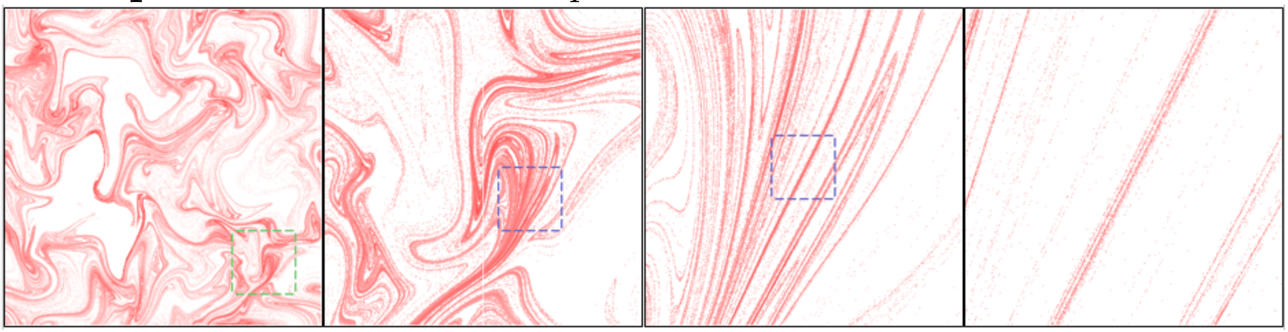
12 Fractals and fractal dimensions

As seen in the last lecture, dissipative chaotic systems may have a strange attractor, i.e. a minimal, attracting, invariant set that is aperiodic with chaotic dynamics. Strange attractors show self-similar structure at arbitrary small scales. It is common to quantify scale-invariant structures using some fractal (non-integer) dimension D (for example D_q below).

Example Successive zoom of particles on surface of turbulent flow:



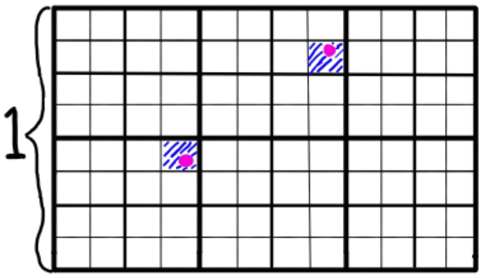
In general D must be calculated numerically. As seen in Lecture 11, strange attractors form in a stretching+folding process and show small-scale filamentary structure. A simple estimate of D of a strange attractor is unity (for the dimension along the filaments) plus the fractal dimension of the pattern on a cross section orthogonal to the filaments. Some mathematical idealisations of such cross-section patterns (e.g. Cantor sets) have analytical expressions for the fractal dimension.

12.1 Box-counting dimension

One way to define the fractal dimension D is the box-counting dimension. It describes how space-filling a geometrical object is.

12.1.1 Regular shapes

Consider three sets in $d = 2$ dimensions.
Cover space by d-dimensional boxes with side length $\epsilon = 2^{-k}$.



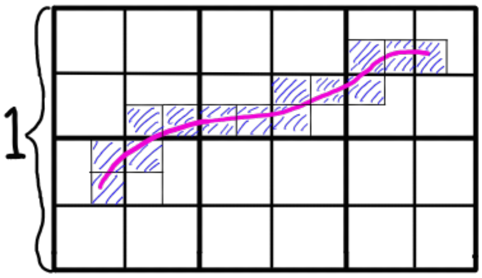
Two points

$N_{\rm box}(\frac{1}{2}) = 2 \qquad N_{\rm box}(\epsilon) = 2$

$N_{\rm box}(\frac{1}{4}) = 2 \qquad$ independent of ϵ

$N_{\rm box}(\frac{1}{8}) = 2 \qquad D_0 = 0$

$\vdots$



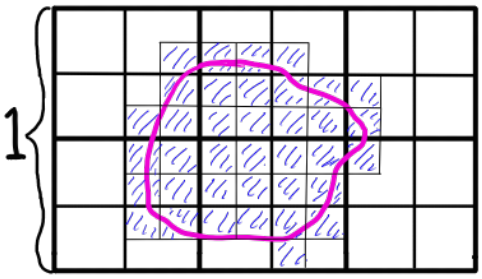
Curve of length L

$N_{\rm box}(\frac{1}{2}) = 4 \qquad N_{\rm box}(\epsilon) \sim \frac{L}{\epsilon}$

$N_{\rm box}(\frac{1}{4}) = 8 \qquad$ for small ϵ

$N_{\rm box}(\frac{1}{8}) = 14 \qquad D_0 = 1$

$\vdots$



Area A

$N_{\rm box}(\frac{1}{2}) = 6 \qquad N_{\rm box}(\epsilon) \sim \frac{A}{\epsilon^2}$

$N_{\rm box}(\frac{1}{4}) = 14 \qquad$ for small ϵ

$N_{\rm box}(\frac{1}{8}) = 37 \qquad D_0 = 2$

$\vdots$

Let $N_{\rm box}(\epsilon)$ = number of boxes with side length ϵ required to cover the set. Then

$$N_{\rm box}(\epsilon) \sim A_0 \epsilon^{-D_0} \text{ for small } \epsilon$$

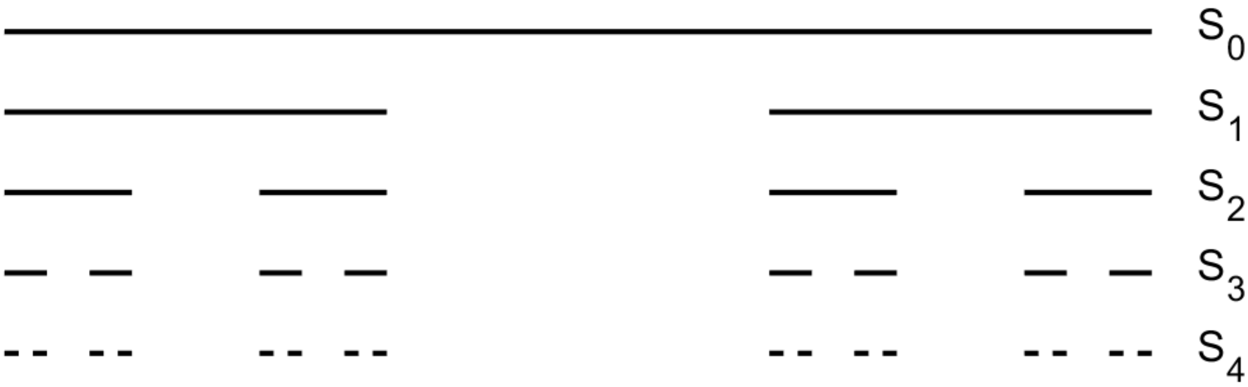
where A_0 is some constant and D_0 is the dimension of the object. Equivalently

$$\begin{aligned} &\ln N_{\rm box}(\epsilon) \sim \ln A_0 + D_0 \ln(1/\epsilon) \\ &\Rightarrow D_0 = \lim_{\epsilon\to 0} \frac{\ln N_{\rm box}(\epsilon) - \ln A_0}{\ln(1/\epsilon)} = \lim_{\epsilon\to 0} \frac{\ln N_{\rm box}(\epsilon)}{\ln(1/\epsilon)} \end{aligned} \tag{1}$$

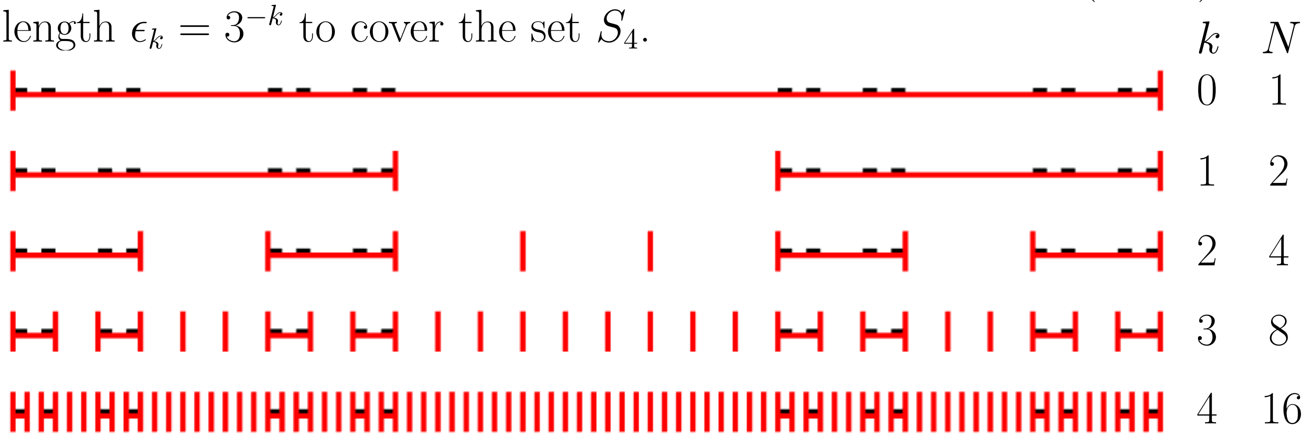
D_0 is the box-counting dimension of a set.
For the three cases above D_0 equals the dimensionality of the sets.

12.1.2 Example: Middle-third Cantor set

Construction by successive generations S_n:



Iterate to get generation n. The Cantor set is the geometrical object with generation $n = \infty$. The total length of generation n is $(2/3)^n$, which goes to zero as $n \to \infty$, the set has zero (Lebesgue) measure. At the same time, it is possible to show that the set consists of an uncountably infinite amount of points (Strogatz 11.2).
Now calculate the box-counting dimension using boxes (sticks) of length $\epsilon_k = 3^{-k}$ to cover the set S_4.



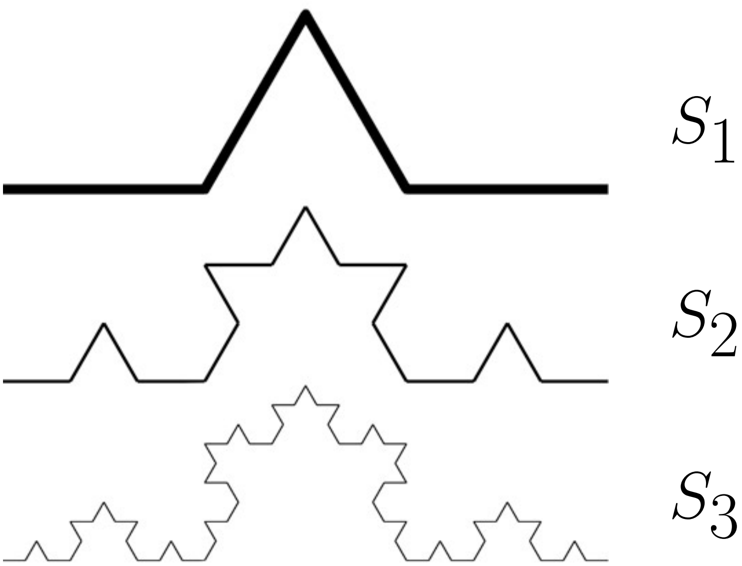
The number of boxes needed to cover the set S_4 is: $N_{\text{box}}(\epsilon_k) = 2^k$. Note that S_4 has a finite resolution 2^{-4} (the lengths of connected intervals). Covering S_4 with smaller boxes, $k > 4$, would yield the wrong result. For general k, we need to at least iterate to generation S_k, and we find $N_{\text{box}}(\epsilon_k) = 2^k$ for all k. The box-counting dimension becomes

$$D_0 = \lim_{\epsilon \to 0} \frac{\ln N_{\text{box}}(\epsilon)}{\ln(1/\epsilon)} = \lim_{k\to\infty} \frac{\ln 2^k}{\ln 3^k} = \frac{\ln 2}{\ln 3} \approx 0.6309$$

The dimension lies somewhere between 0 and 1 (fatter than points, thinner than a line). This result is independent of the choice of gridding ($\epsilon = 3^{-k}$), other choices give the same dimension. A set with non-integer dimension is called a fractal. The cantor set is self-similar: magnification of small pieces reproduces the original set exactly. In general fractals are scale-invariant: successive magnifications reveal more and more structure that is statistically self-similar without resolution limit.

Fractals (but with some small-scale limit) are common in nature: Brownian motion, particle aggregation, networks of rivers and tree branches, coastlines, etc. and they are also common in scale-free networks: WWW, social networks, human brain, metabolic network, protein interaction network, etc. Fractals have applications in for example data compression, pattern generation, CGI, and fractal antenna.

12.1.3 Example: Koch curve



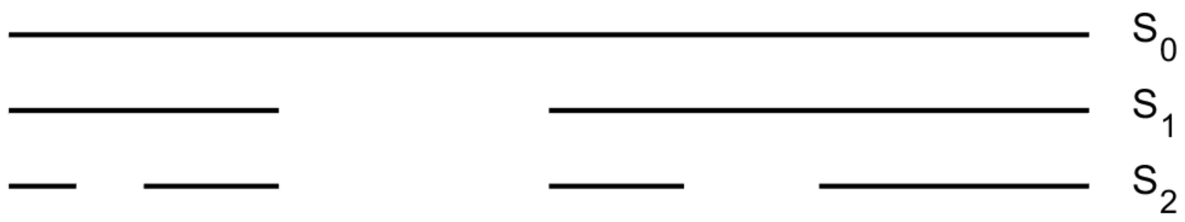
Length of curve: $(4/3)^n \to \infty$ as $n \to \infty$ (infinite length between any two points on curve).

Use boxes of length $\epsilon_k = 3^{-k} \Rightarrow N_{\text{box}}(\epsilon) = 4^k$. Dimension

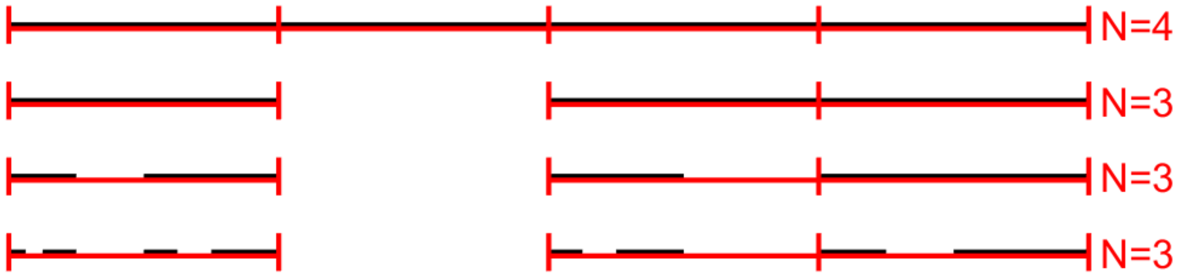
$$D_0 = \lim_{k\to\infty} \frac{\ln 4^k}{\ln 3^k} = \frac{\ln 4}{\ln 3} \approx 1.2619$$

12.1.4 Example: Asymmetric Cantor set

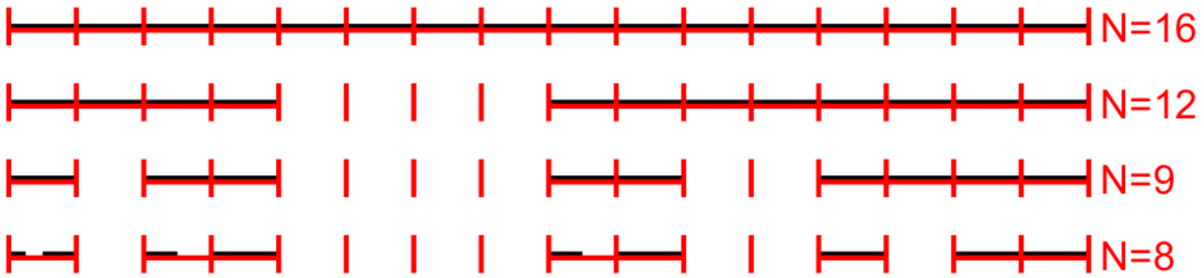
Construct Cantor set by removing second quarter (instead of mid third)



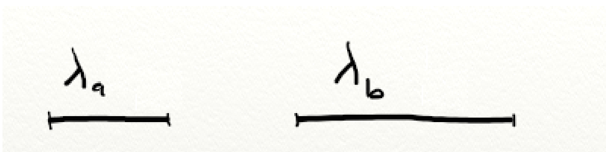
Use $\epsilon_k = 4^{-k}$. From the symmetric Cantor set one could assume $N_{\rm box}(\epsilon_k) = 3^k$, BUT we need to iterate until largest connected interval is smaller than $\epsilon_k = 4^{-k}$ (the set is defined as $n \to \infty$). To show that $N_{\rm box}(\epsilon_k) \neq 3^k$, cover successive generations of the set.
Case $\epsilon_1 = 4^{-1}$:



$\Rightarrow N_{\rm box}(\epsilon = 1/4) = 3$.
Case $\epsilon_2 = 4^{-2}$:



$\Rightarrow N_{\rm box}(\epsilon = 1/4^2) = 8$, i.e. $N_{\rm box}(\epsilon = 1/4^2) \neq 3^2 = 9$.
In conclusion we must iterate to a generation $n > k$ for the asymmetric Cantor set (in previous examples it was enough to iterate to generation $n = k$). The reason is that the asymmetric Cantor set is constructed from two relative length scales $\lambda_a = 1/4$ and $\lambda_b = 1/2$



Let

$$
\begin{aligned}
N_{\text{box}}(\epsilon) &= N_a(\epsilon) + N_b(\epsilon)\\
N_a(\epsilon) &= \text{Number of boxes of size } \epsilon \text{ needed to cover } \lambda_a\\
N_b(\epsilon) &= \text{Number of boxes of size } \epsilon \text{ needed to cover } \lambda_b
\end{aligned}
$$

Self-similarity $\Rightarrow$

$$
\begin{aligned}
N(\epsilon/\lambda_a) &= N_a(\epsilon)\\
N(\epsilon/\lambda_b) &= N_b(\epsilon)
\end{aligned}
$$

- 'To cover strip λ_a with boxes of size ϵ is like covering full unit interval with boxes of size ϵ/λ_a'
D_0 is defined as the scaling exponent $N(\epsilon) = A\epsilon^{-D_0}$. $\Rightarrow$

$$
\begin{aligned}
&\underbrace{N(\epsilon)}_{A\epsilon^{-D_0}} = \underbrace{N_a(\epsilon)}_{N(\epsilon/\lambda_a)} + \underbrace{N_b(\epsilon)}_{N(\epsilon/\lambda_b)}\\
&\Rightarrow A\epsilon^{-D_0} = A(\epsilon/\lambda_a)^{-D_0} + A(\epsilon/\lambda_b)^{-D_0}\\
&\Rightarrow 1 = (1/\lambda_a)^{-D_0} + (1/\lambda_b)^{-D_0}\\
&\Rightarrow 1 = \lambda_a^{D_0} + \lambda_b^{D_0}
\end{aligned}
$$

Here $\lambda_a = 1/4$ and $\lambda_b = 1/2$

$$
\begin{aligned}
&1 = \left(\frac{1}{4}\right)^{D_0} + \left(\frac{1}{2}\right)^{D_0} = \left(\frac{1}{2^{D_0}}\right)^2 + \frac{1}{2^{D_0}}\\
&[\text{Quadratic equation in } x = 1/2^{D_0}]\\
&\frac{1}{2^{D_0}} = \frac{-1 \overset{+}{\scriptstyle(-)} \sqrt{5}}{2}\\
&D_0 = -\frac{\ln\left(\frac{-1+\sqrt{5}}{2}\right)}{\ln 2} \approx 0.69
\end{aligned}
$$

As comparison, the naive result is $\ln 3/\ln 4 \approx 0.79$. Note that the symmetric case, $\lambda_a = \lambda_b = 1/3$, gives the middle-third result as before:

$$1 = (1/3)^{D_0} + (1/3)^{D_0} = 2(1/3)^{D_0}$$
$$\Rightarrow D_0 = \frac{\ln(1/2)}{\ln(1/3)} = \frac{\ln 2}{\ln 3}\,.$$

12.1.5 Numerical evaluation of D_0

In order to numerically evaluate D_0 for a set of numerically evaluated points on a fractal, it is not enough to simply evaluate Eq. (1) for a small value of ϵ:

- We do not know if we have enough points in our data to resolve the chosen value of ϵ.

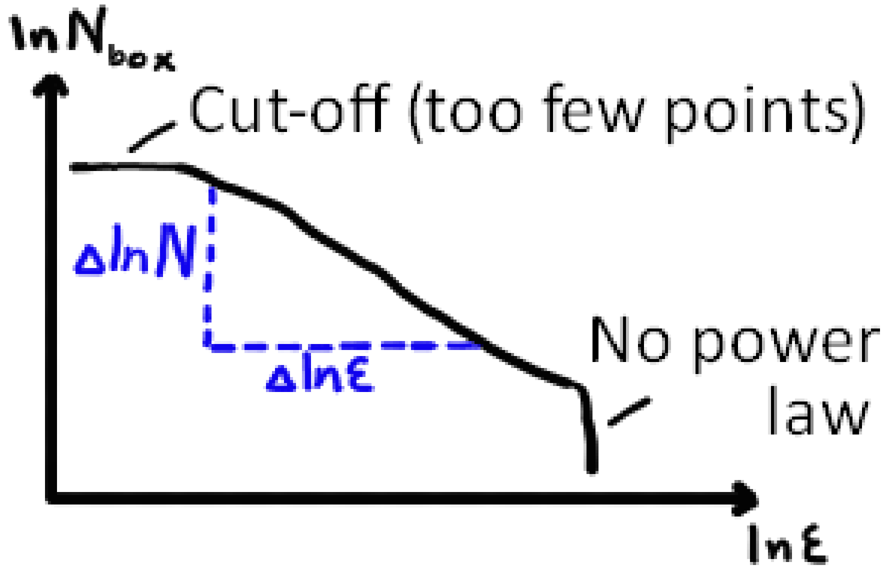
- Even if ϵ is small, the scaling law has an unknown coefficient:

$$N_{\rm box}(\epsilon) = \underbrace{A_0}_{\text{Unknown}} \epsilon^{-D_0}$$
$$\Rightarrow \ln N_{\rm box}(\epsilon) = \ln A_0 + D_0 \ln(1/\epsilon) \to D_0 \ln(1/\epsilon) \text{ as } \epsilon \to 0$$

  The limit $\epsilon \to 0$ is slow because of the logarithm. Since we do not know beforehand how large A_0 is, we do not know how small ϵ needs to be to be able to neglect the contribution $\ln A_0$.

Instead, use slope of curve $\ln N_{\rm box}(\epsilon)$ against $\ln \epsilon$ (A_0 drops out):

$$D_0 = -\frac{\Delta \ln N_{\rm box}(\epsilon)}{\Delta \ln \epsilon}$$



12.1.6 Additional comments

- Box-counting dimension quantifies how space-filling a fractal is
- Straightforward implementation, but limited by memory to store number of visits to each box.
- A minimal cover is in general hard to find, but a uniform grid gives the same dimension

12.2 Generalized dimension sprectrum

The box-counting dimension weighs all boxes equal. This is fine for ideal mathematical fractals (Cantor set, Koch curve, etc.). But many strange attractors show fluctuations in the occupation number of the boxes: often a small percentage of boxes are visited frequently, while most boxes are only sparsely visited. Depending on our purpose we may be interested in boxes where trajectories spend more or less time. We therefore introduce a generalized dimension D_q which weigh boxes differently depending on how large fraction of points they contain.

Let N_{point} be a large number of points, ideally $N_{\text{point}} \to \infty$, on a fractal set or fractal attractor. Label ϵ-sized boxes that contain at least one point by $k = 1, \ldots, N_{\rm box}(\epsilon)$. Let $N_k(\epsilon)$ be the number of points in (point-containing) box k. Let $p_k(\epsilon) = N_k(\epsilon)/N_{\rm point}$ be the fraction of points in box k. Check normalisation:

$$\sum_{k=1}^{N_{\rm box}} p_k = \frac{1}{N_{\rm point}} \sum_{k=1}^{N_{\rm box}} N_k = \frac{1}{N_{\rm point}} N_{\rm point} = 1\,.$$

Generalized dimension D_q (Rényi dimension spectrum)

$$D_q \equiv \frac{1}{1-q} \lim_{\epsilon\to 0} \frac{\ln I(q,\epsilon)}{\ln(1/\epsilon)} \tag{2}$$

with a real parameter q and

$$I(q,\epsilon) = \sum_{k=1}^{N_{\rm box}} p_k^q(\epsilon)\,.$$

The significance of q can be summarized as:

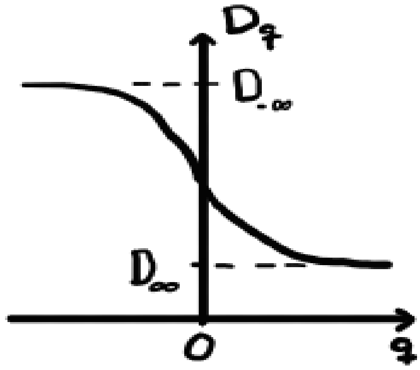
- When $q = 0$ density variations are neglected. In this limit we recover the box-counting dimension (Eq. (1)):

$$D_0 = \frac{1}{1} \lim_{\epsilon\to 0} \frac{\ln I(0,\epsilon)}{\ln(1/\epsilon)} = \lim_{\epsilon\to 0} \frac{\ln\left(\sum_{k=1}^{N_{\rm box}} p_k(\epsilon)^0\right)}{\ln(1/\epsilon)} = \lim_{\epsilon\to 0} \frac{\ln(N_{\rm box})}{\ln(1/\epsilon)}\,.$$

- If $q > 0$ contributions to $I(q,\epsilon)$ from regions of high density on the attractor are amplified compared to low-density regions. D_q with large q therefore characterises clustering of high-density regions.

- When $q < 0$ the opposite is true: low-density regions dominate contributions to $I(q,\epsilon)$ and D_q

It is possible to show that D_q is a non-growing function of q, i.e. $D_q \geq D_{q'}$ if $q < q'$. Typical picture:



The difference between $D_{-\infty}$ and D_∞ determines how large the variations of the fractal dimension are on the strange attractor.

- If $D_{-\infty} = D_\infty$, i.e. a flat D_q spectrum, we have a monofractal.
- If D_q is not constant with respect to q, we have a multifractal, i.e. points are non-uniformly distributed on the fractal.

Monofractal Every point on attractor is equally likely $\Rightarrow D_q = D_0 = \text{const}$. Check:

$$
\begin{aligned}
D_q &= \frac{1}{1-q}\lim_{\epsilon\to 0}\frac{\ln\left(\sum_{k=1}^{N_\text{box}} p_k^q\right)}{\ln(1/\epsilon)} \\
&= \left[\text{`Every point equally likely'} \Rightarrow \underbrace{p_k \sim 1/N_\text{box}}_{\text{if } \epsilon\to 0} \quad \text{and} \quad \textstyle\sum_{k=1}^{N_\text{box}} = N_\text{box}\right] \\
&= \frac{1}{1-q}\lim_{\epsilon\to 0}\frac{\ln(N_{box}(1/N_\text{box})^q)}{\ln(1/\epsilon)} \\
&= \frac{1}{1-q}\lim_{\epsilon\to 0}\frac{\ln(N_{box}^{1-q})}{\ln(1/\epsilon)} = \lim_{\epsilon\to 0}\frac{\ln(N_{box})}{\ln(1/\epsilon)} = D_0
\end{aligned}
$$

12.3 Information dimension

Special case $q = 1$ in Eq. (2): Expression for D_1 diverges $\Rightarrow$ take limit. First expand $\ln I(q,\epsilon)$ around $q = 1$

$$\ln I(q,\epsilon) = \ln\left(\sum_{k=1}^{N_{\rm box}} p_k^q\right)$$

$$\left[\text{Use: } p_k^q = p_k p_k^{q-1} = p_k \exp((q-1)\ln p_k) \approx p_k(1+(q-1)\ln p_k) \text{ for } q\approx 1\right]$$

$$= \ln\left(\sum_{k=1}^{N_{\rm box}} p_k + \sum_{k=1}^{N_{\rm box}} p_k(q-1)\ln p_k)\right)$$

$$\left[\text{Use norm: } \sum_{k=1}^{N_{\rm box}} p_k = 1 \text{ and expand } \ln(1+(q-1)A) \approx (q-1)A \text{ for } q\approx 1\right]$$

$$= (q-1)\sum_{k=1}^{N_{\rm box}} p_k \ln p_k\,.$$

We get

$$\begin{aligned}\lim_{q\to 1} D_q &= \lim_{q\to 1}\lim_{\epsilon\to 0}\frac{1}{1-q}\frac{\ln\left(I(q,\epsilon)\right)}{\ln(1/\epsilon)}\\ &= \lim_{q\to 1}\lim_{\epsilon\to 0}\frac{1}{1-q}\frac{(q-1)\sum_{k=1}^{N_{\rm box}} p_k\ln p_k}{\ln(1/\epsilon)}\\ &= \lim_{\epsilon\to 0}\frac{\sum_{k=1}^{N_{\rm box}} p_k\ln p_k}{\ln(\epsilon)}\,.\end{aligned}$$

D_1 is referred to as the information dimension.

12.3.1 Shannon entropy

The quantity $S = -\sum_k p_k(\epsilon)\ln p_k(\epsilon) = -\langle \ln p(\epsilon)\rangle$ is the Shannon entropy. Given an experiment with r possible outcomes, with probabilities $p_1, p_2, \ldots, p_r$, $\sum_{k=1}^{r} p_k = 1$, the Shannon entropy is

$$S = -\sum_{k=1}^{r} p_k \ln p_k \, ,$$

with $p_k \ln p_k = 0$ if $p_k = 0$.
Simplest case: $p_i = 1$ if $k = i$ and 0 otherwise $\Rightarrow S = 0$.
Case of maximum uncertainty: $p_k = 1/r \Rightarrow S = \ln r$.
In general $0 \leq S \leq \ln r$, the closer S is to $\ln r$ the more uncertain the outcome of an experiment is. For the information dimension: $r = N_{\text{box}}$ and S quantifies the amount of (Shannon) information needed to describe the content of the boxes for a given accuracy ϵ. The information dimension tells how this amount of information scales with resolution ϵ, $S(\epsilon) \sim -D_1 \ln \epsilon$.

12.3.2 Kaplan-Yorke conjecture

It can be shown that almost generally the Lyapunov dimension D_{L} (Lecture 11) is equal to D_1 in Eq. (2), $D_{\text{L}} = D_1$ (but it is possible to construct counterexamples).

12.4 Correlation dimension

Special case $q = 2$ is called the 'correlation dimension'. Rewrite

$$
\begin{aligned}
I(q=2,\epsilon) &= \sum_{k=1}^{N_{\text{box}}} p_k^2 = \sum_{k=1}^{N_{\text{box}}} \left(\frac{N_k}{N_{\text{point}}}\right)^2 = \frac{1}{N_{\text{point}}^2}\sum_{k=1}^{N_{\text{box}}} N_k^2 \\
&= \frac{1}{N_{\text{point}}^2}\sum_{k=1}^{N_{\text{box}}} N_k \sum_{\alpha=1}^{N_{\text{point}}} \begin{cases} 1 & \text{if } \boldsymbol{x}_\alpha \in k\text{:th box} \\ 0 & \text{otherwise} \end{cases} \\
&= \frac{1}{N_{\text{point}}^2}\sum_{k=1}^{N_{\text{box}}} \sum_{\beta=1}^{N_{\text{point}}} \sum_{\alpha=1}^{N_{\text{point}}} \begin{cases} 1 & \text{if } \boldsymbol{x}_\alpha \,\&\, \boldsymbol{x}_\beta \in k\text{:th box} \\ 0 & \text{otherwise} \end{cases} \\
&\quad [\alpha \text{ and } \beta \text{ runs over all particles} \Rightarrow \text{redundant to sum over boxes}] \\
&= \frac{1}{N_{\text{point}}^2}\sum_{\beta=1}^{N_{\text{point}}} \sum_{\alpha=1}^{N_{\text{point}}} \begin{cases} 1 & \text{if } \boldsymbol{x}_\alpha \,\&\, \boldsymbol{x}_\beta \in \text{same box} \\ 0 & \text{otherwise} \end{cases} \\
&\sim \frac{1}{N_{\text{point}}^2}\sum_{\beta=1}^{N_{\text{point}}} \sum_{\alpha=1}^{N_{\text{point}}} \begin{cases} 1 & \text{if } \boldsymbol{x}_\beta \text{within distance } \epsilon \text{ from } x_\alpha \\ 0 & \text{otherwise} \end{cases} \\
&= \frac{1}{N_{\text{point}}^2}\sum_{\beta=1}^{N_{\text{point}}} \sum_{\alpha=1}^{N_{\text{point}}} \Theta(\epsilon - |\boldsymbol{x}_\alpha - \boldsymbol{x}_\beta|)
\end{aligned}
$$

This is the correlation sum. It describes the probability $P(|\boldsymbol{x}_1 - \boldsymbol{x}_2| < \epsilon)$ to find two points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ on the attractor within distance ϵ. The correlation dimension D_{C} can be defined as

$$
\begin{aligned}
D_2 &= \frac{1}{1-2}\lim_{\epsilon\to 0} \frac{\ln\left(I(q=2,\epsilon)\right)}{\ln(1/\epsilon)} \\
&= \lim_{\epsilon\to 0} \frac{\ln\left(P(|\boldsymbol{x}_1 - \boldsymbol{x}_2| < \epsilon)\right)}{\ln \epsilon} \equiv D_{\text{C}}
\end{aligned}
$$

i.e. D_{C} is defined from the scaling $P(|\boldsymbol{x}_1 - \boldsymbol{x}_2| < \epsilon) \sim \epsilon^{D_{\text{C}}}$.
N.B. D_2 is often easier to evaluate using the correlation sum than from box-counting. It is also what you typically measure in experiments.

Similarly D_n with $n = 2, 3, 4, \ldots$ describes scaling of probability to find n particles within separation ϵ.