Information-Theoretic Characterizations of Self-Similar Geometry

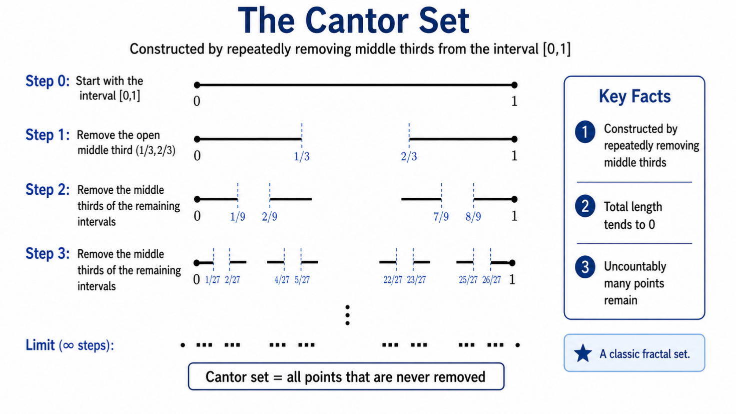
Midnight Sun Summit in Mathematics and Engineering
Narvik, Norway
2026

Joel Ratsaby

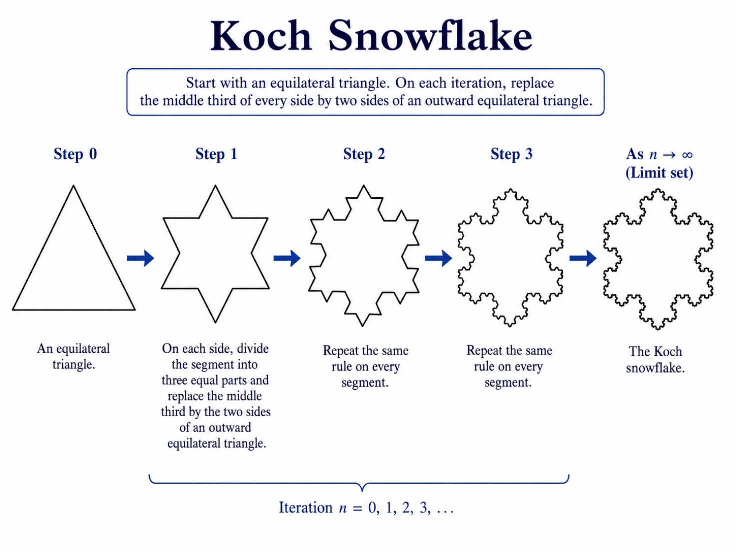
June 19, 2026

Overview

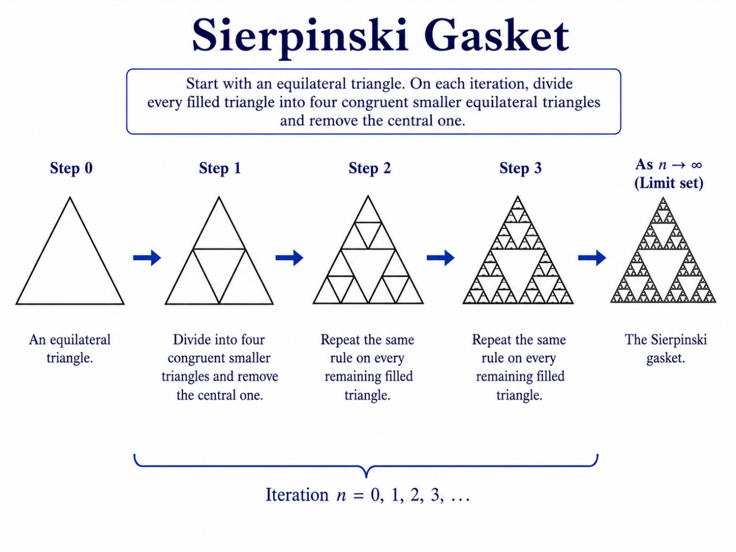
- Fractal sets, self-similar
- Generated by simple recursive rules
- For example:



Overview


Koch Snowflake
Start with an equilateral triangle. On each iteration, replace
the middle third of every side by two sides of an outward equilateral triangle.
Step 0
Step 1
Step 2
Step 3
As $n \to \infty$
(Limit set)
An equilateral triangle.
On each side, divide the segment into three equal parts and replace the middle third by the two sides of an outward equilateral triangle.
Repeat the same rule on every segment.
Repeat the same rule on every segment.
The Koch snowflake.
Iteration $n = 0, 1, 2, 3, \ldots$

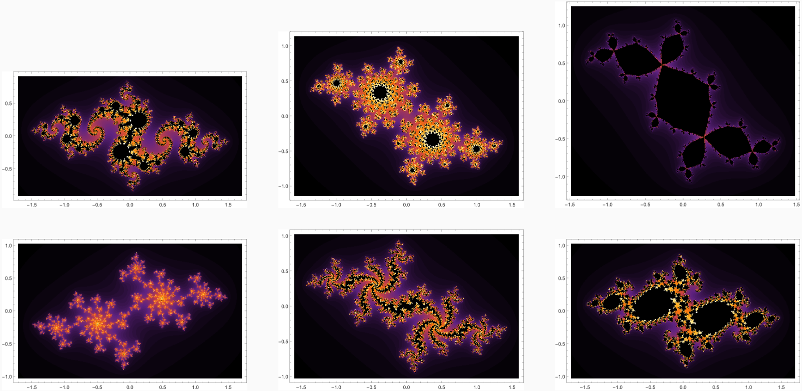
Overview


Sierpinski Gasket
Start with an equilateral triangle. On each iteration, divide every filled triangle into four congruent smaller equilateral triangles and remove the central one.
Step 0
Step 1
Step 2
Step 3
As $n \to \infty$ (Limit set)
An equilateral triangle.
Divide into four congruent smaller triangles and remove the central one.
Repeat the same rule on every remaining filled triangle.
Repeat the same rule on every remaining filled triangle.
The Sierpinski gasket.
Iteration $n = 0, 1, 2, 3, \ldots$

Overview

As we zoom into a fractal set, it retains its complex structure (does not become simpler)

For example, filled-Julia sets:



Fractal Set

A fractal set

- has low Kolmogorov (algorithmic) complexity (length of the shortest program that outputs the set) since it is generated by a short recursive rule
- has high information content when observed as data (without knowing its generating rule)

Motivation

Is there another measure (not algorithmic complexity) that can distinguish different levels of complexity of fractal sets?

To explain this

Develop an information-based set-complexity that quantifies its information richness

Background

ϵ-entropy of a set

Let (X, d) be a compact metric space and let $A \subseteq X$. For $x \in X$ and $\varepsilon > 0$, an open ball is defined

$$B(x, \varepsilon) := \{y \in X : \; d(x, y) < \varepsilon\}.$$

Kolmogorov ε-entropy

For $\varepsilon > 0$, define the covering number of A by ε-balls as

$$N_\varepsilon(A) := \min \left\{ m : \; A \subseteq \bigcup_{i=1}^{m} B(x_i, \varepsilon), \; x_i \in X \right\}.$$

The Kolmogorov ε-entropy of A is

$$H_\varepsilon(A) := \log N_\varepsilon(A).$$

Fractal dimension of a set

Upper and lower box dimensions

The upper and lower box-counting dimensions, equivalently the upper and lower entropy dimensions, are

$$\overline{\dim}_{\mathrm{B}} A = \overline{\dim}_{\mathrm{ent}} A := \limsup_{\varepsilon \downarrow 0} \frac{H_\varepsilon(A)}{\log(1/\varepsilon)},$$

$$\underline{\dim}_{\mathrm{B}} A = \underline{\dim}_{\mathrm{ent}} A := \liminf_{\varepsilon \downarrow 0} \frac{H_\varepsilon(A)}{\log(1/\varepsilon)}.$$

When these two quantities coincide, their common value is denoted

$$\dim_{\mathrm{B}} A = \dim_{\mathrm{ent}} A.$$

(for compact set A, this is called fractal dimension (Barnsley, 1993).)

A related dimension, is the Hausdorff dimension (in general, difficult to estimate)

Our approach

Consider a compact set A in the complex plane $\mathbb{C}$ (or the real plane $\mathbb{R}$).
To define an information-based complexity of a set A, we

- construct a binary encoding of A
- approximate A to an arbitrary finite precision
- extend a classical information-theoretic notion of compression ratio of binary sequences, to compression ratio of A,
- define information density of A.
- estimate the information density and compression ratio of some example sets

The approach can be easily extended to n-dimensional Euclidan space

Binary string

- x, finite-length binary string
- $\ell(x)$ is length of x (number of bits)
- m, an integer,
- $x(m)$ is binary string representing the signed magnitude of m using $\lfloor \log |m| \rfloor + 1$ bits for its magnitude and one bit for its sign.
- Example: If $m = -12$ then $x(m) = 11100$
- 1^l, a string of l bits of value 1. For two strings x, y we let xy denote their concatenation. For a binary string x of length $\ell(x)$ denote by

  $$s(x) := 1^{\ell(x)}0x$$

  a prefix codeword of the string x.
- $s(x)$ is a self-delimiting codeword for x

Integer

- $s(m) := 1^{\ell(x(m))}0x(m)$, a string that encodes integer m.
- complexity of m is defined as $\ell(m) := \ell(s(m)) = 2\lfloor \log |m| \rfloor + 5$,

Complex number

- Let $p, q, u, v \in \mathbb{Z}$, where $q \neq 0,\ v \neq 0$.
- rational complex number is denoted by

$$z := \frac{p}{q} + i\left(\frac{u}{v}\right).$$

- $\hat{\mathbb{C}}$ set of rational complex numbers (Gaussian rational field).
- (p, q), the greatest common divisor (GCD) of p and q,
- lowest term representation (or reduced form) of p/q is $\frac{p}{q} = \frac{(p,q)p'}{(p,q)q'} = \frac{p'}{q'}$.
- encode z by a prefix codeword $s(z) := s_\nu(p')s_\nu(q')s_\nu(u')s_\nu(v')$, where $s_\nu(p')$, ..., $s_\nu(v')$ encode p', $\ldots$, v', using ν bits (ν sufficiently large to encode the largest absolute value μ of the four integers)
- The complexity of z is defined as $\ell(z) := \ell(s(z)) = 4\ell(\mu)$,

Encoding

Grid

- F_N, Farey sequence of reduced fractions between 0 and 1 with denominators N or less, in increasing order.
- For instance, for $N = 6$, $F_6 = \left\{ \frac{0}{1}, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \frac{1}{1} \right\}$.
- $F_N^{(ext)}$, extended Farey sequence, consists of elements of F_N, their inverse, their negation, and their negative inverse.
- For instance, for $N = 6$,

$$F_{\mathbf{6}}^{(ext)} = \left\{ -6, -5, -4, -3, -\frac{5}{2}, -2, -\frac{5}{3}, -\frac{3}{2}, -\frac{4}{3}, -\frac{5}{4}, -\frac{6}{5}, -1, -\frac{5}{6}, -\frac{4}{5}, -\frac{3}{4}, -\frac{2}{3}, -\frac{3}{5}, -\frac{1}{2}, -\frac{2}{5}, -\frac{1}{3}, -\frac{1}{4}, -\frac{1}{5}, -\frac{1}{6}, 0, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, 1, \frac{6}{5}, \frac{5}{4}, \frac{4}{3}, \frac{3}{2}, \frac{5}{3}, 2, \frac{5}{2}, 3, 4, 5, 6 \right\}$$

- Farey grid, $\mathcal{G}_N \subset \hat{\mathbb{C}}$, $\mathcal{G}_N := \left\{ z \in \hat{\mathbb{C}} : \Re(z), \Im(z) \in F_N^{(ext)} \right\}$
- Properties: $\mathcal{G}_N \subset \mathcal{G}_{N+1}$, and for $z \in \mathcal{G}_N$, $\ell(z) \leq 4\ell(N)$

Approximation

Approximation

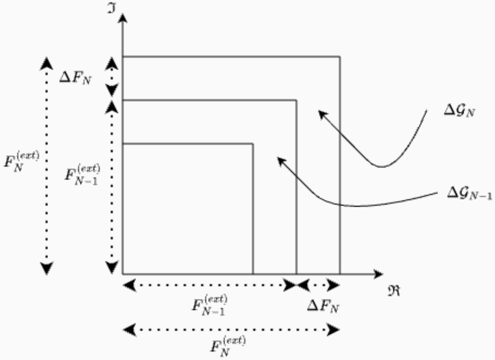
- K, fractal set
- $\hat{K} := K \bigcap \hat{\mathbb{C}}$, rational approximation of K
- $\hat{K}_N := \hat{K} \bigcap \mathcal{G}_N$, a finite approximation of $\hat{K}$
- $\hat{K}_N$ consists of elements $z \in \hat{K}$ of complexity $\ell(z) \leq 4\ell(N)$,
- can encode $z \in \hat{K}_N$ by $\leq 4\ell(N)$ bits
- $\lim_{N\to\infty} \hat{K}_N = \hat{K}$

Aim

to devise a lossless encoding of $\hat{K}_N$, where $\hat{K}_N$ is an approximation of $\hat{K}$ which improves with increasing N

Partition the grid

- $\Delta F_N := F_N^{(ext)} \setminus F_{N-1}^{(ext)}$, all reduced fractions of the form $\pm\frac{p}{N}, \pm\frac{N}{p}$, where $1 \leq p \leq N$. For $i \neq j$, $\Delta F_i \bigcap \Delta F_j = \emptyset$.
- Let $\Delta\mathcal{G}_N := \mathcal{G}_N \setminus \mathcal{G}_{N-1}$, $N \geq 1$, where $\mathcal{G}_0 := \emptyset$, then $\Delta\mathcal{G}_N = \left\{F_N^{(ext)} \times \Delta F_N\right\} \bigcup \left\{\Delta F_N \times F_{N-1}^{(ext)}\right\}$.
- Partition $\mathcal{G}_N = \bigcup_{i=1}^{N} \Delta\mathcal{G}_i$



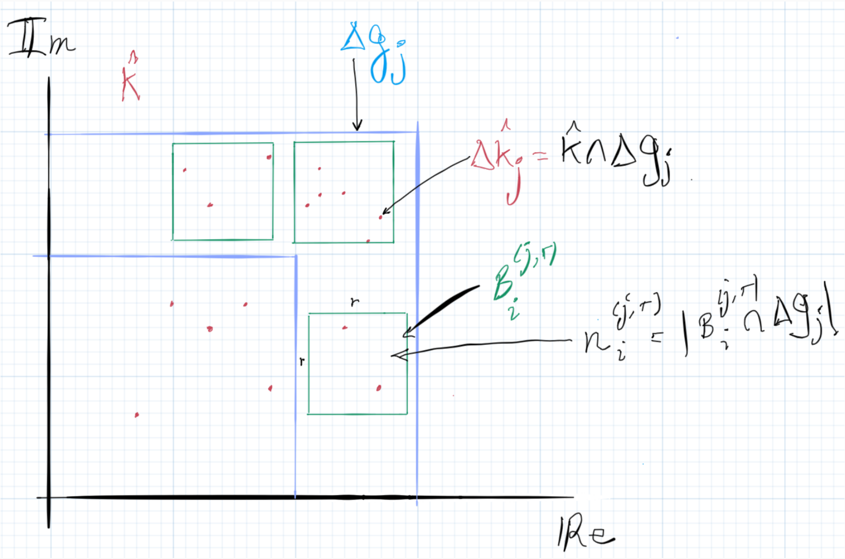
In reality, $F_{N-1}^{(ext)}$ and ΔF_N have also negative elements and their elements interleave

Cover

- Let $\Delta\hat{K}_j := \hat{K} \bigcap \Delta\mathcal{G}_j$, $1 \le j \le N$,
- then, $\hat{K}_N = \bigcup_{j=1}^{N} \Delta\hat{K}_j$, because $\hat{K} \cap \mathcal{G}_N = \bigcup_{j=1}^{N} \left(\hat{K} \cap \Delta\mathcal{G}_j\right)$
- To encode $\hat{K}_N$ we encode each of the sets $\Delta\hat{K}_j$ separately
- $\mathcal{C}(\Delta\hat{K}_j; r)$, minimal-size cover of $\Delta\hat{K}_j$ by square r-boxes $B_r^{(j)}(z) \subset \Delta\mathcal{G}_j$ of side-length r, with top left vertex $z \in \Delta\mathcal{G}_j$
- $\mathcal{N}(\Delta\hat{K}_j; r) := |\mathcal{C}(\Delta\hat{K}_j; r)|$ is r^{th} covering number of $\Delta\hat{K}_j$
- Calculating $\mathcal{N}(\Delta\hat{K}_j; r)$ is NP-hard
- polynomial time algorithm (Chvatal, 1979) computes a cover $\hat{\mathcal{C}}(\Delta\hat{K}_j; r) := \left\{B_1^{(j,r)}, \ldots, B_{\nu_{j,r}}^{(j,r)}\right\}$ (not necessarily minimal size) by r-boxes $B_i^{(j,r)}$ with top-left vertex z_i, such that its cardinality $\nu_{j,r} \le \mathcal{N}(\Delta\hat{K}_j; r)\left(\ln\left|\Delta(\hat{K}_j)\right| + 1\right)$.

We now use this to encode $\hat{K}_N$ by a binary sequence

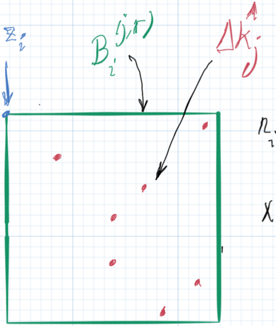
Cover



Uncompressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$

- $n_i^{(j,r)} := \left| B_i^{(j,r)} \bigcap \Delta \mathcal{G}_j \right|$, for $1 \leq i \leq \nu_{j,r}$
- $x_i^{(j,r)} := s(z_i) s\left(x^{(n_i^{(j,r)})} \right)$, is uncompressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$, where $x^{(n_i^{(j,r)})}$ is an $n_i^{(j,r)}$-bit string of indicators of the set $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$ over the elements of the set $B_i^{(j,r)} \bigcap \Delta \mathcal{G}_j$.

Uncompressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$



$n_i^{(j,r)}$ = # grid points

$$x^{(n_i^{(j,r)})} = \underbrace{000 \cdot 010010 \cdot 0100 \cdot 010 \cdot 010 \cdot 010 \cdot 010 \cdot 01}_{n_i^{(j,r)}}$$

Uncompressed Encoding of $B_i^{(j,r)} \cap \Delta \hat{K}_j$:

$$x_i^{(j,r)} = s(z_i)\, S\left(x^{(n_i^{(j,r)})}\right)$$

Uncompressed encoding of $\hat{K}_N$

- $b^{(j,r)} := s\,(\nu_{j,r})\,x_1^{(j,r)} x_2^{(j,r)} \cdots x_{\nu_{j,r}}^{(j,r)}$ is uncompressed encoding of $\Delta \hat{K}_j$, where $s(\nu_{j,r})$ encodes the block length $\nu_{j,r}$

Uncompressed encoding of $\hat{K}_N$

binary sequence $\xi^{(N,r)} := b^{(1,r)}\, b^{(2,r)} \ldots b^{(N,r)}$

Uncompressed description length of $\hat{K}_N$

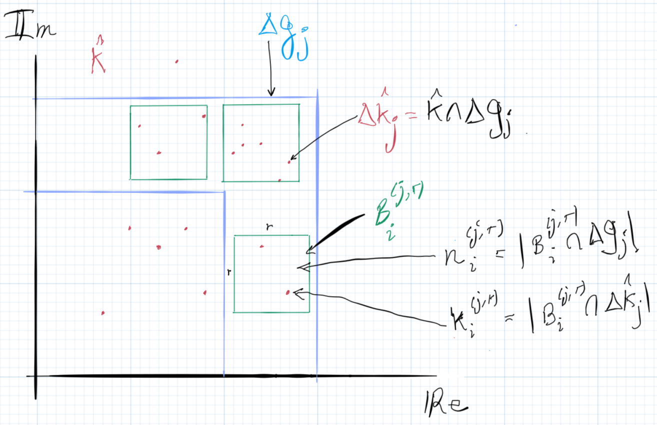
- Uncompressed description length of $\Delta\hat{K}_j$,
  $\Lambda\left(\Delta\hat{K}_j; r\right) := \ell\left(b^{(j,r)}\right) = \ell\left(\nu_{j,r}\right) + \sum_{i=1}^{\nu_{j,r}}\left(4\ell(j) + 2n_i^{(j,r)} + 1\right)$,
  where $4\ell(j)$ is length of codeword describing $z_i \in \Delta\mathcal{G}_j$

Uncompressed description length of $\hat{K}_N$

$\Lambda\left(\hat{K}_N; r\right) := \ell\left(\xi^{(N,r)}\right) = \sum_{j=1}^{N}\Lambda\left(\Delta\hat{K}_j; r\right)$.
It is number of description bits of the set $\hat{K}_N$.

Compressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$

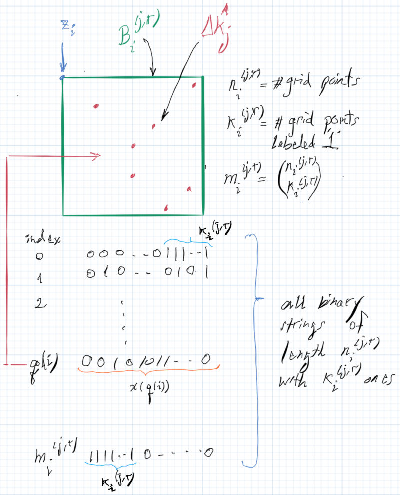
- $n_i^{(j,r)} := \left| B_i^{(j,r)} \bigcap \Delta \mathcal{G}_j \right|$
- $k_i^{(j,r)} := \left| B_i^{(j,r)} \bigcap \Delta \hat{K}_j \right|$



Compressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$

- $m_i^{(j,r)} := \binom{n_i^{(j,r)}}{k_i^{(j,r)}}$
- $q(i)$ the index of indicator sequence that specifies elements of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$
- $y_i^{(j,r)} := s(z_i)s(n_i^{(j,r)})s(k_i^{(j,r)})s(q(i))$, is compressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$,

Compressed encoding of $B_i^{(j,r)} \bigcap \Delta \hat{K}_j$



Compressed Encoding of $B_i^{(j,r)} \cap \Delta \hat{K}_j$:

$$y_i^{(j,r)} = s(z_i)\, s(n_i^{(j,r)})\, s(k_i^{(j,r)})\, s(q(i))$$

Compressed encoding of $\hat{K}_N$

- $w^{(j,r)} := s(\nu_{j,r})\, y_1^{(j,r)} y_2^{(j,r)} \cdots y_{\nu_{j,r}}^{(j,r)}$ is compressed encoding of $\Delta \hat{K}_j$

Compressed encoding of $\hat{K}_N$

binary sequence $\zeta^{(N,r)} := w^{(1,r)} w^{(2,r)} \ldots w^{(N,r)}$

Compressed description length of $\hat{K}_N$

- Compressed description length of $\Delta\hat{K}_j$,

$$\lambda\left(\Delta\hat{K}_j; r\right) := \ell\left(\left(w^{(j,r)}\right)\right.$$
$$= \ell\left(\nu_{j,r}\right) + \sum_{i=1}^{\nu_{j,r}}\left(4\ell(j) + \ell\left(n_i^{(j,r)}\right) + \ell\left(k_i^{(j,r)}\right) + \ell\left(m_i^{(j,r)}\right)\right)$$

Compressed description length of $\hat{K}_N$

$\lambda\left(\hat{K}_N; r\right) := \ell\left(\left(\zeta^{(N,r)}\right)\right. = \sum_{j=1}^{N}\lambda\left(\Delta\hat{K}_j; r\right)$.
It is number of information bits of the set $\hat{K}_N$.

Compression ratio of $\hat{K}_N$

- $(N, r)^{th}$ compression ratio of $\hat{K}_N$

$$\rho\left(\hat{K}_N; r\right) := \frac{\lambda(\hat{K}_N; r)}{\Lambda\left(\hat{K}_N; r\right)}$$

Compression ratio

Number of information bits per description bit.

The more complex the set $\hat{K}_N$, the higher the compression ratio.

Compression ratio of $\hat{K}_N$, as $N \to \infty$

- Uncompressed encoding of $\hat{K}$,

$$\xi^{(r)}\left(\hat{K}\right) := \lim_{N\to\infty} \xi^{(N,r)} = \left\{b^{(j,r)}\right\}_{j=1}^{\infty} \quad \text{(see)}$$

- Compressed encoding of the fractal set $\hat{K}$ sequence,

$$\zeta^{(r)}\left(\hat{K}\right) := \lim_{N\to\infty} \zeta^{(N,r)} = \left\{w^{(j,r)}\right\}_{j=1}^{\infty} \quad \text{(see)}$$

- limit of compression ratio,

$$\lim_{N\to\infty} \rho(\hat{K}_N; r) = \lim_{N\to\infty} \frac{\ell(\zeta^{N,r})}{\ell(\xi^{N,r})}$$

Information density of $\hat{K}$

Information density of $\hat{K}$

$\mathcal{I}\left(\hat{K}\right) := \limsup_{N\to\infty} \inf_{r>0} \rho\left(\hat{K}_N; r\right)$

It measures the minimum number of information bits per description bit, over all values of box side length $r > 0$

- $\mathcal{I}\left(\hat{K}\right)$ exists and is finite

Boundary compression ratio of $\hat{K}_N$

- a set is complex due to its boundary
- As a variant, consider the description length of its boundary
- $\partial \Delta \hat{K}_j$ boundary of $\Delta \hat{K}_j$, the set of points in $\Delta \hat{K}_j$ that have at least one adjacent point $z \in \Delta \mathcal{G}_j$ such that $z \notin \Delta \hat{K}_j$
- boundary of $\hat{K}_N$, $\partial \hat{K}_N := \bigcup_{j=1}^{N} \partial \Delta \hat{K}_j$
- $(N, r)^{th}$ boundary compression ratio of $\hat{K}_N$

$$\rho\left(\partial \hat{K}_N; r\right) := \frac{\lambda(\partial \hat{K}_N; r)}{\Lambda\left(\partial \hat{K}_N; r\right)}$$

Boundary compression ratio

Number of information bits per description bit.
The more complex the border of $\hat{K}_N$, the higher the boundary compression ratio.

r^{th} boundary compression ratio of $\hat{K}$

- Let $\partial b^{(j,r)}$ and $\partial w^{(k,r)}$ be defined as in $b^{(j,r)}$ and $w^{(j,r)}$, except, they are based on the cover of $\partial \Delta \hat{K}_j$
- uncompressed and compressed description lengths of $\partial \Delta \hat{K}_j$ are defined as $\Lambda\left(\partial \Delta \hat{K}_j; r\right) = \ell\left(\partial b^{(j,r)}\right)$ and $\lambda\left(\partial \Delta \hat{K}_j; r\right) = \ell\left(\partial w^{(j,r)}\right)$
- instead of summing over j (as here and here), define non-cumulative boundary compression ratio (easier to compute numerically)

$$\rho_{NC}(\hat{K}; j, r) := \frac{\lambda\left(\partial \Delta \hat{K}_j; r\right)}{\Lambda\left(\partial \Delta \hat{K}_j; r\right)}$$

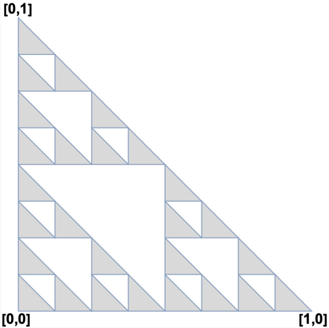
- if $\lim_{j\to\infty} \rho_{NC}(\hat{K}; j, r)$ exists, it equals r^{th} boundary compression ratio,

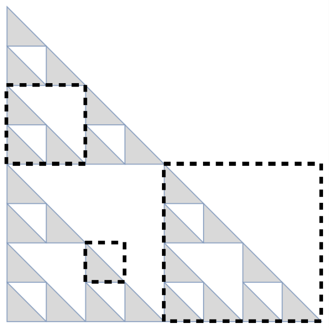
$$\rho\left(\partial \hat{K}; r\right) = \limsup_{N\to\infty} \rho(\partial \hat{K}_N; r)$$

Sierpinski triangle

Sierpinski set $K = \bigcap_{l=1}^{\infty} K^{(l)}$ with finite approximation $K^{(l)}$ (depth level l)

- For instance, $l = 3$:





- Cover using boxes of side length $r = \frac{1}{s}$, where $s = 2^i$, $i \geq 0$

Sierpinski triangle

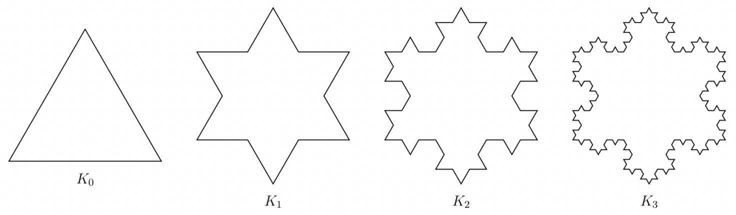
- $(l, s)^{th}$ compression ratio

$$\rho(K^{(l)}; s) \simeq H\left(\frac{s^{\beta}}{2}\left(\frac{3}{4}\right)^{l}\right)$$

  where $H(p) := -p\log p - (1-p)\log(1-p)$ and $\beta = 2 - \log 3 \simeq 0.415$
- fractal dimension (Barnsley, 1993) of K is $\frac{\log 3}{\log 2}$
- information density $\mathcal{I}(K) \simeq 0$

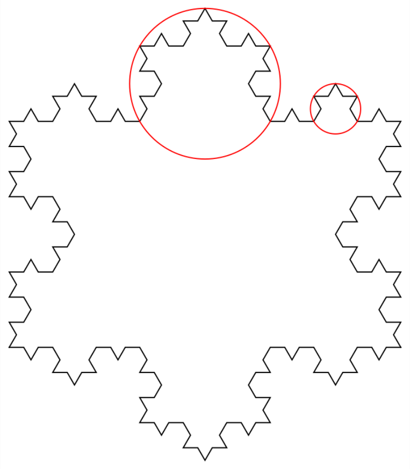
Koch snowflake

filled-Koch snowflake $K = K_0 \bigcup \left(\bigcup_{j=1}^{\infty} \Delta K_j \right)$,
where $\Delta K_l := K_l \setminus K_{l-1}$, with finite approximation K_l (depth level l),



- Side length $s_l = (1/3)^l$, $l \geq 0$

Koch snowflake



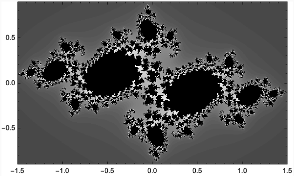
- Cover using balls B_i of radius $r_i = \frac{s_i}{\sqrt{3}}$, $i \geq 0$

Koch snowflake

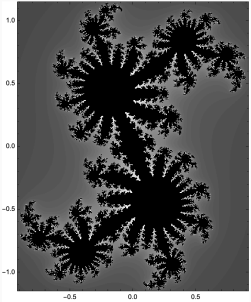
- $(l, i)^{th}$ compression ratio, $\rho(K_l, i)$, some decreasing function with respect to i
- Information density $\mathcal{I}(K) \simeq 0$
- Boundary information density $\mathcal{I}(\partial K) \simeq 0.639$

More sets

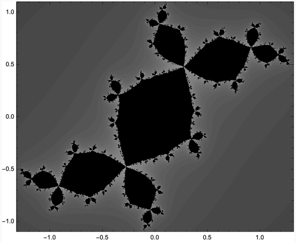
- Evaluate numerically the compression ratio for several sets



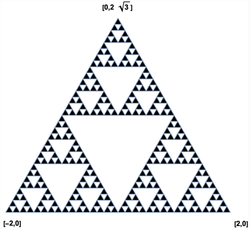
(a) $\kappa = -0.7269 + i0.1889$



(b) $\kappa = 0.377 - i0.248$

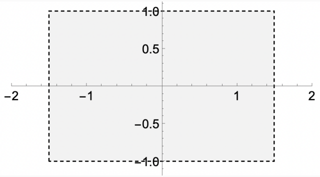


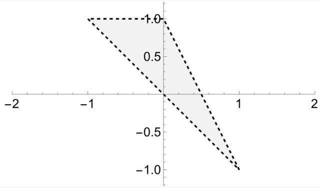
(c) $\kappa = -0.1226 - i0.7449$



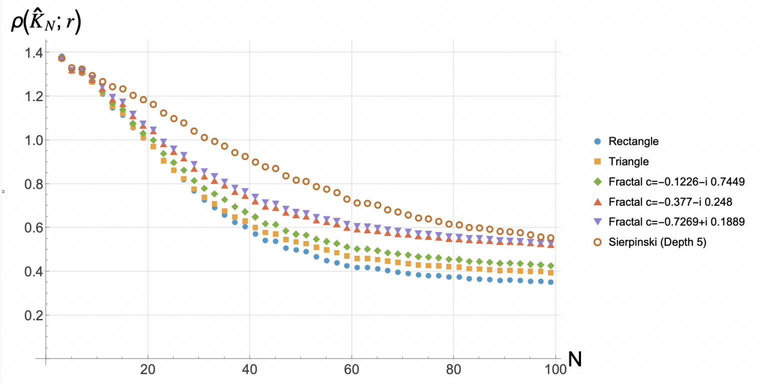
(d) Sierpinski set $S^{(5)}$

More sets





Comparing



r^{th} boundary compression ratio	Rectangle	Triangle	(c)	(b)	(a)	(d)
$\rho\left(\partial\hat{K};r\right)$	0.176	0.204	0.305	0.411	0.415	0.464

Conclusions

- We introduce information-based measures of complexity of fractal sets (extend classical notion of compression ratio of binary sequences (Ziv & Lempel, 1978))
- different fractal sets have different non-zero compression ratio values (although their algorithmic complexity is essentially zero)
- compression ratios are useful as a complexity measure for fractal sets
- straightforward to estimate (r^{th} boundary compression ratio)

Thanks for your attention.

Appendix

Hausdorff dimension

For $s \geq 0$ and $\delta > 0$, define

$$\mathcal{H}^s_\delta(A) := \inf \left\{ \sum_i \big(\operatorname{diam} U_i\big)^s : A \subseteq \bigcup_i U_i,\ \operatorname{diam}(U_i) \leq \delta \right\}$$

where inf is over all countable covers of A, and U_i are arbitrary subsets of X used to cover A.

The s-dimensional Hausdorff measure is

$$\mathcal{H}^s(A) := \lim_{\delta \downarrow 0} \mathcal{H}^s_\delta(A).$$

The Hausdorff dimension of A is

$$\dim_{\mathrm{H}} A := \inf\{s : \mathcal{H}^s(A) = 0\} = \sup\{s : \mathcal{H}^s(A) = \infty\}.$$