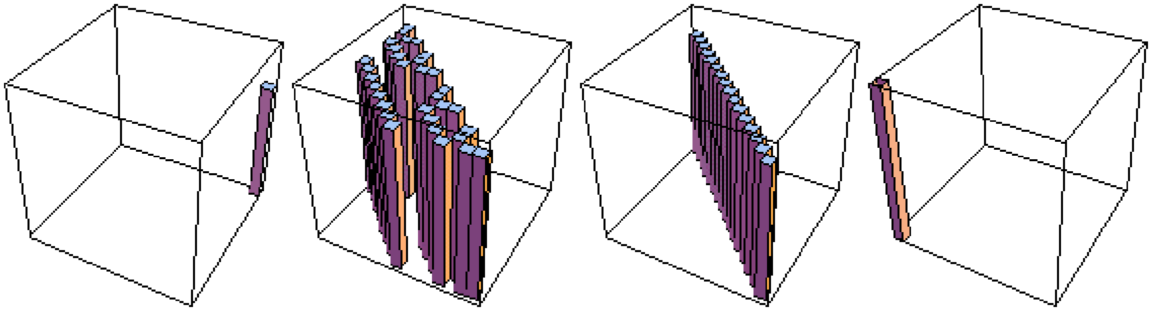
Now define

$$E_{alpha} = \{(x, y): d_{loc}(x, y) = alpha\},$$

the collection of all points of the fractal having local Holder exponent alpha.

As alpha takes on all values of the local Holder exponent, we decompose the fractal into these sets E_{alpha}.

Here are examples, E_{alpha} (alpha = column height) for the lowest value of alpha (on the left), two intermediate values, and the highest value.

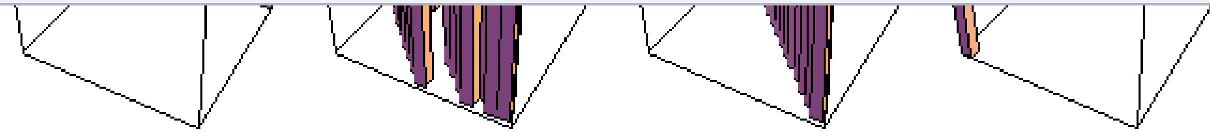


Click here for an animation scanning through all the values of alpha, from lowest to highest, resolved to boxes have side length $1/2^4$.

Because each local Holder exponent alpha is the exponent for a power law, a multifractal is a process exhibiting scaling for a range of different power laws.

The multifractal structure is revealed by plotting $\dim(E_{alpha})$ as a function of alpha.

(In general, a dimension more subtle than the box-counting dimension must be used. We ignore this complication here.)



Click here for an animation scanning through all the values of alpha, from lowest to highest, resolved to boxes have side length $1/2^4$.

Because each local Holder exponent alpha is the exponent for a power law, a multifractal is a process exhibiting scaling for a range of different powe

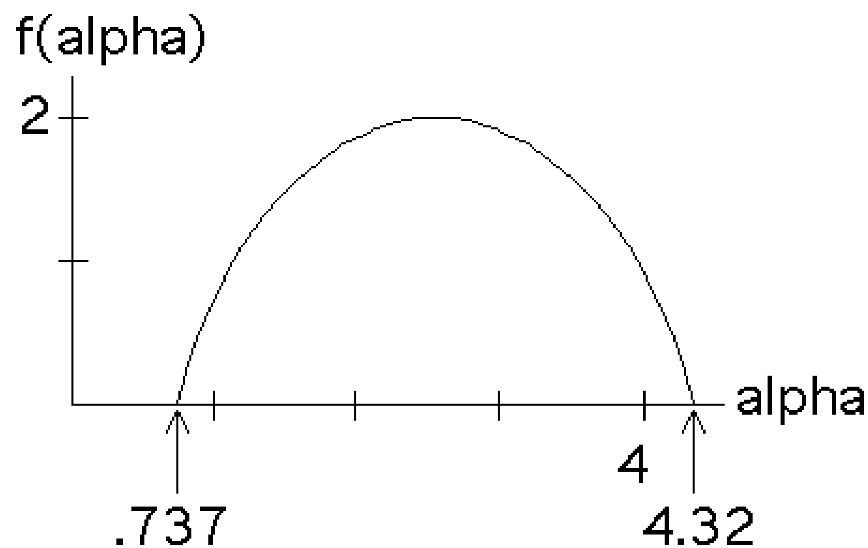
The multifractal structure is revealed by plotting dim(E_{alpha}) as a function of alpha.

(In general, a dimension more subtle than the box-counting dimension must be used. We ignore this complication here.)

This graph is called the f(alpha) curve.

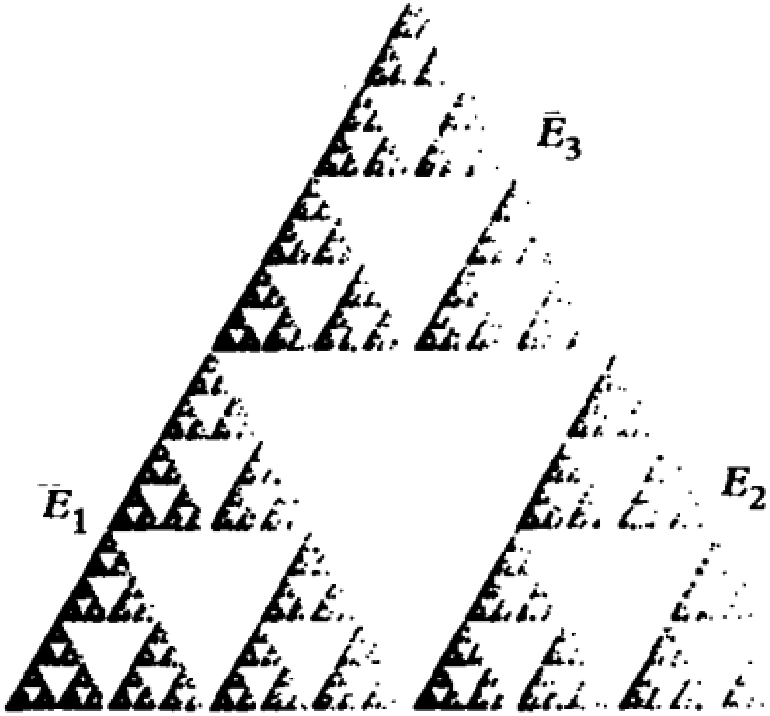
Here is the f(alpha) curve for the example with $p_1 = 0.2$, $p_2 = p_3 = 0.25$, and $p_4 = 0.3$.

At least in this example, sets E_{alpha} for the lowest and highest values of alpha reduce to points in the limit, hence have dimension f(alpha) = 0. This is represented in the left and right endpoints of the curve lying on the x-axis.



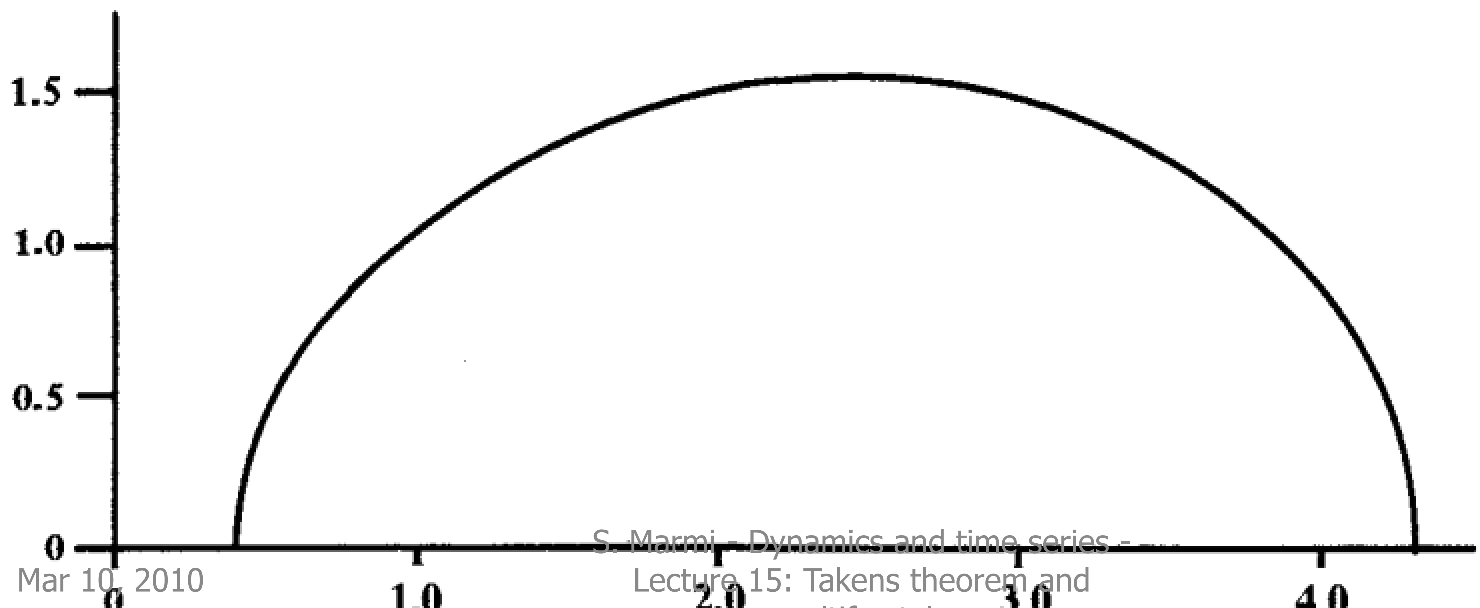
This result is derived under more general conditions in a later section.

Return to Multifractals.



K. Falconer, Techniques in Fractal geometry

P=(0.8,0.05,0.15)



The Legendre transform of f(α)

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

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

The dimension d of the attractor $\mathcal{A}$ is the solution of the equation

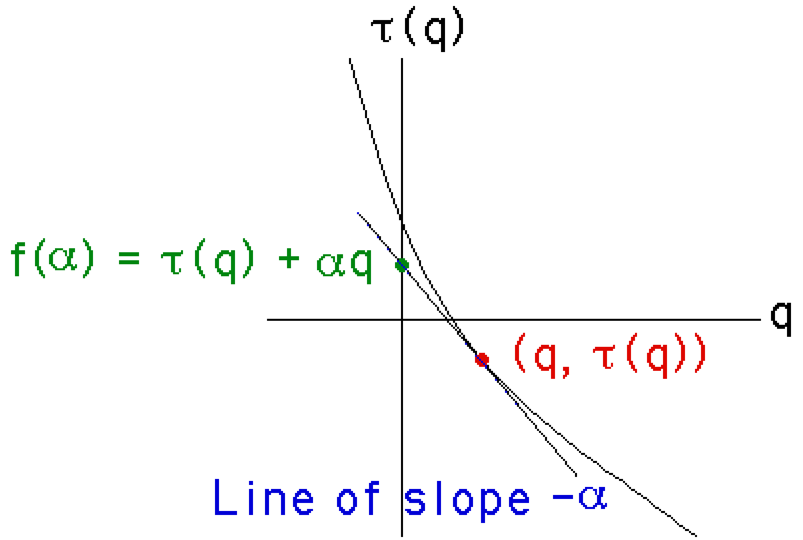
$$s_1^d + s_2^d + \ldots + s_N^d = 1$$

The singularity spectrum $\alpha \to f(\alpha)$ of a probabilistic i.f.s. is the Legendre transform of the function $q \to \tau(q)$ obtained solving the functional equation

$$p_1^q s_1^{\tau(q)} + p_2^q s_2^{\tau(q)} + \ldots + p_N^q s_N^{\tau(q)} = 1$$

Defining $f(\alpha)$

For each point $(q, \tau(q))$ say the slope of the tangent line is $-\alpha$. That is, $\alpha = -d\tau/dq$.



This tangent line passes through the point $(q, \tau(q))$ and the point $(0, y)$. Consequently,

$$-\alpha = (y - \tau(q))/(0 - q)$$

Solving for y,

$$y = q\cdot\alpha + \tau(q)$$

Call this y-value $f(\alpha)$:

$$f(\alpha) = q\cdot\alpha + \tau(q)$$

Return to Multifractals from IFS.

The singularity spectrum $\alpha\rightarrow f(\alpha)$ of a probabilistic i.f.s. is the Legendre transform of the function $q\rightarrow\tau(q)$